SB-QTVW: AN ALTERNATE INDEX TO DIFFERENTIATE BETWEEN HEALTHY AND ACUTE ISCHEMIC/INFARCTED PATIENTS USING VECTOR MAGNITUDE DERIVED QT VARIABILITY

By

RAHIM RUKNUDIN SEWANI

Bachelor of Science in Aerospace Engineering

Iowa State University
Ames, Iowa
1995

Doctor of Osteopathic Medicine
Oklahoma State University
Tulsa, Oklahoma
2020

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of

DOCTOR OF PHILOSOPHY
July, 2021

SB-QTVW: AN ALTERNATE INDEX TO DIFFERENTIATE BETWEEN HEALTHY AND ACUTE ISCHEMIC/INFARCTED PATIENTS USING VECTOR MAGNITUDE DERIVED QT VARIABILITY

Dissertation Approved:

Bruce A. Benjamin, Ph.D.

Dissertation Adviser

Alexander Rouch, Ph.D.

Thomas Curtis, Ph.D.

Randy S. Wymore, Ph.D.

Dursun Delen, Ph.D.

ACKNOWLEDGEMENTS

I wish to take the opportunity to thank all of the faculty and staff at Oklahoma State University (OSU) Center for Health Sciences (CHS) and College of Medicine (COM.) I wish to especially thank my committee members for their untiring efforts and guidance in seeing me through my journey at OSU.

I wish to thank Randy Wymore, Ph.D., for his different perspective at each turn; Thomas Curtis, Ph.D., for his reality perspective to every aspect of research and education; Alexander Rouch, Ph.D., for believing in me, and without whose guidance I would never have been admitted to OSU-CHS (so you can blame him, haha); Dursun Delen, Ph.D. for his extreme patience and guidance; and, last but not least, Bruce Benjamin, Ph.D., without whom I would not have been able to enter this marathon, kept on running, and reached the finish line eventually. Dr. Benjamin truly saw me through this whole process each day at a time.

I also wish to thank Satish Bukkapatnum, Ph.D., and Late Rama Komunduri, Ph.D. for initially being on my committee and guiding me through my initial years of research.

I also wish to thank all my colleagues and fellow students throughout my years at OSU. All of the medical and graduate students taught me and helped me get through almost all my classes, quizzes, exams and whatnot at OSU CHS and COM.

I wish to express my gratitude to Randy Freeman, D.O./M.S.; Thomas Troung D.O.; and Rupesh Agrawal, Ph.D., for helping me with several parts of my research. Without their contribution, I would not have been able to complete my research.

Name: RAHIM RUKNUDIN SEWANI

Date of Degree: JULY, 2021

Title of Study: SB-QTVW: AN ALTERNATE INDEX TO DIFFERENTIATE BETWEEN HEALTHY AND ACUTE ISCHEMIC/INFARCTED PATIENTS USING VECTOR MAGNITUDE DERIVED QT VARIABILITY.

Major Field: BIOMEDICAL SCIENCES

Abstract: Myocardial ischemia alters ventricular repolarization through a variety of mechanisms. Studies show that QTVI, an index of QT variability, is elevated during acute ischemia and cardiomyopathy. This indicates an increased QT variability in ischemic patients. Preliminary studies from our laboratory showed that the length of vector magnitude derived QT interval varies by less than 16 ms in healthy controls, and frequently varies more than 16 ms in ischemic patients. With this study, we introduced an alternative index—QT Variability Window (QTVW), and tested the hypothesis that ischemia causes the length of resting QT intervals in vector magnitude to vary more than a QTVW of 16 ms. Results from our study support our hypothesis. Furthermore, in certain cases, our algorithm was able to differentiate between the healthy controls and acute ischemic patients using as few as 20 consecutive beats. These findings provide an alternative index, QTVW, to assess ventricular repolarization following ischemia.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Cardiovascular disease (CVD) and coronary heart disease (CHD) are chronic health problems in the United States. Nearly 126.9 million Americans had some form of cardiovascular disease between 2015 – 2018 [1]. CVD resulted in over eight hundred thousand deaths in US in the year 2017 [1]. Coronary heart disease is the single leading cause of death and claimed the lives of over 42% of Americans in 2018 [1]. Early detection and treatment of ischemia may help mitigate CVD/CHD and its associated mortality rate. However, improved medical tests to detect early symptoms of ischemia, especially non-symptomatic ischemia, are still in works [2].

STEMI

At the present time, one of the standard criteria for detecting ischemia, using 12-lead EKG, is T-wave inversion. Similarly, the criterion for detecting MI in EKG is evidence of ST segment elevation or depression [2]. The gold standard for MI detection utilizes presence of elevated levels of Troponin I (>0.1 ng/mL, 5-6 hours from onset of chest pain) and/or Troponin T enzyme (>0.2 ng/mL, 8-10 hours from onset of chest pain)

in blood as confirmation of cardiac cell damage [3]. Sensitivity and specificity can be as high as 96% and 80% for Troponin T, and 90% and 95% for Troponin I, respectively [3].

Evidently, all current available diagnostic options provide information after the injury to cardiac muscle cells has occurred. Hence the need for an early ischemia detection test, which is both relatively reliable and inexpensive, cannot be stressed enough.

Alternative Methodologies

Several scientists have grappled with this problem, and have tried to find ways to extract relevant precursor information from the EKG. In that context, several studies have concentrated on using heart rate variability as a key surrogate measure to predict early onset of ischemia. Relatively few studies have concentrated on the QT interval length and its significance to ventricular repolarization as ways to predict cardiac abnormalities. While all such studies and solutions have produced mixed results in terms of sensitivity and specificity, they have helped advance our knowledge and gain more understanding in the area of CVD/CHD [2]. In our lab, we concentrated on QT variability and its behavior following acute ischemia and infarction. We focused on the QT interval because it is a reflection of the length of the cardiac action potential which is sensitive to changes in myocardial oxygen supply (myocardial ischemia). Theoretically, QT interval variability should be an important surrogate index of myocardial ischemia. Specifically, we concentrated on analyzing QT intervals using vectorcardiography (VCG), instead of using the standard 12 lead EKG, to help identify and differentiate

healthy versus acute ischemic/infarct patients. Our reasons for using VCG instead of EKG were based on past studies that highlight limitations of using traditional 12 lead EKG feature extraction for quantitative analysis. Specifically, anatomical positions of the heart, left arm characteristics, individual cardiac electrical dipole movement and torso characteristics vary from individual to individual. In addition, EKG feature extraction is highly dependent on EKG lead selection. These variations can produce quantitative errors within EKG interpretation [4]. VCG analysis overcomes these limitations. Our initial results—while not directly resolving the early ischemia detection question—do provide a key step toward that direction. Specifically, our criterion is able to identify and categorize healthy versus acute ischemic/infarct patients with reasonable accuracy. It is our belief that with future development and continued research, perhaps in combination with other existing criteria, our results may be successful in providing an effective test to help predict early onset, or risk of ischemia.

Background

Myocardial cells are affected differently by the ionic balance in their local environment, and the reflexive control exacted by a combination of endocrine, nervous and intrinsic cardiac systems [5]. The intrinsic cardiac system provides input for neuronal compensatory mechanism in case of cardiac abnormalities [6]. Studies have shown that cardiac abnormalities and the intrinsic system compensatory mechanism affect various electrocardiographic features [2]. Particularly, studies have shown that QT variability increases following acute cardiac ischemia and or infarction [7]. Since several

studies have investigated heart rate variability as a predictor of ischemia, we concentrated on QT variability in our study.

QT Variability

Myocardial ischemia/infarction alters ventricular repolarization through a variety of mechanisms. These alterations are depicted in patients' EKG through the variation of their beat to beat QT interval lengths [8, 9]. Studies have shown that, an index of QT variability (QTVI), is elevated during acute myocardial ischemia and cardiomyopathy [8]. This finding indicates that variability in the QT interval may be used as an index of myocardial ischemia. However, studies also show that QT interval variability is influenced by other factors. Most notable is the fact that the measured value of QT interval is dependent on the EKG lead selection. As a result, QT variability is strongly influenced by lead selection [5, 10].

While qualitatively effective when used on the appropriate EKG lead, QTVI is limited in its results: it does not define a clear quantitative threshold to differentiate healthy versus ischemic patients. This shortcoming of QTVI is partly due to the fact that it has so far has been derived using standard 12 lead EKG. Studies have shown that the standard 12 lead EKG contains errors arising from inherent factors, such as: torso shape variation, dipole location variability, left arm characteristics and individual anatomy that restrict EKG's use for pure quantitative analysis [4]. The remaining problem with QTVI rests on the coupling of heart rate (HR) with QT variability [5, 10]. QT variability is

strongly lead selective as well [5, 10]. Nonetheless, QTVI has helped solidify the fact that QT variability increases following ischemia. Therefore, in order to avoid the inherent quantitative shortcomings of the standard 12 lead EKG, coupling of HR with QTV, and lead selection, our lab investigated pure QT variability using Frank XYZ Vectorcardiography system. The approach to EKG analysis using the Frank XYZ system offers advantages that are detailed in the next section. It should be noted that the conventional 12-lead EKG system was not adopted as the standard for EKG analysis because it's a superior system. The 12-lead system was adopted because it was used more frequently in the middle of the 20th Century when both systems were used and conventions were being established. Perhaps if digital computers were widely available at the time, it is likely that the Frank XYZ system would have become the most widely used system.

Frank XYZ Lead System

The Frank XYZ lead system provides a three dimensional, orthogonal component voltage recording that can be used with considerable accuracy for effective quantitative analysis. This system measures 3-dimensional electrical activity of the heart using a lead system based on a 3-dimensional Cartesian coordinate system with each axis 90 degrees from the others. The x-axis is comparable to Lead I in the conventional 12-lead system, the Y-axis is comparable to lead aVf and the Z-axis is front to back and is similar in direction to lead V2. The magnitude of the cardiac vector in the X, Y, and Z direction has the same calibration. This is not true for the 12-lead EKG system. It arguably provides a

more complete ‘electrical picture’ of the heart, and hence may be considered as the best available system for quantitative analysis [4]. We, therefore, investigated an alternate approach to measure QT variability using QT Variability Window (QTVW) based on vector magnitude derived from the Frank XYZ leads.

Vector Magnitude

Vector magnitude is calculated using orthogonal leads; and is defined to be the magnitude of the cardiac vector as the electrical dipole progresses through the cardiac cycle. At the start of the atrial depolarization, the vector magnitude is essentially zero volts. As the action potential progresses through the atrium, the magnitude of the vector increases to a maximum and decreases back toward zero. Coincidentally, the shape of the vector magnitude wave during atrial depolarization is very similar to the shape of the P wave in an EKG recording. During ventricular depolarization, the calculated vector magnitude increases to a magnitude that is greater than the magnitude during atrial depolarization. The size and general shape is reminiscent of the QRS complex. Finally, as the ventricle relaxes, the calculated vector magnitude rises and falls as it does during ventricular depolarization, however, the shape approximates the shape of a typical T wave in an EKG recording. In summary, vector magnitude is calculated from the length of the vector in the X, Y, and Z planes. Vector magnitude captures 3-dimensional information from the heart and is not subject to the lead specific calculations in the 12-lead system (see Figure 2 page 24).

The shape of the vector magnitude transformation resembles the shape of a frontal plane EKG lead. Transforming three dimensional lead voltage into vector magnitude allows intervals of EKG to be determined without measurement error caused by the projection of the cardiac vector to the axis of the lead system (12-lead or Frank X, Y, Z system). Therefore, we believe that vector magnitude provides the best approach to investigate and study electro-cardiologic features compared to any single lead readout at any given location.

QT Variability Window (QTVW)

Based on our analysis of vector magnitude data, in comparison to QTVI, our studies showed a promising new criterion that could better assess ventricular repolarization during acute myocardial ischemia/infarct. Specifically, our QT variability window (QTVW) analysis showed that the length of the QT interval varied within a window of 16 ms (i.e. ±8 ms) in healthy subjects and QT interval frequently varied by greater than 16 ms in diseased patients.

We conducted our study to test the hypothesis: that vector magnitude derived QT variability exceeds 16 ms in acute ischemic/infarcted diseased patients.

CHAPTER II

LITERATURE REVIEW

Background

Normal heart beats are directed by the spontaneous depolarization of the sinoatrial nodal cells. Acute ischemia or infarct of any coronary artery can result in local myocardium hypoxia and can lead to localized myocardial cell death. Preceding such an event, it is believed that the neuronal control system—through the aid of various intrinsic and extrinsic receptors—detects hypoxic conditions and modulates heart rate to compensate for this condition [6]. Hence, heart rate variation, as detected from EKG, and the neuronal control and compensation thereof, have been the subject of several extensive and complex studies to decipher and predict impending ischemia.

The other aspect of ischemic/infarct heart condition that has been a subject of fewer studies is QT variation. Studies have shown that QT interval of EKG is altered primarily due to changes in ventricular repolarization—as depicted by morphological changes in T wave [2, 5]. Furthermore, studies have shown that QT variability increases following acute ischemia/infarct [7, 8].

The physiological mechanism underlying increased QT variability following acute ischemia or infarction isn't clearly understood. The cardiac myocytes depolarize in a chain event, with one depolarizing cell leading to depolarization of the other through the cell to cell gap junctions. However, the repolarization process is individually achieved by each cell via the cell membrane potassium ion (K^+) channels and the sodium potassium ion exchange (Na^+/ K^+) pumps. It has been argued that in the event of ischemia, localized hypoxia caused due to inadequate blood supply leads to localized electrophysiological changes in both the intra and extra cellular fluids [5]. These localized ionic imbalances subsequently result in a sub or supra optimal performance of the cellular channels and pumps which is depicted as abnormalities in the overall cardiac electrical vector movement [11]. Studies have shown that the ischemia/hypoxia generated abnormalities in the depolarization/repolarization cycles are depicted in the EKG as increased QT interval variability [12-14]. In 2005, FDA issued recommendations for QT/QTc study to be done with all non-antiarrhythmic medications. According to FDA, ventricular repolarization delay is depicted in QT interval prolongation in surface EKG [15]. Delays in cardiac repolarization may lead to cardiac or ventricular arrhythmias, and, in certain cases, sudden death [11, 15]. Ischemia has been known to cause electrophysiological conditions leading to ventricular arrhythmias and sudden death [5]. Therefore, ventricular repolarization variations as depicted through QTV are being given increasing importance.

Electrophysiology

Normal cardiac myocytes receive depolarizing action potential generated by sinoatrial cells through the conduction bundles and Purkinje fiber system. Myocyte to myocyte propagation of action potential, from endocardium to epicardium, during depolarizing cycle is conducted through intercellular gap junctions. Repolarization follows a reverse path, from epicardium to endocardium, and is dependent on each myocyte's own ionic channels and pumps. Repolarization is different from depolarization in the sense that it does not require a central action potential generated from sinoatrial node—or other likewise cells—to initiate the chain; each myocyte performs repolarization after its own specific plateau period.

Ischemia alters cardiac myocytes electrophysiology in several different ways. Studies using porcine models have shown that the normal resting membrane potential of approximately -80 to -90 mV is reduced to -65 to -60 mV in myocytes during ischemia [15]. During global ischemia, resting membrane potential has been shown to reduce as low as -49 mV using isolated animal heart models [16]. At these reduced membrane potentials, the myocytes are considered inexcitable [16]. The likely factor attributed to the reduced membrane potential is the elevated extracellular K^+ concentration [2]. Studies using animal models have also shown that in the event of reperfusion within 15 minutes of occlusion, normal resting membrane potential can be restored [17]. However, studies show that after 20 minutes of sustained occlusion, the extracellular K^+ concentration increase cannot be reversed, even with reperfusion, and the occlusion event leads to permanent cellular damage [2].

Under normal conditions, extracellular and intracellular K^+ balance is maintained by Na^+/ K^+ pumps. These Na^+/ K^+ pumps require a continuous supply of ATP to ensure continuous pump activity. In the event of ischemia/hypoxia, the available intracellular ATP reserves are depleted rapidly. Hence, Na^+/ K^+ pumps are still able to maintain some functionality within the first 10-15 minutes of an ischemic event [5]. Neuronal compensation with catecholamine has been shown to prolong the functionality of Na^+/ K^+ pumps and action potentials. However, resting cellular membrane potential is unaffected by catecholamine [6]. Due to reduced oxygen availability in an ischemic event, cellular ATP production is affected. With reduced ATP availability, fast Na^+ channels are blocked, Na^+/ K^+ pump activity is reduced, and ATP sensitive K^+ channels open. Hypoxia also leads to decreased intracellular pH. Studies have shown that in the event of decreased pH, Na^+/ Ca^+ exchangers are blocked. This event leads to further reduced Na^+ conductance through the membrane and affects the overall ionic balance [5].

Studies have shown that certain metabolites that have been found to exist in ischemic myocardium may have detrimental effects on the myocytes [18]. Specifically, lysophosphoglycerides at high concentrations have been shown to cause membrane loss, resulting in massive Ca^+ influx and cell death [19].

Since at reduced membrane potential, fast conducting Na^+ channels are impeded and Na^+/ Ca^+ exchangers are blocked due to decreased intracellular pH, the Na^+ influx is not considered to be the major cause of membrane potential reduction. At reduced ATP levels, ATP sensitive K^+ channels open and lead to considerable K^+ efflux. This efflux of K^+ is considered to be the chief reason for the membrane depolarization and overall cell

inexcitability [5]. Consequently, around eight minutes from an ischemic event, elevated extracellular K^+ levels start to diffuse toward healthy myocardial zones and may reach near normal levels after 18-20 minutes near the ischemic/healthy zone juncture [20]. This overall process presents as a delayed recovery of ischemic myocytes [21].

The inexcitable or slightly depolarized state of the ischemic myocardium has also been termed as "postrepolarization refractoriness" (PRR) [5]. Experiments have shown that myocytes in such PRR states may present delayed recovery—by several milliseconds—even after local repolarization process has completed [22, 23]. Ischemia has been shown to induce beat-to-beat action potential alternations which may cause refractory period to alter as well. Furthermore, due to localized differences of extracellular K^+ concentration levels between ischemic and healthy areas of myocardium, spatial inhomogeneity of recovery exists within myocardium in the presence of ischemia [24]. Spatial inhomogeneity may become apparent in shorter cycle lengths due to increased heart rate or premature depolarizations [25]. Such localized inhomogeneities may cause inhomogeneity in membrane potentials, action potential length, and the recovery time of the myocyte. Studies suggest that such inhomogeneities are embodied in the observations where the ventricular time to recover varies [5]. It is suggested that this variation of ventricular time to recover is due to a combination of delayed conduction time and the difference in effect ischemia has on various localized regions of myocardium. Another contributing factor to slower conduction velocities is decreased cellular electrical coupling or increased resistance. After 12-18 minutes intracellular resistance increases rapidly; and after 24 minutes of ischemia, majority of gap junctions in the affected region dissociate [5, 26].

It has been suggested that cardiac ganglia, with the aid of intrathoracic neurons, and input from intrinsic mechanoreceptors, chemo and baroreceptors, can control cardiac functions independently of the central nervous system during certain instances [27]. Ischemia has been known to activate cardiac receptors and in turn cause a response from the central nervous system and the cardiac ganglia [28]. The compensatory afferent commands may lead to heart rate variation and QT variability [29].

Electrocardiography

Between 1902 and 1903, Willem Einthoven published his findings on human ECG recordings. Since then, "the basic principles of this technique have remained unchanged" [30]. While ECG does monitor cardiac electrical functionality in real time, its diagnosing capability is not considered as a good as biomarkers in detecting life threatening ischemia (see Table 1). However, biomarkers, especially Troponin I take anywhere from 12 to 24 hours after a myocardial infarct to reach significant levels in blood plasma. The delay in definite infarct diagnosis leads to delay in treatment, which can be significant at times [30].

ECG is a surface readout of the heart's overall electrical dipole movement through the heart. Individual anatomical variations and low level cellular electrophysiology adds complexities to the overall dipole and its propagation. As such, using ECG to unravel underlying cardiac pathologies is a challenge. Nonetheless, over

the past century, studies have helped enhance our understanding, interpretation and use of this very powerful diagnosing tool (see Figure 1.)

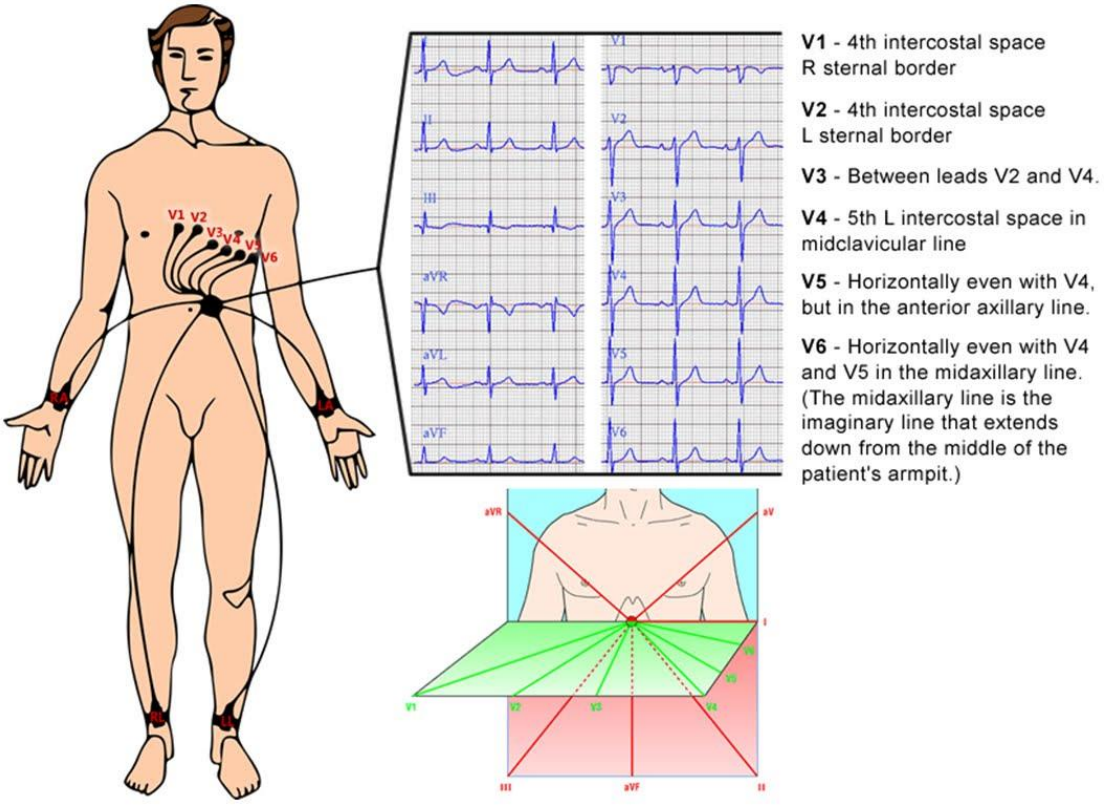

Figure 1: Conventional 12 Lead EKG electrode placement and resultant surface electrical dipole travel projection. 12 Lead EKG offers a morphological clinical use in cardiac assessment which has been favored by physicians. (*Reprinted from Bembook*)

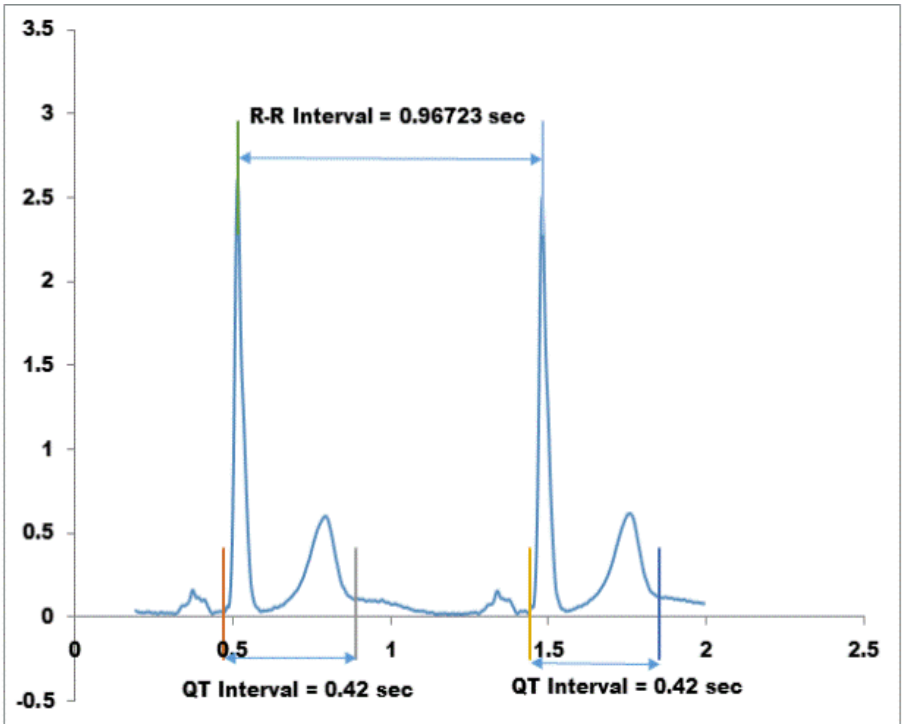

Figure 2: QT and RR intervals on two consecutive vector magnitude beats of a healthy patient. The shape of the vector magnitude waves resembles the shape of the P, QRS, and T wave of a conventional EKG which facilitates measurements of the RR and QT intervals.

Studies have shown that a small but significant percentage of patients with ischemia or infarction present with little or no ECG changes [31]. Established ischemia diagnosis criteria for ECG present a variety of sensitivities and specificities. The Minnesota code describes an ST elevation of ≥ 0.1 mV as significant. The Minnesota code has been shown to have a sensitivity of 56% and a specificity of 94%. Other criteria have an overall sensitivity ranging from 45% to 69%, and specificity of 81% to 98% [30]. See table 1. Hence, there still exists a need for a better predicting and real-time diagnostic solution. Several scientists have worked toward using complex mathematical analyses to extract more information from existing 12 lead ECG.

Table 1: Sensitivity and Specificity of ST Elevation According to Methodology Used

Method	**Sensitivity**	**Specificity**
McClelland et al	34-45%	98-94%
Minnesota code	56%	94%
GUSTO IIa trial	45-69%	98-81%
Selvester et al	32-72%	95%

New techniques utilized so far have involved computerized measurements and analysis of ST segment and Q wave changes in the 12 lead ECG [10]; various forms of scoring techniques by combining ECG information with clinical history [32]; and, analyzing HRV and utilizing its coupling with ST elevation [33, 34]. In the recent past, exercise tolerance tests (ETT) combined with various computerized complex analysis have also proven to be noteworthy [7]. Two new technique called the Athens score and ST-HR hysteresis coupled with ETT have also proven to be powerful tools in diagnosing ischemia, especially transient ischemia [7]. Additionally, some scientists have suggested analyzing QT interval variability to diagnose and predict ischemic conditions. One such technique is QT variability index (QTVI) which makes use of QTV coupled with HRV [29].

By definition, QT interval depicts the depolarization and repolarization cycle. Studies have shown that QT variability is mostly induced by repolarization instability [29]. Studies show that ventricular arrhythmia—a leading cause of sudden cardiac death—may be caused by ventricular repolarization instability [5]. The underlying

physiology of this instability has been discussed earlier in this chapter. QTVI was proposed as a method to gauge the QTV with respect to HRV [29].

QTVI, while being a powerful tool, fails to provide a clear threshold or cutoff between healthy and diseased patients. Reasons for this failure include particular lead selection and HRV coupling [13, 35]. QTV has been shown in studies to be highly dependent on selected lead [13]. Specifically, QTV is not the same in every lead. Analysis of our own data also showed similar results (See figure 3 & figure 4). HRV coupling has also been shown in studies to have its own drawbacks. The relationship between HRV and QTV is non-linear: insofar, there is hysteresis involved between HR and QT interval [35]. We attempted to negate the setbacks suffered by QTVI by analyzing pure QT variability derived from vector magnitude using Frank XYZ lead system.

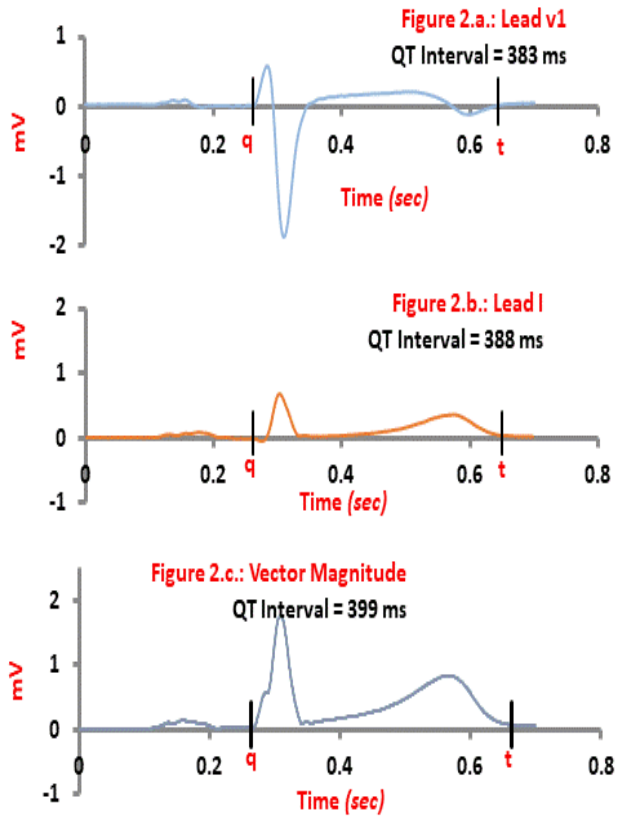

Figure 3: QT interval comparison between leads v1, I and VM. QT interval varied from 383 in lead v1 to 399 in VM. 12 lead ECG is two dimensional projections, while VM is a more complete representation of the three dimensional EHV. QT interval length derived from VM in most cases retains the longest interval length of all leads. VM also retains other EKG boundary features.

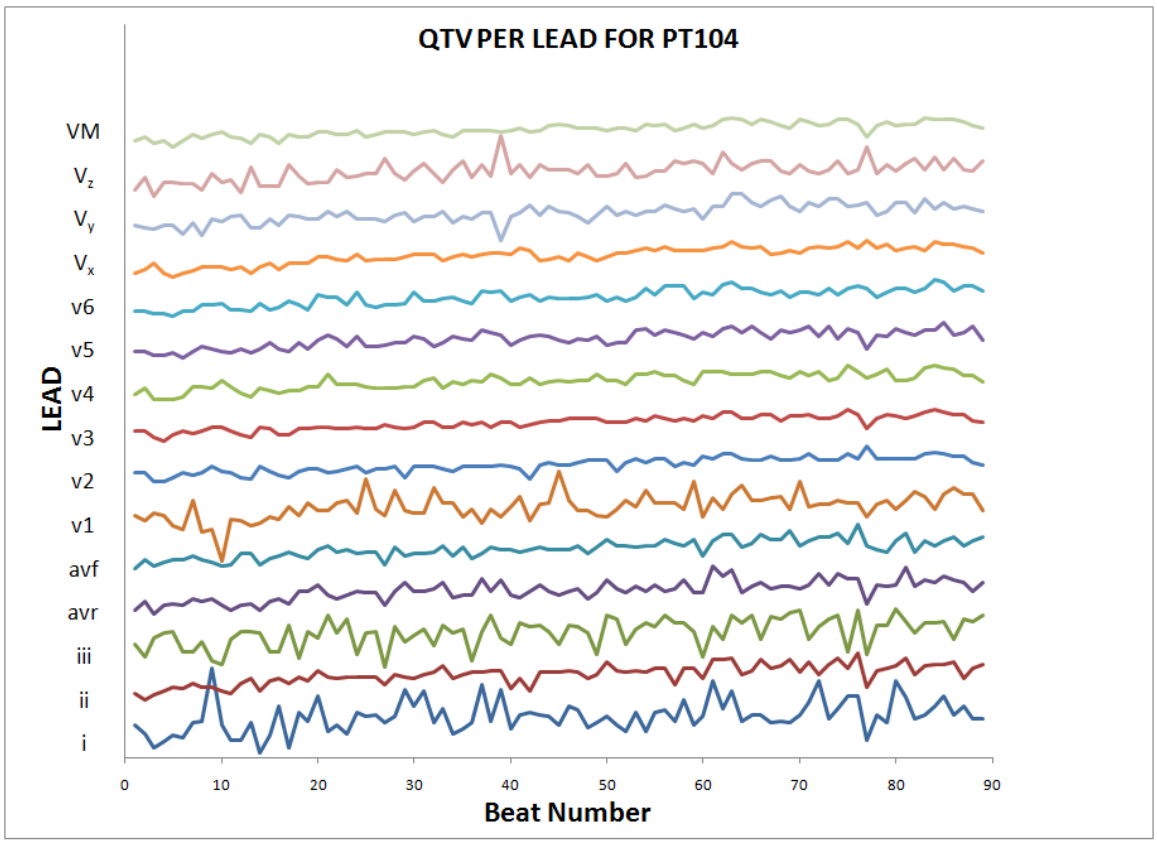

Figure 4: Beat-to-Beat QT interval comparison between EKG leads. QT interval length differs for lead to lead. Hence QTV depends on chosen EKG lead.

Vectorcardiography

In 1956, Ernest Frank introduced a vectorcardiographic system that was argued as superior to other likewise systems for quantitative analysis [4]. It is based on the use of 7 electrodes that help derive 3 orthogonal leads (See figure 5).

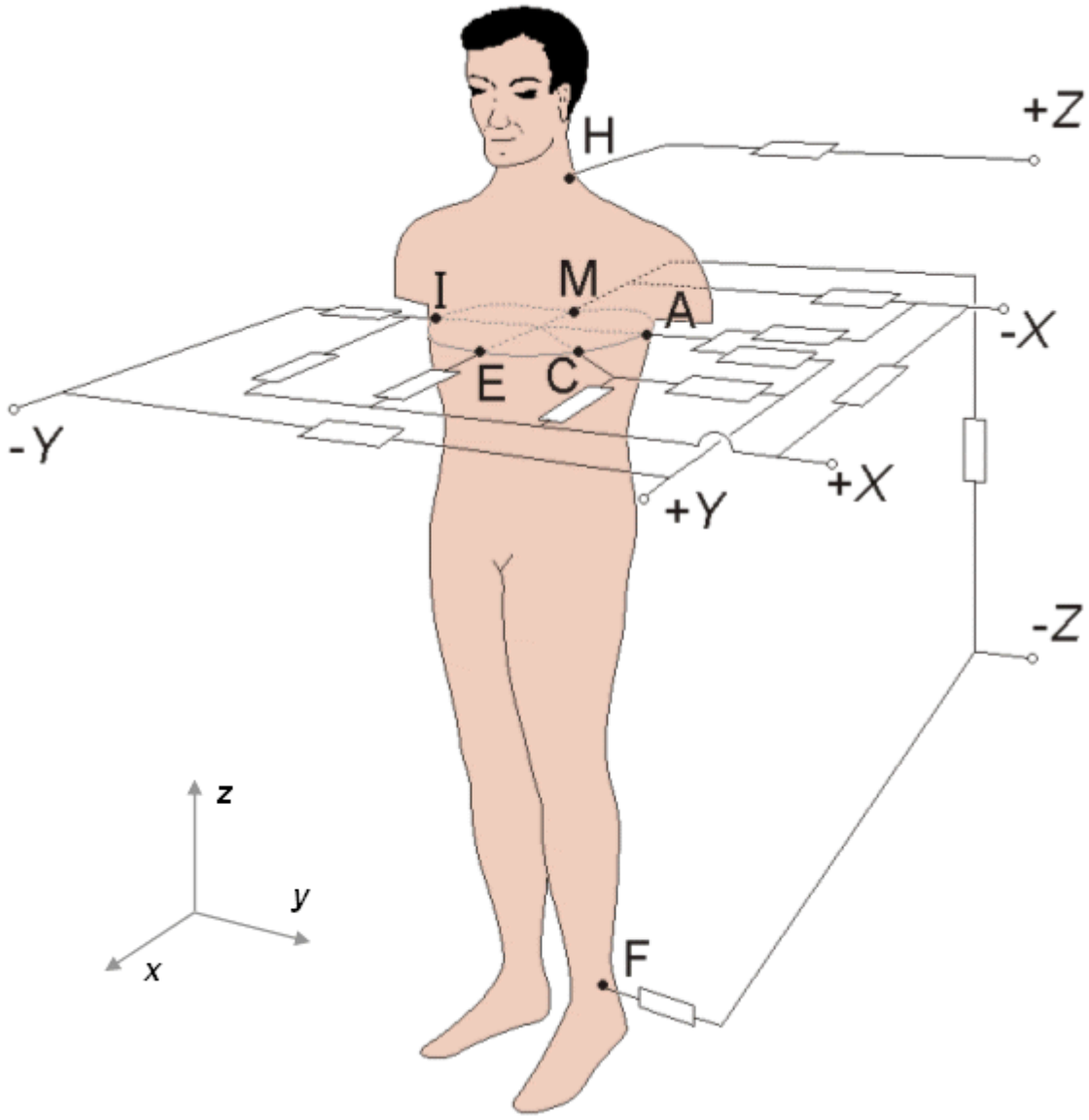

Figure 5: Frank XYZ VCG electrode placement. Inline boxes represent normalizing electronic resistance. Electrode C is placed to normalize anatomical variations and body axis positions. (*Reprinted from Bembook*)

Frank devised this system for accurate quantitative analysis of surface electrocardiography, and foretold the need for advances in mathematical analyses of the electrocardiograms to further medical diagnosis of cardiac diseases [4].

Vector magnitude is calculated using orthogonal leads, and is defined to be the magnitude of the cardiac electrical dipole vector as it progresses through the cardiac cycle. The shape of vector magnitude transformation resembles the shape of a frontal plane ECG lead. Transforming lead voltage into vector magnitude allows intervals of ECG to be determined without measurement errors caused by the projection of the cardiac electrical vector to the axis of the lead system. Therefore, vector magnitude provides the best approach to investigate and study electrocardiologic features compared to single lead readout at any given location. Hence, we utilized vector magnitude for our analysis and computations.

Pseudo-vectorcardiography

Since VCG is not a common undertaking under taking during clinical settings, we wanted to investigate if we could extract data similar to VCG from conventional 12 lead EKG. As the conventional 12 lead EKG does not provide true orthogonal leads, we identified the "pseudo-orthogonal" leads—and hence the resultant vector magnitude—"pseudo-vectormagnitude." Furthermore, the resultant QTV Window was also renamed "pseudo-QTVW" or "pQTVW."

To extract pseudo-vector magnitude (pVM), a comparison of standard leads to Frank XYZ leads led to 5 reasonable lead choices: I, aVf, V1, V2, and V6. Figure 6 explains the arrangement of leads for both systems.

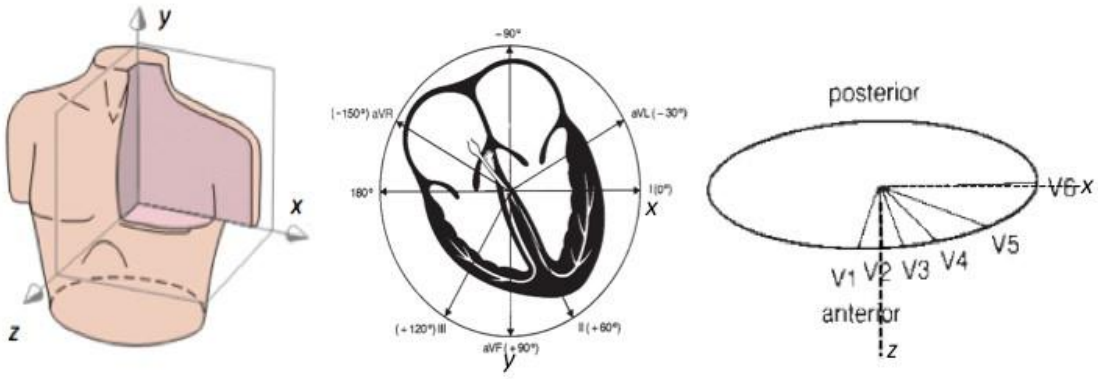

Figure 6: Comparison of orthogonal planes to conventional 12 lead EKG electrode placement. Evidently aVf resembles y axis. (*Reprinted from Bembook*)

Based on its orthogonal position, aVf was the clear choice to represent Y; additionally, figure 7 demonstrates that aVf is an approximation of Vy. In Eindhoven's original naming of the coordinates he followed the mathematical conventions of x representing left to right (from the physician perspective of the patient), y representing foot to head. When Frank introduced vector cardiography, x and y remained the same convention and z represented the anterior to posterior. Thus the frontal plane is described by x & y and the horizontal plane by z & x and the sagittal plane defined by y & z. aVf is the only lead in the y direction thus it was the natural lead choice and starting point for choosing representative leads. All precordial leads are perpendicular aVf since they are in a perpendicular plane. This provides four possible permutations which are summarized in Table 2.

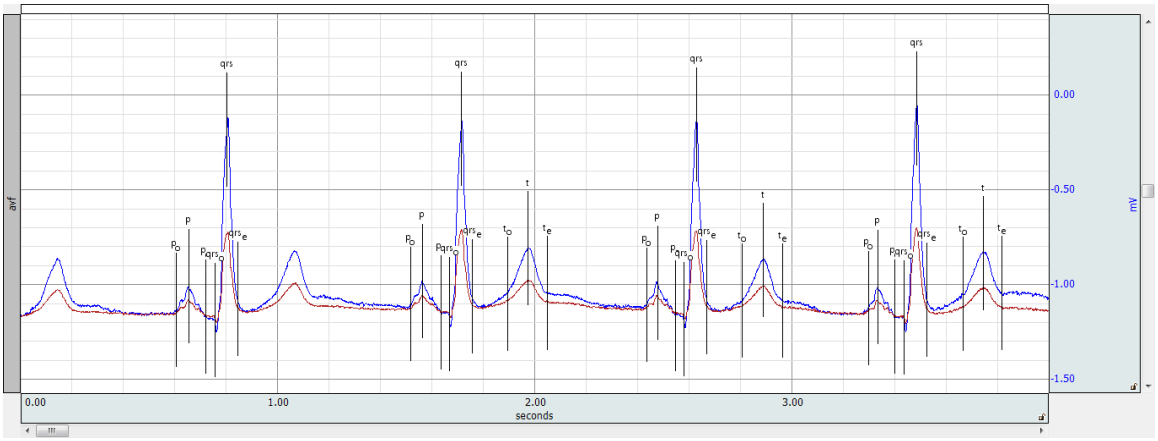

Figure 7: Comparison of aVf (blue) of the conventional 12 lead EKG and Vy (red) of the Frank XYZ. Intervals trace very closely in comparison of one to the other. Tracing illustrated using Acknowledge software.

Table 2: Summary of Lead Selections

Selection Number	Lead Selections for x,y,z	Frontal Plane	Horizontal Plane	Sagittal Plane
1	I, aVf, V1	I, aVf	V1	aVf
2	I, aVf, V2	I, aVf	V2	aVf, V2
3	V6, aVf, V1	aVf	V6, V1	aVf
4	V6, aVf, V2	aVf	V6, V2	aVf, V2

Selection 1 provides two leads in the frontal plane (I & aVf), one lead in the horizontal plane (V1), and one lead in the sagittal plane (aVf). Selection 2 provides two leads in the frontal plane (I & aVf), one lead in the horizontal plane (V2), and two leads in the sagittal plane (aVf & V2). Selection 3 provides one lead in the frontal plane (aVf), two leads in the horizontal plane (V6 and V1), and one lead in the sagittal plane (aVf). And selection 4, provides one lead in the frontal plane (aVf), two leads in the horizontal

plane (V6 & V2), and two leads in the sagittal plane (aVf & V2). Preliminary experiments were performed using V6, aVf, & V2. Results were similar to the work using the Frank XYZ system. Using Biopac software, EKG tracings were overlaid to choose the closest representation of the X Z components (see figures 8 and 9.) Using the same sample of Vx both I and V6 were compared to Vx and again both I and V6 had very similar shapes but V6 was a better approximation of Vx's amplitude. In fact, it was almost an identical tracing of Vx.

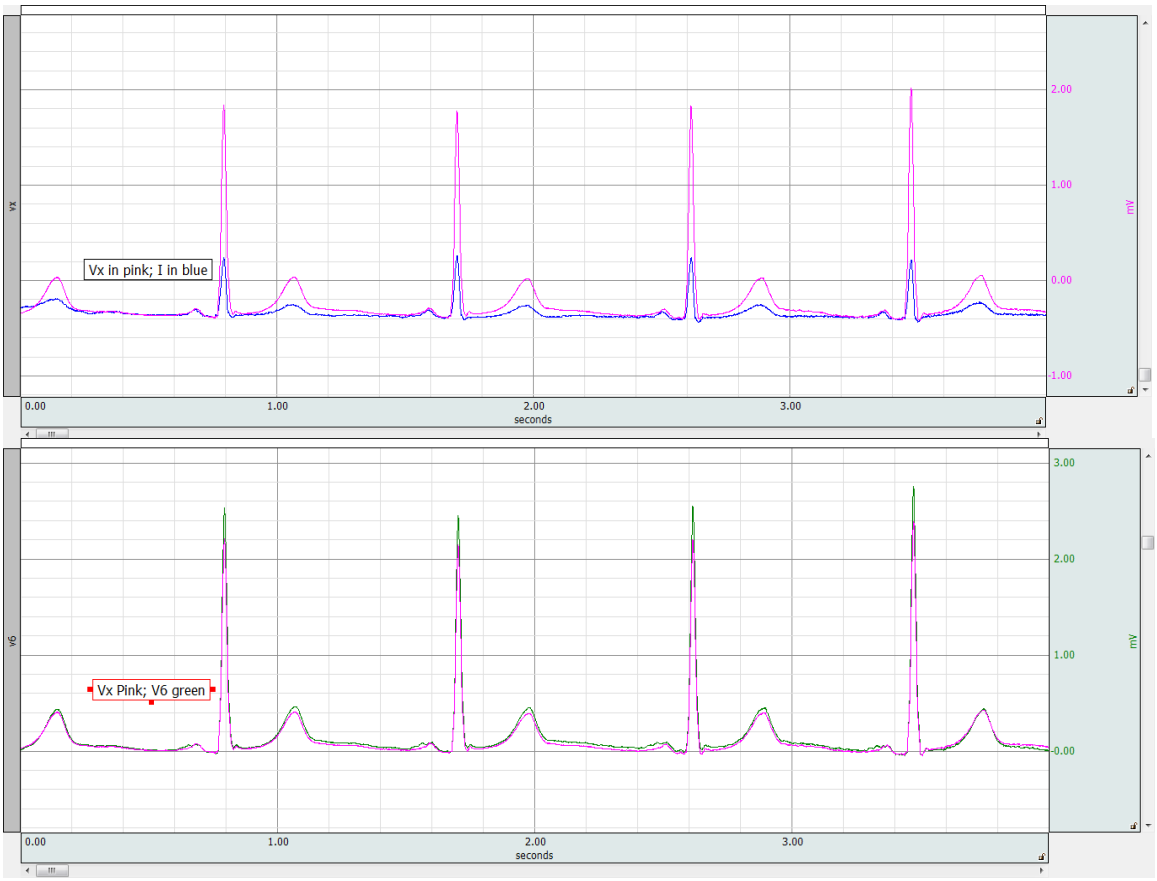

Figure 8: Comparison of V1 and V6 of conventional 12 lead EKG to Vx of Frank XYZ system. Tracing illustrated using Acknowledge software.

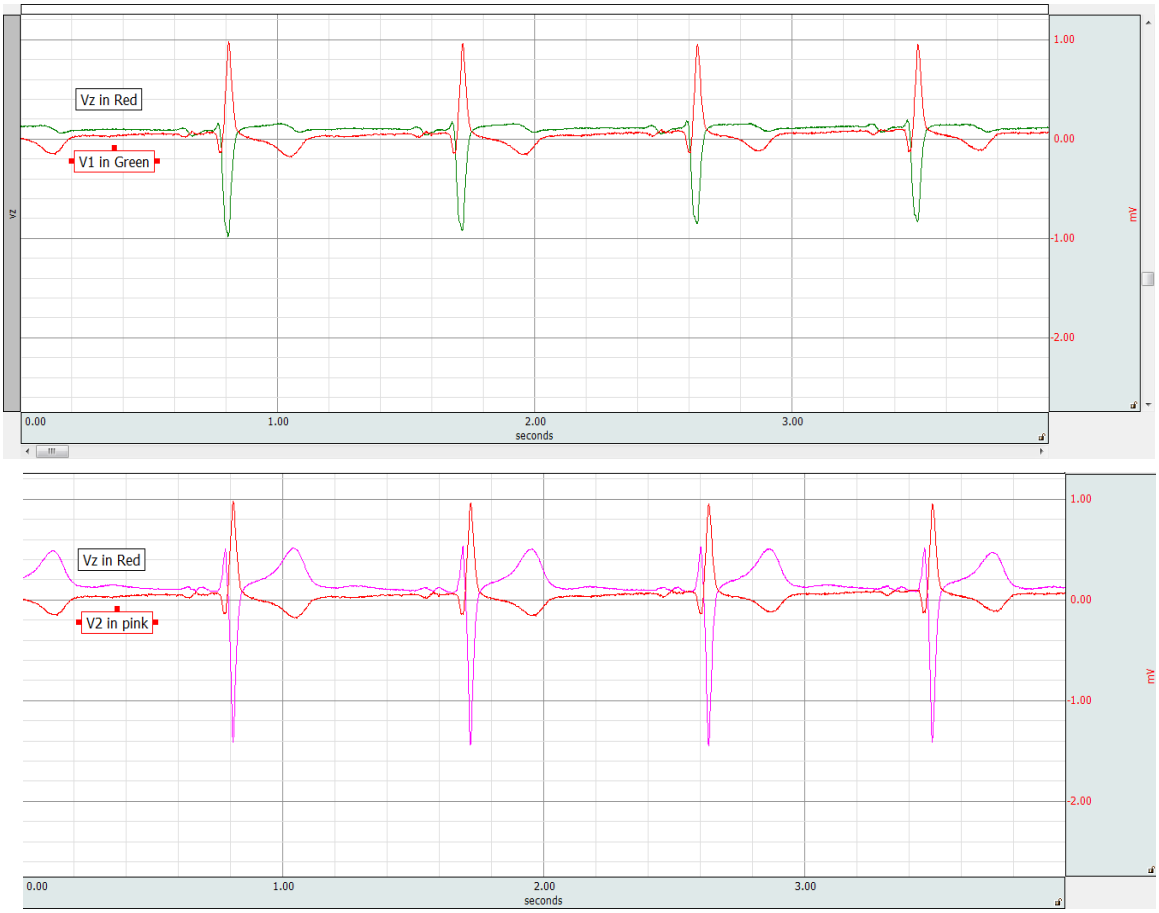

Figure 9: Comparison of V1 and V2 of conventional 12 lead EKG to Vz of Frank XYZ system. Tracing illustrated using Acknowledge software.

While matching the amplitudes is not really important since we are evaluating QT interval, it was thought that better consistency could be achieved in choosing the T end interval. Using the same sample of Vz, both V1 and V2 were overlaid and compared to Vz to look for the closest waveform which best compares in shape and amplitude. V1 and V2 both reciprocated the shape of Vz as expected but V1 was the best approximation of Vz's amplitude (See figure 9.)

After several permutations, and comparing derived pVM to VM, leads V6, aVf, and V1, were chosen to represent Frank XYZ component leads. Thus the pVM was calculated in the same manner using Pythagorean's theorem. Figure 10 compares the vector magnitude from the Frank XYZ to the pseudo-vectormagnitude derived from the pseudo-orthogonal leads of the conventional 12 lead ECG.

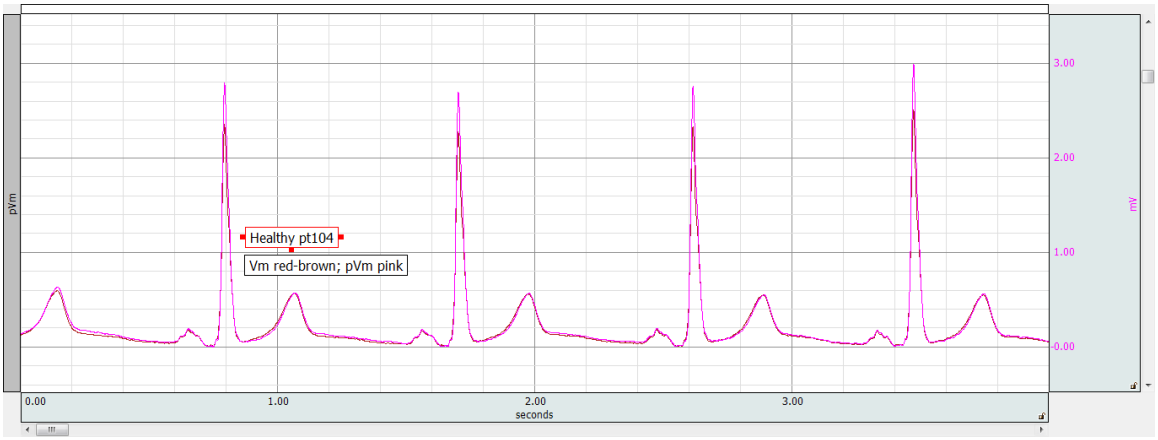

Figure 10: Comparison of VM and pVM. Tracing illustrated using Acknowledge software.

Data Mining

Data mining (DM) and machine learning techniques (ML) are powerful tools in efficiently extracting underlying relationships between data that may otherwise be hidden from investigators. We elected to apply DM and ML techniques to our post-processed data extracted from VCG derived VM, containing QT and RR intervals, and their statistical variables. Our initial goal was to identify obvious patients undergoing MI—as evident from their 12-lead conventional ECG depicting an ST elevation. Using ML techniques, our goal was to systematically evaluate QT and RR variables to see if they

can help identify healthy and ischemic patients; and which variables out of the two and their extracted features prove to be the most significant in predicting healthy and ischemic patients. Our method shows promise in automatically identifying MI patients based on their QT and RR variables with accuracy of 98.31%, sensitivity of 100% and specificity of 96.55% (see results and discussion.)

Data Mining and Machine Learning Techniques

Data mining tools and methods have made significant progress in extracting useful information from unstructured data. Data mining has made significant progress by augmenting machine-learning techniques to help with complex decisions in several domains e.g. prediction of medical events, forecasting financial trends, customer choices etc. Data mining and machine learning techniques have been synergized with exponentially increasing computing power and decreasing cost of computing hardware-software [36]. Data mining and machine learning techniques have evolved over time, with diverse characterization. Data mining is defined as a process to find patterns and relationships in the data. The results of data mining are directly affected by the quantity and quality of the data [37]. Per Alizadehsani et al. [38], data mining is a process of extracting patterns and relationship from single or multiple datasets.

Machine Learning has evolved significantly since development of first computer ENIAC was in 1946. ML is a sub-discipline of artificial intelligence (AI). SAS Institute explained machine learning to be a method for automating iterative analytical model to learn inductive rules from complex data patterns. Machine learning has been defined as a scientific method to develop an optimum program which helps make intelligent decisions

by recognizing, learning and improving decisions from complex patterns [39]. Recent amalgamation of data mining, machine learning, and cost effective computing power is enabling analysis of big medical data to solve practical applications in clinical medicine [40] and explore the clinical domain knowledge to support clinical diagnosis and prognosis. Data mining along with machine learning methods use various techniques to solve wide-ranging problem that stretches to wide-ranging domains. Some of the popular methods are prediction, clustering, and classification. Classification analysis occurs in a wide range of human activity to make decisions using available information and repeatedly making such judgments in new situations [41]. Spam filtering is a classic example of machine-learning based classification tools to identify spams and filter them from non-spam emails.

Predicting medical outcomes with increasing accuracy, is a goal of clinical and data science researchers. Prediction of severe health manifestation before it deteriorates to permanent damage, may help in managing and reversing the trend of health outcome with timely treatment. Dangare and Apte showed that heart diseases may be predicted using individual behavioral and personal data [42]. Alizadehsani et al, proposed inclusion of discriminative features to increase prediction accuracy for diagnosing coronary artery disease [38]. There is, however, an opportunity cost for adding a new feature to the mix for performing prediction using data mining techniques; and therefore equilibrium is essential to optimize the use of prediction techniques by utilizing minimum parameters and predicting with highest accuracy. Appropriate attention towards data preparation and selection of machine learning techniques helps limit false positives and negatives in the results; and allowing the complete classification process to be more robust in its

deployment. This study is limited to using classification tools to identify infarcted patients and heathy patient. As such, it does not focus on any indicator to classify the various levels of myocardial ischemia or blockage, or the severity of Myocardial Infarction.

The next section covers methods of data collection; data preparation; introduction to various machine learning subroutines and evaluation methodologies utilized in this study; and K-Fold cross validation techniques.

CHAPTER III

METHODS

Data Source

Data for our study were acquired from the PTB database, which is a part of Physionet – an online repository of medical research data maintained by MIT Laboratory for Computational Physiology [43]. Both conventional 12 lead, and Frank XYZ, data of healthy controls and diseased patients are contained in the PTB database. PTB database contains records from 290 patients: 52 are healthy controls, while the remaining are diseased patients with or without co-morbid abnormalities associated with their hearts.

Data Preparation

We limited our study to include only acute myocardial ischemic and infarcted patients for our initial investigation. Our study excluded patients with ectopic beats or arrhythmias; non-infarcted patients; patients without catheterization data; and patients with excessively noisy signals. We selected 29 diseased patients (mean age 54.4 ± 9.8 years, 7 females) and 30 healthy controls (mean age 53.6 ± 16.8 years, 5 females). A

cardiologist independently read subject patient's 12 lead conventional ECG data to verify presence or absence of ischemia or infarction, and also identify and exclude patients that matched our exclusion criteria. Data in PTB database are limited in the number of beats per patient. We therefore investigated and included 90 beats per patient to standardize RR and QT extractions per patient. Selected patients' XYZ lead data was downloaded and analyzed using Biopac's Acqknowledge software [44]. To eliminate a common low frequency noise inherent in most ECG signals, (called baseline wander) [45], individual VCG data for each patient was filtered using bandpass filter (Blackman -61dB window, low pass at 0.6 Hz and high pass at line frequency 60 Hz), and zeroed with respect to vertical (mV) axis using Acqknowledge's built-in functions. Although other researchers have advocated use of Blackman window [46], we arrived at the choice after multiple iterations of various filters available in the Biopac software on multiple patients' ECG signals. Blackman window with settings stated above proved superior to all other filters in our testing. Subsequently Vector magnitude (VM) was derived from the orthogonal VCG leads. QT intervals were marked on the vector magnitude using Acqknowledge's ECG boundary locating function based on ECGPUWave algorithm [43], and identified by manual inspection by both the principal and one of the co-investigators.

Approximately 5% of the intervals were misidentified and manually corrected. Computer assisted and manually marking method of picking Q and T points is an acceptable method as suggested by FDA [47, 48]. As previously stated, ninety beats were selected and QT intervals for each subject were marked, verified, and analyzed. Feeny et al [45], showed that QT variability is evident in as little as ten seconds of ECG data which amounts to less than twenty beats. Use of ninety beats afforded us an adequate data

set for our feature extraction required for machine learning analysis. Figure 2 provides a visual representation of QT and RR interval extraction from two consecutive beats from a healthy patient's VM signal.

To avoid effects of hysteresis, we looked towards studies that have shown that effects of hysteresis are eliminated when short intervals of patients' data are studied [45]. Therefore, we opted to conduct our study over 90 beats per patient to avoid the hysteresis effect.

Machine Learning Classification Routines

The correct choice of algorithms to be used, amongst several available, in ML technique is essential. While several researchers have addressed qualifying best classification algorithm, all classification algorithms have their pros and cons, depending upon the data and the domain being analyzed. Our team used IBM SPSS Modeler 17.0 (SPSS) and chose the following four algorithms: Artificial Neural Networks (ANN), Support Vector Machine (SVM), Decision Tree-C5, and Ensemble method for classification analysis.

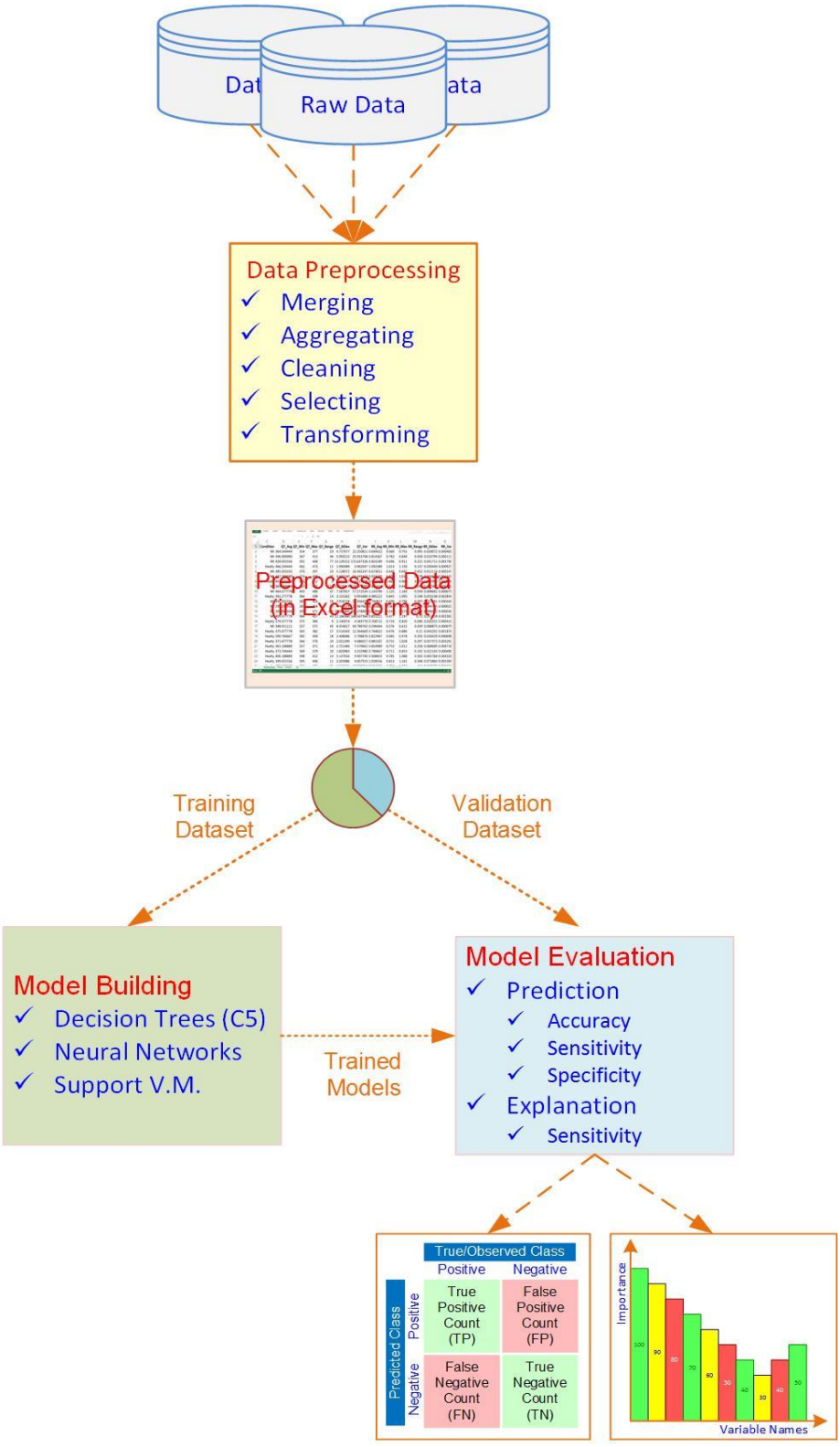

Figure 11: Flow chart representation of Data Mining process flow.

Artificial Neural Networks (ANN)

Artificial neural networks (ANNs) derive its foundation from the brain's cognitive neural connection. Neural Networks are highly sophisticated non-linear statistical data modeling techniques. Neural networks can be supervised or non-supervised [49], to model complex relationships between inputs and outputs, or to find patterns in a given dataset [38]. We used supervised learning Multi-layer perceptron (MLP) ANN with back-propagation a powerful function for prediction and classification problems [36].

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a robust classification and regression technique used to optimize model for accuracy and overfitting data with very large numbers (for example, thousands) of predictor fields (IBM). We chose predictor importance for model evaluation and raw propensity scores to analyze the data.

Decision Tree- C5

Decision trees algorithm is another popular machine-learning algorithm and works by splitting the sample based on the field that provides the maximum information gain (IBM). Process works by splitting subsample and the process repeats until the subsamples cannot be split any further. We ran several decision tree model, e.g. CHAID;

and based on accuracy we chose C5.0 for our final analysis. We chose predictor importance for model evaluation and raw propensity scores to analyze the data.

Ensemble

Advancement in computing power and convergence of data mining and machine learning techniques have allowed development of a very powerful software algorithm that automatically runs multiple algorithms simultaneously; compares their results and predictive strength of each algorithm; and finds the best algorithm amongst all used ones. We chose Ensemble method for obvious reasons: to compare different models for binary outcomes and automatically choose the best approach for the analysis by comparing the measure of sensitivity, specificity, and accuracy against the other three methods selected for analysis.

Beat-to-beat derived QT and RR variables are represented as time series with non-linear, non-stationarity features [7, 50]. Feature extraction is an essential step in data mining process for proper application of meaningful ML techniques [50]. We used QT and RR time series to extract first and second order statistics (FSST) features to represent the original characteristics for better classification results [50]. FSST used in this study represents average, minimum, maximum, range, standard deviation, and variance. FSST features are derived from all 59 patients' QT and RR time series records. Researchers have established that ECG is a nonstationary and nonlinear time series physiological signal [51]. Challis et al defined stationarity as a quality of a process in which the statistic

measures such as mean and standard deviation do not change with time [52]. Therefore, it is important to select features that accurately represent non-linear and non-stationary characteristics of the ECG signal. Kanters *et. al.* proved that standard deviation of the R-R intervals explains the non-linearity characteristics of the heart rate [50]. Similar, in this study we chose condition (Healthy and MI - 0 and 1) as our target response variable.

Analysis

Each patient's XYZ data was downloaded and analyzed using Biopac's Acqknowledge software [44]. To reduce high frequency noise, individual XYZ data for each patient was filtered, using bandpass filter (Blackman -61dB window, lowpass at 0.6 Hz and highpass at line frequency 60 Hz), and zeroed with respect to vertical (mV) axis using Acqknowledge's built-in functions. The three dimensional vector magnitude (VM) was then calculated using the three dimensional Pythagorean theorem represented below in equation (1):

$$VM = \sqrt{V_x^2 + V_y^2 + V_z^2} \qquad (1)$$

To pick 'QRS' onset and 'T' end points, we looked for an established standard. FDA has established guidelines for picking QT intervals to conduct long QT studies for pharmaceutical approval procedures [15, 47]. Per FDA guidelines, QT intervals may be picked with the assistance of computer software, followed by an independent verification and correction as required.

Hence, we used BIOPAC to pick QT intervals on the vector magnitude using software's EKG boundary locating function which is based on ECGPUWave. QT intervals for all 59 patients were visually verified by the principal investigator and independently verified by a second investigator. A total of 90 beats for each subject were marked and analyzed. If the signals were deemed too 'noisy' (difficult for software or the investigator to pick or verify Q or T points), we opted to reject that signal.

The pseudo vector magnitude (pVM) was calculated using three dimensional Pythagorean theorem represented below in equation (2):

$$pVM = \sqrt{V_6^2 + aVf^2 + V_1^2} \qquad (2)$$

To compare our study to the established QTVI [8], we calculated QTVI, and its requisite statistical parameters from vector magnitude of same patients' data. QTVI is a ratio of QT variability to HR variability. A logarithm of the ratio is taken to introduce Gaussian statistics [8]. Statistical Mean QT interval length (QT_m), Mean HR (HR_m), and statistical QT and HR variances (QT_v and HR_v, respectively) were also calculated from vector magnitude. These values were used for QTVI calculation using formula represented in equation (3):

$$QTVI = log_{10}\left[\frac{(QT_v)/(QT_m)^2}{(HR_v)/(HR_m)^2}\right] \qquad (3)$$

It should be noted, however, that our SB-QTVW method uses beat-to-beat QT variability and is designated in the paper as QTV; while QTVI uses statistical QT variance of a patient's data set, and is designated as QT_v.

Our QTVW was established as QTV boundaries that we observed when QTV data was plotted versus beat number for each patient after initial experiments were conducted. For each patient data, after extracting QT interval on a beat to beat basis, we plotted the intervals on a time interval versus beat number basis. We observed that QT variations for each patient had a mean value or baseline different than others. Therefore, we subtracted each patient's respective mean QT interval from their individual beat to beat QT intervals and plotted those (See figure 13). As normal EKGs usually have millisecond resolution, we rounded our resultant data to the nearest millisecond and then subsequently analyzed it.

After initial experiments, it was evident that QTV of healthy controls appeared to be constrained within a certain time frame (See figure 13, Panel A), while of diseased patients appeared to exceed such a constraint—specifically a range of ±8 milliseconds (See figure 13, Panel B). We then proceeded to define a healthy QTV boundary and conducted further experiments. We double checked our methods and concluded that our initial data indeed showed an existence of a constraint that healthy controls' QTV lies within. We termed that constraint as Sewani-Benjamin QT Variability Window (SB-QTVW)—a window of time representing the maximum variations of QT intervals in healthy heart. We, therefore, propose that SB-QTVW is a method that provides QT variability threshold between healthy controls and infarcted patients, and that it may be used to separate healthy subjects from patients with acute myocardial ischemia and/or infarction. We then conducted blind experiments of patients, plotted their QTV accordingly, and analyzed results accordingly.

Methodology for data analysis for pSB-QTVW was exactly similar, except we used leads V1, V6 and aVf instead. From the previously selected list of patients, a total of 29 healthy controls and 25 diseased patients were selected for analysis using pVM. Reduction in numbers was primarily due to excessive noise in the lead signals for the rejected patients.

All data analyzed for this study was acquired from PTB database and can readily be accessed from its repository (https://www.physionet.org/physiobank/database/ptbdb/).

Accuracy, Sensitivity, and Specificity

In this study, we used three performance measures: Accuracy, Sensitivity, and Specificity. Refer to equations 4-6 below:

$$Accuracy = \frac{True\ Positive+True\ Negative}{True\ Positive+True\ Negative+False\ Postive+False\ Negative} \quad (4)$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive+False\ Negative} \quad (5)$$

$$Specificity = \frac{True\ Negative}{True\ Negative+False\ Positive} \quad (6)$$

Sensitivity measures the proportion of correctly predicted positives outcomes; also sometimes refer to as “true positive rate”. Specificity (also called the true negative rate) measures the proportion of correctly predicted negatives outcomes also sometimes

refer to as “true negative rate”. In our study, the sensitivity calculates percentage of correctly predicted healthy people, while specificity calculates percentage of correctly predicted diseased patient. Accuracy measures proportion of correctly predicting positive and negative outcomes and measures correctly predicting healthy and diseased patients in our study. Theoretically, results should be of high accuracy, sensitivity, and specificity. However, optimized solutions do compromise one or the other factors.

K-Fold Cross-Validation

According to Kohavi, a good estimation method should have low bias and low variance [53]. We chose k-fold cross-validation method vs random sampling validation method to support high classification accuracy, also called rotation estimation [36]. We also used stratified 10-fold cross-validation method for all techniques used for classification of myocardial infarction. There are three cross-validation methods: a) Random Subsampling, b) K-Fold Cross-Validation, c) Leave-one-out Cross-Validation. Our dataset was divided into ten mutually exclusive train/test sets (folds) samples for training and testing (Refer to figure 12). Also for each set 10% of the data was allocated for training the algorithm. Stratified cross validation method breaks the dataset into equal portion for every fold. Ten-fold cross validation method repeats each fold ten times to remove bias from random sampling. We followed the three-step 10-fold cross-validation procedure to estimate the error rate; a technique explained by Delen et al [36]. Accuracy of 10-fold cross validation is defined by equation: (7), where CVA is overall average validation accuracy of k number of fold.

$$CVA\ (CrossValidationAccuracy) = \frac{1}{k}\sum_{i=1}^{k} Accuracy_i \qquad (7)$$

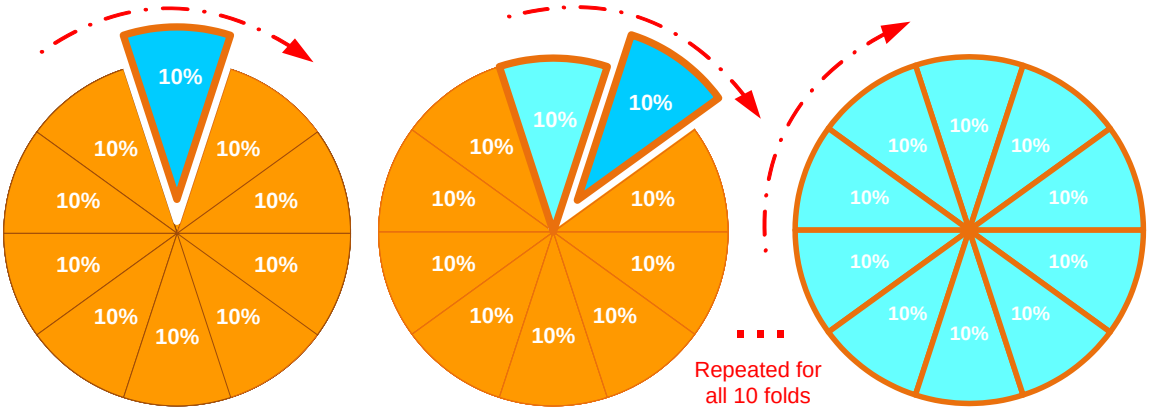

Figure 12: 10-fold Cross Validation Procedure

Statistical Analysis

All statistical analyses were conducted using SPSS software version 19.0 [54]. All data was expressed at the mean ± standard deviation. Each parameter values were calculated separately for each group. Data was compared using t-test. Statistical significant level was accepted at $P<0.05$. Sensitivities and Specificities were calculated for various thresholds and plotted using Receiver Operator Curve (ROC) for further analysis.

CHAPTER IV

FINDINGS

SB-QTVW

As figure 13, plate A shows that majority of healthy controls' QTV varied within ±8 ms window from their mean QT interval length. 4 out of 30 healthy controls exceeded the ±8 ms window limitation, resulting in 4 false positives; while 2 out of 29 diseased patients remained within the window, resulting in 2 false negatives. Overall mean QT interval lengths of healthy and diseased patients were determined to be 395 ± 28.66 ms and 391 ± 38.79 ms, respectively.

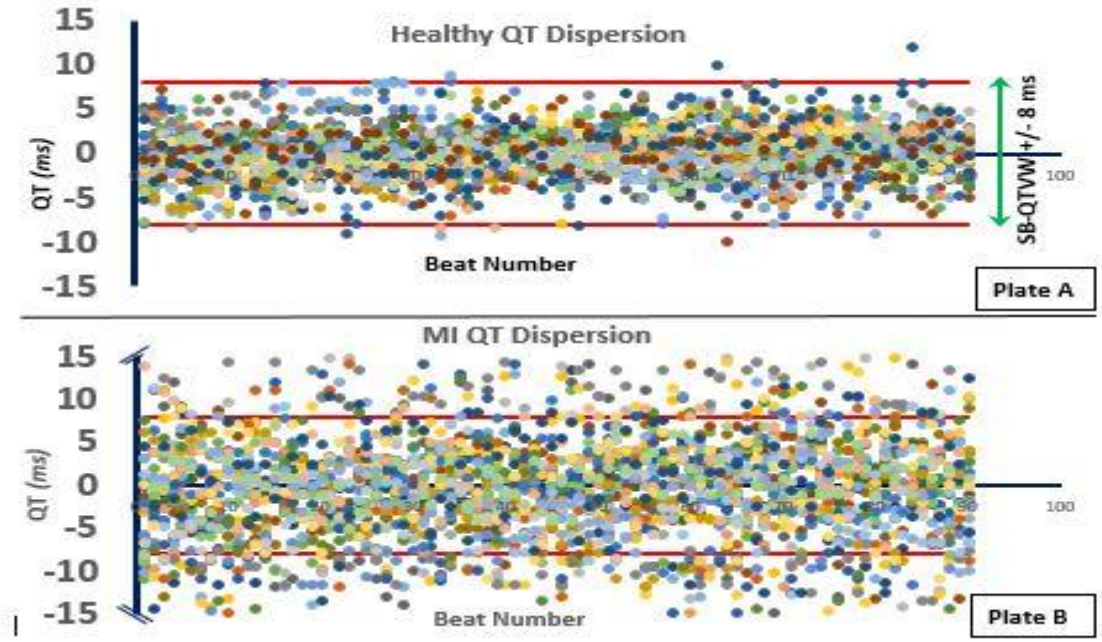

Figure 13: 90 beat QT interval dispersion comparison. Mean QT interval lengths of healthy were 395±28.66 ms, and diseased patients were 391± 38.79 ms, respectively. Plate A, top panel, shows healthy QT intervals. Plate B, bottom panel, shows QT interval dispersion of MI patients on a beat to beat basis. Healthy controls' QTV appears constrained within a window (SB-QTVW +/-8ms), while diseased patients' QTV exceeds the window. MI QT dispersion extended to +/-150ms. However, for purpose of visual comparison we truncated outliers beyond +/- 15ms.

To further explore the QTV, we conducted histogram analysis of our data. As figure 14 shows, 99.7% of healthy controls' QTV data was within ±8 ms window. 87.5% of healthy controls' QTV data was within -2 and +4 ms window. As expected, diseased patients' QTV data is spread wider than that of healthy controls. 99.2% of diseased QTV data falls within -20 ms to +40 (See figure 14). 87.5% of diseased patients' data fell within ±10 ms. Healthy and diseased QTV means were 0.457 and -0.194 ms, and standard deviation 2.48 and 8.23 ms, respectively (See figure 14).

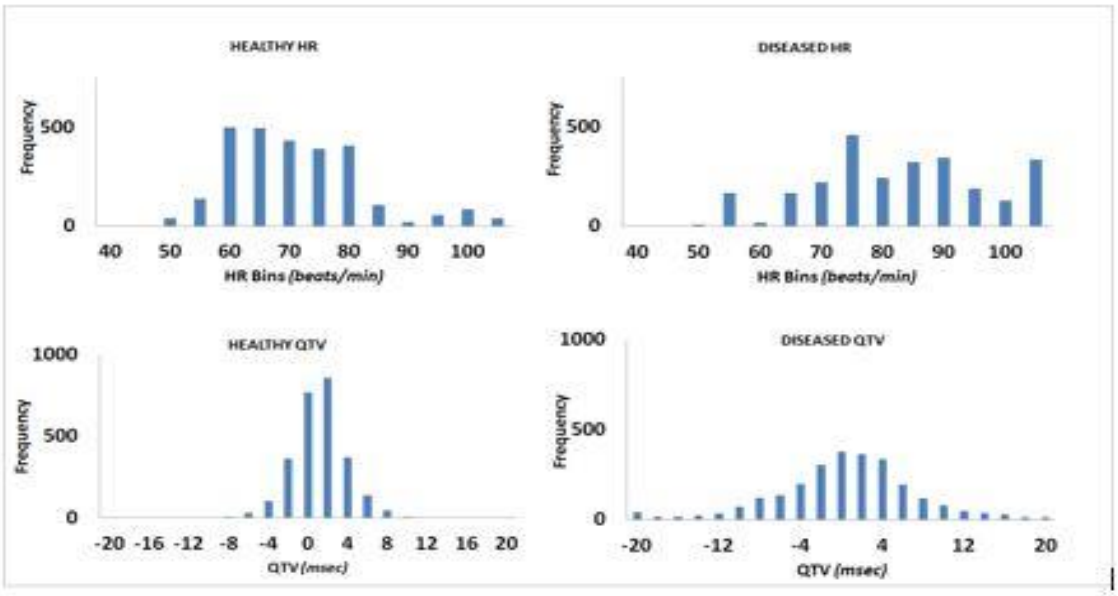

Figure 14: HR and QTV histogram comparison. VM derived HR and QTV histogram comparison of healthy and diseased patients. Total number of 2700 and 2610 HR and QTV data points for healthy and diseased patients respectively used for analysis.

HR histogram analysis showed healthy controls centered at approximately 70 beats per minute. Diseased patients were centered at approximately 80 beats per minute (See figure 14). Healthy and diseased mean HR was 68.6 and 80.5 beats per minute (bpm), and standard deviation 10.83 and 14.5 bpm, respectively.

QTVI analysis of our data shows diseased patients having elevated variability numbers than healthy controls. Healthy controls' index numbers ranged from -1 to -3.5, while diseased patients' index numbers ranged from -2.8 to +1.5 (See figure 15). Healthy and diseased mean QTVI was -1.825 ± 0.584, and -0.4595 ± 0.894, respectively.

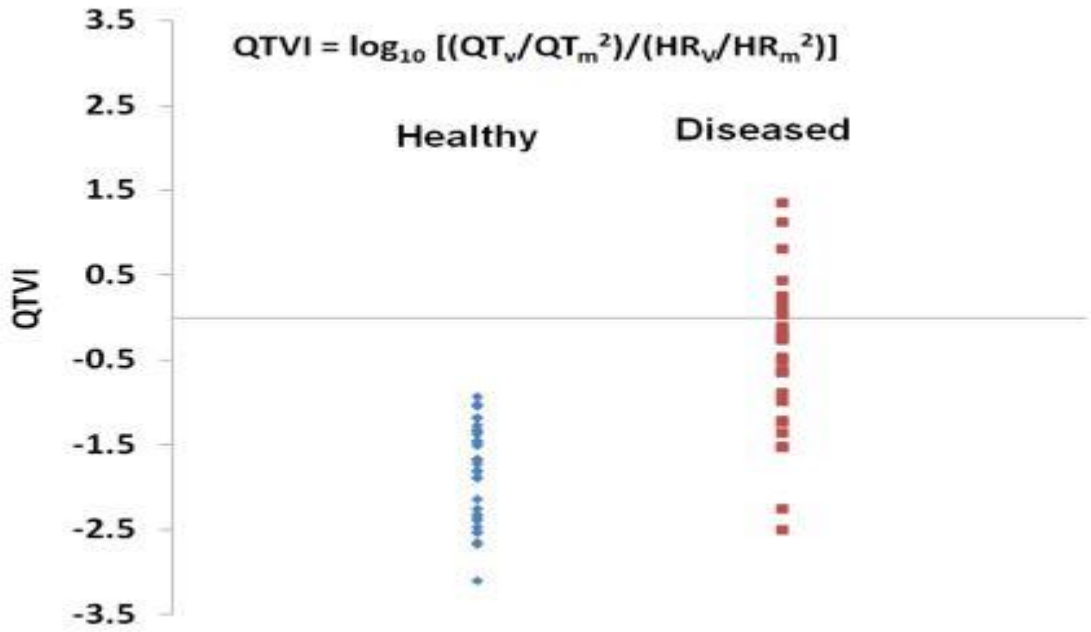

Figure 15: QTVI. VM derived QTVI comparison of healthy and diseased patients. Total number of 30 healthy controls and 29 diseased patients respectively were used for analysis. Diseased patients show elevated QTVI. Healthy and diseased mean QTVI were -1.825 ± 0.584 and -0.4595 ± 0.894 respectively.

Comparing average QTV standard deviations of healthy versus diseased patients, we noticed a significant difference ($p<0.01$). Average standard deviations of healthy controls and diseased patients were 2.48 and 8.23 ms respectively (See figure 16). Overall average means did not depict a significant difference. Average means for healthy and diseased patients were 0.0457 and -0.1934 ms respectively.

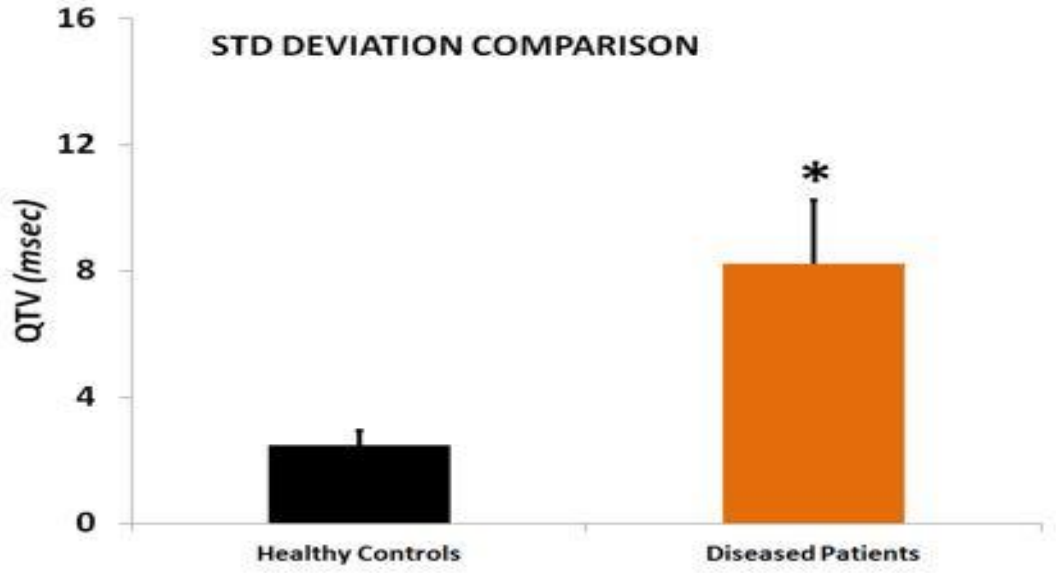

Figure 16: Standard deviation comparison of VM derived QTV of healthy and diseased patients. Total number of 30 and 29 healthy and diseased patients respectively used for analysis. Healthy and diseased QTV standard deviations were 2.48 ± 0.45 and 8.23 ± 2.0, respectively.

Sensitivity and specificity analysis showed a 93.1% sensitive and 86.7% specific test using ±8 ms cutoff for SB-QTVW. Overall results show 26 true negatives (healthy controls), 4 false positives (mis-identified healthy controls), 27 true positives (diseased patients), and 2 false negatives (mis-identified diseased patients) (See table 3).

Table 3: Sensitivity and Specificity*

RESULTS	DISEASED	HEALTHY
POSITIVE	27 (True Positives)	4 (False Positives)
NEGATIVE	2 (False Negatives)	26 (True Negatives)

*Sensitivity (93.10%) and specificity (86.67%) test.

A ROC curve analysis appears to optimize at a value of ±9.5 ms with a sensitivity of 93.3% and specificity of 89.7%. With a ±9.5 ms (19 ms) window, we achieve 26 true positives, 2 false positives, 28 true negatives and 3 false negatives (See figure 17).

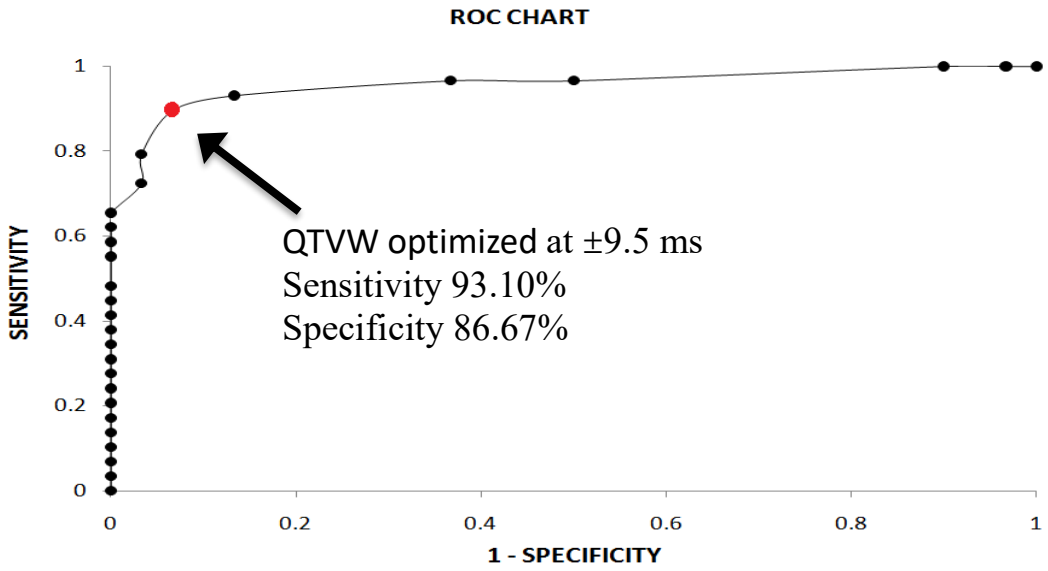

Figure 17: ROC analysis. VM derived results. True positive rate (Sensitivity) vs false positive rate (1 – Specificity). Results optimized at ±9.5 ms.

Further analysis using SPSS' logistic regression using maximum QTV for each patient produces a 91.5% overall predictability, with a sensitivity of 93.3% and specificity of 89.7%. The log-odd equation has QTV coefficient calculated to 0.868 and intercept at -8.370.

Pseudo-QTVW

Pseudo-vector magnitude analysis from the standard 12 lead ECG produced results similar to those found using vector magnitude from the Frank XYZ system, albeit with a different QTVW size. From the previously selected list of patients, a total of 29 healthy controls and 25 diseased patients were selected for analysis using pVM. Reduction in numbers was primarily due to excessive noise in the lead signals for the rejected patients.

Using the ROC analysis, the pVM derived QTVW optimized at a value of ±13 ms (see figure 18); however, a value of ±10 was chosen as a value more suited for a screening test producing a sensitivity and specificity of 100% and 76% respectively; with 7 false positives and no false negatives. Figure 18 shows the QTVW and can be compared to figure 17 for similarities. The size of the QTVW depends on the specific purpose for the test.

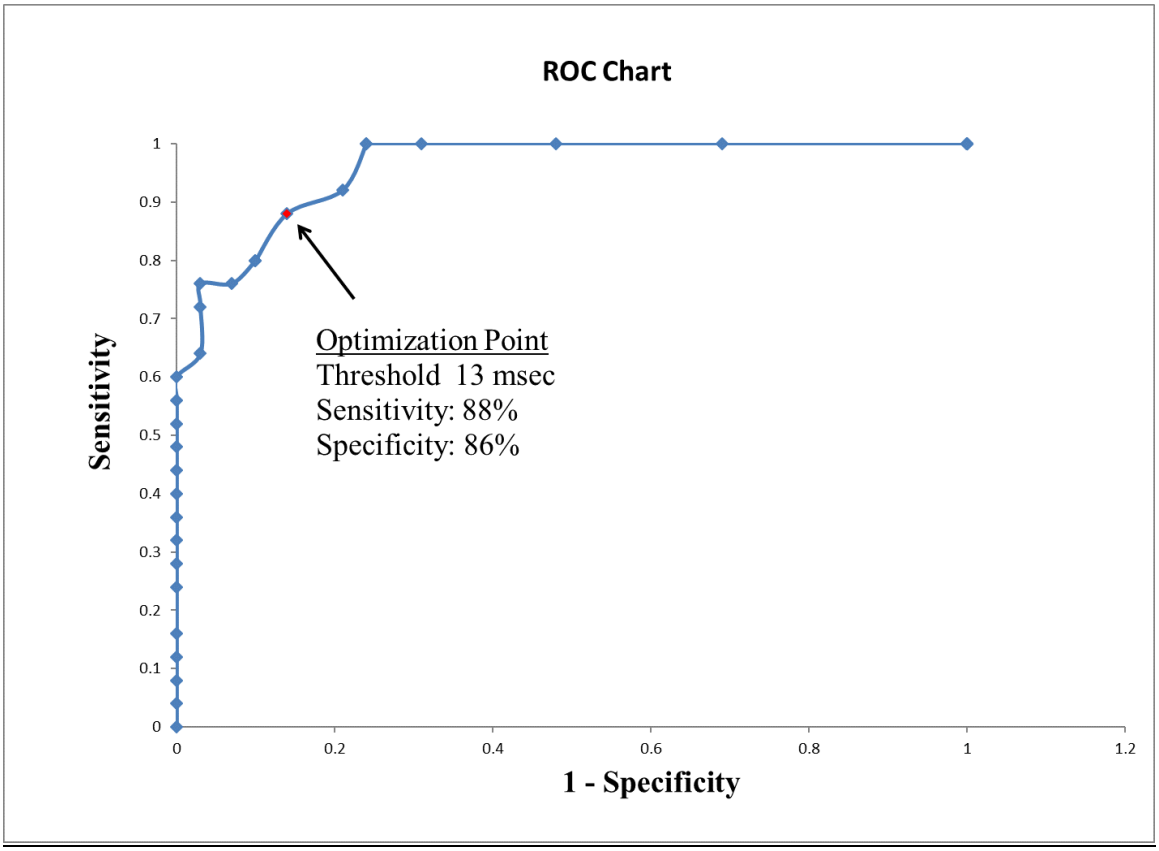

Figure 18: ROC analysis. pVM derived results. True positive rate (Sensitivity) vs false positive rate (1 – Specificity). Results optimized at ±13 ms.

Table 4: Diagnostic Implications of QTVW Size

QTVW	SENS	SPEC	PPV	NPV
10	100%	76%	78%	100%
11 or 12	88%	76%	77%	88%
13	84%	86%	84%	86%
14	80%	90%	87%	84%
15	76%	90%	86%	81%

Histogram analysis of pVM showed that 99.0% of data points lie within ±10 ms window; with 87.5% data points falling within +6 and -2 ms window for healthy controls. 97.5% of diseased QTV data fell within +60 and -20 ms window; with 87.5% QTV data falling within +40 and -12 ms (figure 19). A comparison of the standard error for QTV for healthy and ischemic patients demonstrates that there is increase variance between subjects in the QT interval in the diseased state (figure 18).

Selecting a pQTVW of 10 msec was more useful when ruling out the disease than ruling in the disease, as that yielded a negative predictive value of nearly 100%. Optimization of a test for sensitivity and specificity parameters are defined based on the needs of a particular test. The advantage of selecting a lower pQTVW value sets the test better as a diagnostic test than as a confirmatory one: the test increases in sensitivity. In general, the less the QT variability, more the subsequent confidence in the test that ischemia is not preset; and, vice versa, the greater the QT variability, the more the likelihood for presence of ischemia.

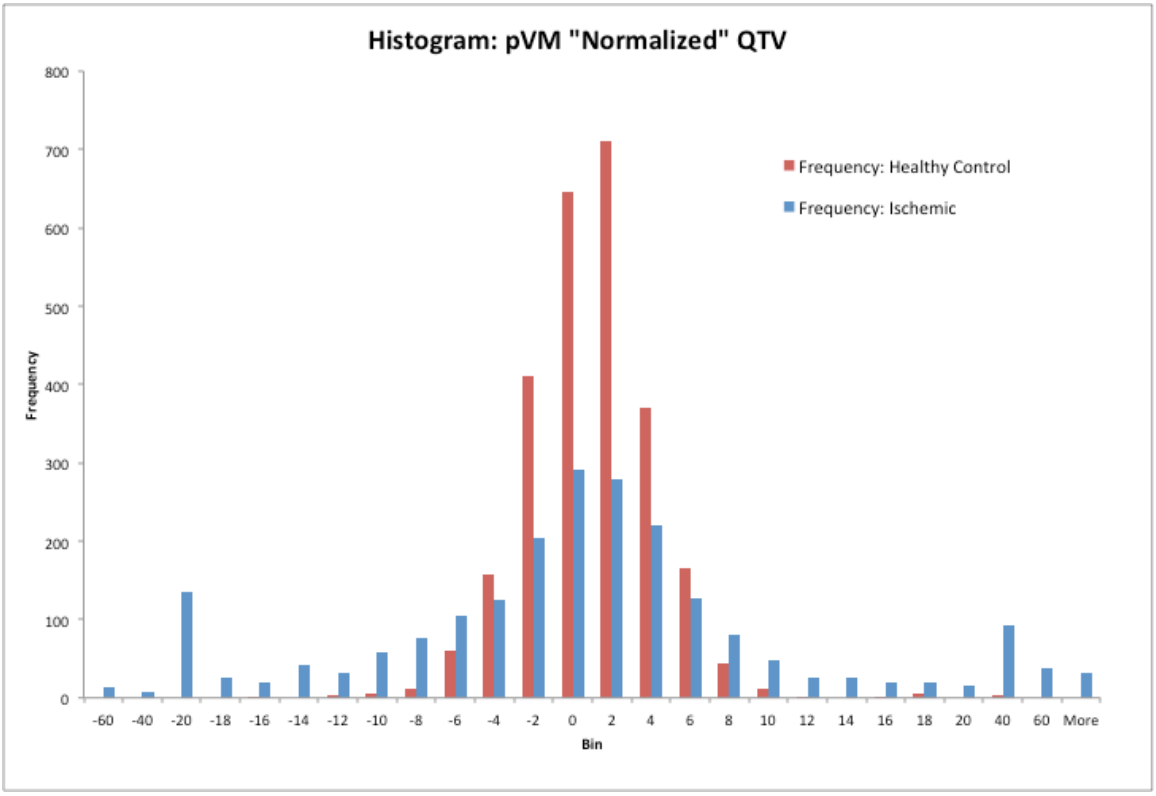

Figure 19: HR and QTV histogram comparison. pVM derived HR and QTV histogram comparison of healthy and diseased patients. Total number of 2700 and 2610 HR and QTV data points for healthy and diseased patients respectively used for analysis.

Machine Learning Classification Results

The results indicated that the Decision tree (C5-DT) method is the best predictor with 98.31% accuracy on the holdout sample. Decision tree (C5-DT) and ensemble method had accuracy of 98.31%, however the sensitivity is 100% and 96.55% for C5-DT, Support Vector Machine (SVM) with 88.14% accuracy and the Artificial Neural Network models has the lowest accuracy of 86.4% among the selected method for the study. Sensitivity and specificity of SVM is the second highest, however the accuracy is lower

by ~ 10%. Refer to Table 5 for accuracy, sensitivity, and specificity summarizing details of classification results: accuracy, sensitivity, and specificity for all ML methods utilized for the study. Refer to equation 1-3 showing expressions used to calculate the measures of classification results. 10-fold cross validation is applied to the data set to predict the performance of machine learning classification techniques.

Table 5: Confusion Matrix-Machine Learning classification performance

Method	**Confusion Matrix** (Rows: predictions)		**Accuracy**	**Sensitivity**	**Specificity**
Decision Tree (C5)	30	0	98.31	100.00	96.55
	1	28			
SVM	28	2	88.14	93.33	82.76
	5	24			
ANN	27	3	86.44	90.00	82.76
	5	24			
Ensemble	29	1	98.31	96.67	100.00
	0	29			

Predictor Importance

One of the main criticisms of machine learning methods is the inability to explain the results and the underlying mechanism used to derive the classification result. Refer to Tables 6 and 20 showing predictor (or feature) importance for all models used in the

study. Predictor importance is simply ratio of the model error (R2) without the predictor variable and model error (R2) with all predictor variable included in the model [55].

Table 6: Predictor Importance (Normalized)

	ANN	SVM	C5-DT	Ensemble	FUSED
QT_StDev	1.000	1.000	0.844	1.000	1.000
QT_Range	0.421	0.686	1.000	0.643	0.715
RR_StDev	0.353	0.711	0.000	0.286	0.351
RR_Max	0.242	0.619	0.000	0.286	0.298
QT_Max	0.369	0.260	0.000	0.500	0.294
RR_Range	0.150	0.128	0.324	0.286	0.231
RR_Var	0.266	0.154	0.000	0.286	0.183
RR_Min	0.160	0.279	0.000	0.214	0.170
RR_Avg	0.207	0.274	0.000	0.071	0.144
QT_Var	0.124	0.196	0.000	0.214	0.139
QT_Min	0.172	0.092	0.000	0.000	0.069
QT_Avg	0.000	0.000	0.000	0.000	0.000

Predictor importance results where normalized for easy comparison between different methods and relativity importance of predictor. Table 6 summarizes the normalized predictor importance results for all models. The central idea behind the predictor importance is that higher particular variable contributes to the accuracy of the

model, higher would the ratio of error (importance) in the absence of the predictor variable and consequently decreased performance.

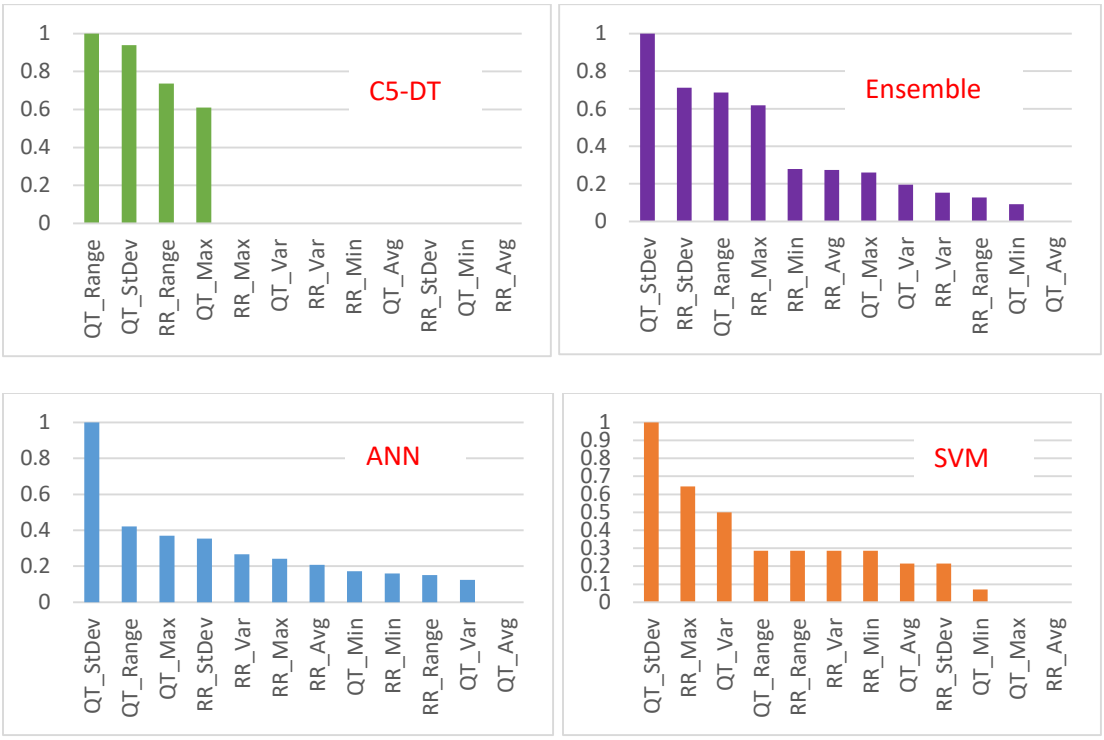

Figure 20: Relative importance of variables—per model type

QT features are more prominent e.g. QT_StDev is most importance predictor for ANN, SVM and Ensemble method and second important predictor for Decision Tree (C5) method. QT_Range is in top four importance predictors for all models. Three out of four top predictor are QT features for C5, SVM and ANN. Decision Tree model with highest accuracy (98.3%), Sensitivity (100%) and Specificity (96.55%) has only four predictor variable that have normalized importance above zero (0) and with importance indictor value above "0.5". (Please refer to 20). Results indicate that QT variability (or its features) bears more predictability value than RR variability. Model when studied independently with each indicator features (RR and QT) resulted in a lower accuracy,

sensitivity, and specificity. In conclusion, we can conclude with confidence that "QT" is a more importance predictor of heart health and adding variable features in addition to "RR" increases results with sensitivity, specificity, and accuracy between 96 % to 100%.

Classification Performance

Accuracy is number of correct prediction results by a given model; however, accuracy can be misleading as it includes correct prediction of positive and negative outcomes (Refer to equation (1)). Hence the need for sensitivity for correct prediction of favorable (true, positive) outcome and need for specificity for correct prediction of unfavorable (false, negative,) outcome.

We can see from Table 5, that although the accuracy from model C5 and Ensemble are equal (98.31% / 98.31%), sensitivity (100% / 90%) and specificity (96.55% / 82.76%) measures are lower for Ensemble model. We chose ensemble model to compare weighted prediction accuracies of different models to enhance confidence in our results. First phase or input phase includes pre-processed data after merging, aggregating, cleaning, selecting right patients, and finally transformed data to use for analysis via ANN, SVM, DT/C5, and ensemble methods. Second phase of the process or Training & Testing phase splits the 10 folds sample (Refer to figure 12) for developing the model for training as each fold or sample of data is processed and simultaneously calibrating of the model. All methods (ANN, SVM, C5, and Ensemble) follow the same process of training and testing. Calibrated model is then used to test the model and finally calculates the predictor importance for each model. Predictor importance is calculated to test the model

performance in absence of the predictor variables. Final phase or Output phase compiles all the data to output the classification result (Table 5: Confusion Matrix-Machine Learning classification performance and normalized predictor importance summary (Table 6) to help make conclusion from the model analysis.

As stated earlier machine learning helps us detect patterns; rules in an unstructured data and the ability of machine learning method to detect relationship between variables helps refine prediction accuracy. In our study we learned that C5-decision tree used 66% less predictor variables as compared to other model and generated highest accuracy of correctly prediction true and false classes. We started with the hypothesis that QT interval enhances the accuracy of the prediction for diagnosis of myocardial ischemia as studies have established an abnormal dynamicity of QT interval during ischemic conditions [8] and from predictor importance results, indicators representing QT interval, show higher contribution towards increased accuracy of the prediction class.

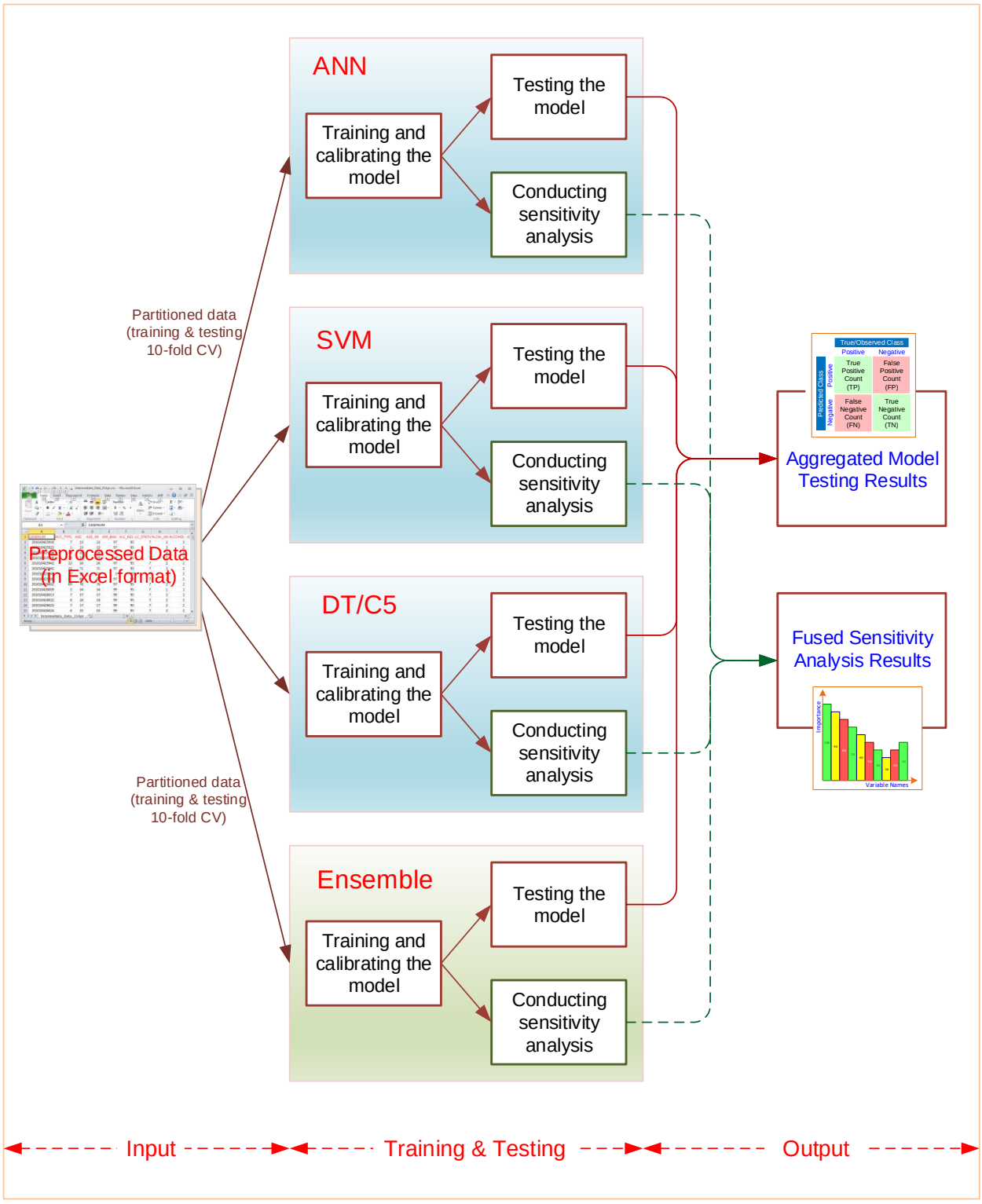

Figure 21: A Graphical Depiction of the Model Building and Testing Process

CHAPTER V

CONCLUSION

SB-QTVW

Myocardial ischemia/infarction alters ventricular repolarization in various ways. Studies have shown that this repolarization variability is depicted in the QT interval variations of surface electrocardiograms. We set about to devise a method by which this variability can be gauged between healthy and diseased patients. We utilized three dimensional vector magnitude derived from Frank XYZ leads, instead of standard 12 lead EKG, to overcome quantitative limitations and lead selection issues. We also analyzed QT variability independent of HR to avoid limitations incurred due to hysteresis effect between the two factors. Our goal was to achieve two distinct and identifiable groups: healthy and diseased, based on their QT variations. Our initial experiments showed promise and further experimentation led to a criterion we named as Sewani-Benjamin QT Variability Window (SB-QTVW)—a window of time representing the maximum variations of QT intervals in healthy heart EKG. Our results show that it is possible to parse the two groups based on SB-QTVW.

Results from 30 healthy and 29 diseased patients supported our preliminary studies. While using three times the standard deviation for the QTV for healthy controls gave us a cutoff of 7.44 ms, choosing ±8 ms not only adapts well to contemporary

electrocardiographic equipment millisecond resolution, but also encompasses 99.7% of healthy QTV data. We investigated the reasons that caused SB-QTVW to misclassify patients, resulting in false positives and false negatives. However, limitations in the type and amount of data available constrained this analysis.

It is interesting to note that the ROC curve optimized at 9.5 ms. With a 9.5 ms selection, the sensitivity and specificity improves by 0.2% and 3% respectively. However, since ±8 ms encompass 99.7% healthy controls, we opted to retain the more conservative definition for our upper and lower window limit.

With the window set with limits as stated above, healthy controls mostly fell within a SB-QTVW of 16 ms (±8 ms) and QT variability for ischemic patients frequently exceeded the SB-QTVW of 16 ms. Furthermore, in a considerable number of cases, our algorithm was able to differentiate between the healthy controls and ischemic patients using as few as 20 consecutive beats.

HR histogram analysis showed diseased patients with a significantly higher heart rate than healthy (mean HR of 80.5 vs 68.6 bpm; $p<0.01$) (See figure 14). This may suggest a shift in autonomic regulation favoring sympathetic tone in presence of ischemia or infarct. The average standard deviation comparison of healthy controls versus diseased patients showed a significant difference. This difference provides an indication of increased QT variability in diseased patients as compared to healthy controls. However, no significant difference was observed between healthy controls and diseased patients while comparing overall mean lengths of QT intervals. This may lend to the notion that QT interval length by in itself may not be the significant marker, unless it's beyond FDA recommended corrected values (>500 ms) [15]. Histogram analysis of our

data does provide us that notion that HR and QTV vary differently than each other. Hence, coupling the two may be problematic.

QTVI

QTVI analysis does depict elevated QT variability. However, we did not observe a clear threshold for healthy versus diseased patients (See figure 13.) Furthermore, no significant difference in QT variability means suggests that a single summarizing number index may be unable to capture the beat-to-beat dynamics of QT variability—which is the basis of our argument in devising the alternate SB-QTVW methodology to analyze beat-to-beat QT variability dynamics. Therefore, we believe that our SB-QTVW is more effective in providing tangible thresholds to separate healthy and diseased patients.

Ectopic Beats

Although we rejected patients with ectopic beats, we did conduct a limited study of those patients' data. We observed that the QT intervals of ectopic beats exceeded the SB-QTVW bounds as well. Furthermore, the pattern of beat-to-beat QT variation data dispersion immediately before, during and following the ectopic beats appeared different than ischemic patients. We surmised that the dispersion of QTV data in itself may contain characteristic signatures which may pinpoint to the likely probability of pathologies. This obviously is a vast subject in itself and may warrant future investigations.

Post-Cath Analysis

Since decreased circulation due to acute ischemia or infarct showed increased QT variability, out of curiosity, we analyzed available post-cath EKG data from of our selected patients to see if following intervention and resolution of coronary circulation, any decrease in QT variability was observed as well. Data for only two patients from our selected patient group was available. Both patients were deemed "resolved" by our team cardiologist still showed excursions outside the SB-QTVW in their post-cath EKG data. However, we did observe a reduction in their overall QT variability: standard deviation pre-cath and post-cath of 8.99 and 4.54, and 24.85 and 6.26 for the two patients respectively.

Statistical Analysis

Statistical analysis and optimizations support our hypothesis to be correct and increase confidence in use of our SB-QTVW system to differentiate healthy and ischemic patients. However, our data was limited to stabilized, supine patients. Non-stationarity, resulting from exercise or data extracted from holter readings, may not necessarily reflect the same SB-QTVW rule. We also excluded arrhythmic, ectopic beat, bundle branch block, and other patients not categorized as ischemic or infarct.

It was interesting to observe that healthy controls' data depicted range of QT variability within the window. This may lend to a variable threshold SB-QTVW

depicting varied levels of coronary health. However, to achieve a correlation between various thresholds and percentage occlusion will need intensive clinical data and analysis.

Conclusion and Future Work

Further studies may be needed to assure validity and confidence in SB-QTVW system. Future controlled studies in clinical settings may also provide beneficial understanding of various levels of SB-QTVW thresholds and their clinical significance. If in fact the SB-QTVW system holds validity, it may be able to provide an independent novel alternative criterion to STEMI, and Non-STEMI, for diagnosis of ischemia and infarction. Additionally, further studies within clinical settings, may deem patients falling between 2 and 3 standard deviations in risk zone and target of pre-emptive intervention.

Therefore, we believe that our SB-QTVW system warrants further studies to validate its performance under various clinical conditions.

Machine Learning

As the results showed, we have solutions via ML techniques which can correctly identify MI and healthy patients. Although computer assisted, this methodology now provides an alternate to ST elevation criteria on a limited basis. The accuracy is 98.31%, sensitivity is equal to 100% and specificity equal to 96.55% with the C5 algorithms. Other methods did not have as high accuracy as C5, however showed more than 90%

sensitivity (ability to correctly predict true positive class.) Although we limited our initial study, we feel that it is an encouraging initial step towards developing confidence in our methodology and proof of concept. Furthermore, results of this study show that QT variability bears far more importance than RR variability in predicting coronary health condition. RR variability, though heavily researched so far, proves to be a lesser predictor than the QT variability. However, combined variables predict the best overall results. Therefore, the synergistic value of the two variables combined is proved with this study; and we believe that this approach would be the best for future heart health predictability studies.

Data Collection

In medical research data fidelity is very important, especially when dealing with life threatening disease such as one considered by this study—myocardial infarction. However, it is extremely difficult to find clean data set for such diagnostics. We started with a total of 295 patients' records, and, after exclusion criteria, were left with only 59 patient records that were effective for our analysis. Although the sample size of 59 was limited, it isn't unusual in clinical data mining settings since in clinical settings, each set of corresponding data points bears a lot more meaning and importance compared to say a marketing or business dataset [56]. Using 10-fold cross-validation reduced some of our concerns related to bias and overfitting, we believe a larger dataset may provide increased confidence to support our research findings. Physionet's PTB database [43] provided de-identified data. Hence, our study process did not include an IRB approval process.

Classification Methods

Our research explored various literatures on prediction of heart failure or diagnosis. Recent effort has shown tremendous influence of machine learning methods to predict or diagnose the heart health. Refer to Table 7 for all methods used for comparing our selected model for accuracy, sensitivity, and specificity.

Table 7. Ensemble- Machine Learning Classification Model

C5
Logistic regression
Decision List
Bayesian Network
Discriminant
KNN Algorithm
SVM
Neural Net
C&R Tree
Quest
CHAID

We believe past research have focused on comparing different techniques and optimizing methods. However, clean and cost effective cardiac variable selection to use for analysis has been absent. It had been widely established that in data mining or any other such analytical process: “garbage in and garbage out”. We ensured to focus our

study to ensure that raw data, although medical grade, was pre-processed before it was used for classification prediction. Data preparation built our confidence in including only selected and the most preferred methods. Moreover, we also ran Ensemble method—which is an aggregate of several other methods—to ensure choice of best methodologies amongst several available options.

Conclusion and Future Work

Our study shows that, given appropriate emphasis on predictor selection, ensuring high fidelity of data especially when we are dealing with human health prediction and diagnosis, we can develop a high accuracy of prediction, along with high sensitivity and specificity. High accuracy (98.31%) and sensitivity (100%) increases the enthusiasm to develop non-invasive diagnostics methods that can help in managing and may reverse the trend of coronary artery disease [57] as the number one killer of humans with timely treatment. We started with the hypothesis that QT interval enhances the accuracy of the prediction for diagnosis of myocardial ischemia, as studies have established an abnormal dynamicity of QT interval during ischemic conditions [8]; and from predictor importance results, indicators representing QT interval, show higher contribution towards increased accuracy of the prediction class. We also believe that VCG derived QT interval also aided towards the increased accuracy of the results. Furthermore, we also showed that QT variability bears a higher intrinsic value as compared to RR variability—a variable highly researched hitherto in studies.

Above observations and conclusion strengthens our conviction to continue the research and we plan to collect a larger sample of patient data that would help broaden our current study to include additional patient condition and reduce limitations placed in this study. Generalizing our results to a wider patient population will provide our method a wider acceptance as a clinical tool.

REFERENCES

1. Association, A.H., *2021 Heart Disease and Stroke Statistics Update Fact Sheet At-a-Glance* 2021.
2. Herring, N. and D.J. Paterson, *ECG diagnosis of acute ischaemia and infarction: past, present and future.* QJM, 2006. **99**(4): p. 219-30.
3. Ebell, M.H., D. Flewelling, and C.A. Flynn, *A systematic review of troponin T and I for diagnosing acute myocardial infarction.* J Fam Pract, 2000. **49**(6): p. 550-6.
4. Frank, E., *An accurate, clinically practical system for spatial vectorcardiography.* Circulation, 1956. **13**(5): p. 737-49.
5. Couderc, J.P., *Cardiac regulation and electrocardiographic factors contributing to the measurement of repolarization variability.* J Electrocardiol, 2009. **42**(6): p. 494-9.
6. Huang, M.H., et al., *Effects of transient coronary artery occlusion on canine intrinsic cardiac neuronal activity.* Integr Physiol Behav Sci, 1993. **28**(1): p. 5-21.
7. Murabayashi, T., et al., *Beat-to-beat QT interval variability associated with acute myocardial ischemia.* J Electrocardiol, 2002. **35**(1): p. 19-25.
8. Berger, R.D., *QT variability.* J Electrocardiol, 2003. **36 Suppl**: p. 83-7.

9. Berger, R.D., et al., *Beat-to-beat QT interval variability: novel evidence for repolarization lability in ischemic and nonischemic dilated cardiomyopathy.* Circulation, 1997. **96**(5): p. 1557-65.

10. Malik, M., *Beat-to-beat QT variability and cardiac autonomic regulation.* Am J Physiol Heart Circ Physiol, 2008. **295**(3): p. H923-H925.

11. Downar, E., M.J. Janse, and D. Durrer, *The effect of acute coronary artery occlusion on subepicardial transmembrane potentials in the intact porcine heart.* Circulation, 1977. **56**(2): p. 217-24.

12. Clerico, A., et al., *High-sensitivity troponin: a new tool for pathophysiological investigation and clinical practice.* Adv Clin Chem, 2009. **49**: p. 1-30.

13. Coronel, R., et al., *Distribution of extracellular potassium and its relation to electrophysiologic changes during acute myocardial ischemia in the isolated perfused porcine heart.* Circulation, 1988. **77**(5): p. 1125-38.

14. Couderc, J.P., *Measurement and regulation of cardiac ventricular repolarization: from the QT interval to repolarization morphology.* Philos Trans A Math Phys Eng Sci, 2009. **367**(1892): p. 1283-99.

15. Food and H.H.S. Drug Administration, *International Conference on Harmonisation; guidance on E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs; availability. Notice.* Fed Regist, 2005. **70**(202): p. 61134-5.

16. Han, J. and G.K. Moe, *Nonuniform Recovery of Excitability in Ventricular Muscle.* Circ Res, 1964. **14**: p. 44-60.

17. Hirche, H., et al., *Myocardial extracellular K+ and H+ increase and noradrenaline release as possible cause of early arrhythmias following acute coronary artery occlusion in pigs.* J Mol Cell Cardiol, 1980. **12**(6): p. 579-93.

18. Janse, M.J., et al., *Variability of recovery of excitability in the normal canine and the ischaemic porcine heart.* Eur Heart J, 1985. **6 Suppl D**: p. 41-52.

19. Janse, M.J. and A.G. Kleber, *Electrophysiological changes and ventricular arrhythmias in the early phase of regional myocardial ischemia.* Circ Res, 1981. **49**(5): p. 1069-81.

20. Janse, M.J. and A.L. Wit, *Electrophysiological mechanisms of ventricular arrhythmias resulting from myocardial ischemia and infarction.* Physiol Rev, 1989. **69**(4): p. 1049-169.

21. Katz, A.M., *Membrane-derived lipids and the pathogenesis of ischemic myocardial damage.* J Mol Cell Cardiol, 1982. **14**(11): p. 627-32.

22. Kleber, A.G., *Resting membrane potential, extracellular potassium activity, and intracellular sodium activity during acute global ischemia in isolated perfused guinea pig hearts.* Circ Res, 1983. **52**(4): p. 442-50.

23. Kleber, A.G., et al., *Mechanism and time course of S-T and T-Q segment changes during acute regional myocardial ischemia in the pig heart determined by extracellular and intracellular recordings.* Circ Res, 1978. **42**(5): p. 603-13.

24. Kleber, A.G., et al., *Changes in conduction velocity during acute ischemia in ventricular myocardium of the isolated porcine heart.* Circulation, 1986. **73**(1): p. 189-98.

25. Kodama, I., et al., *Combined effects of hypoxia, hyperkalemia and acidosis on membrane action potential and excitability of guinea-pig ventricular muscle.* J Mol Cell Cardiol, 1984. **16**(3): p. 247-59.

26. Lamfers, E.J., et al., *Prehospital versus hospital fibrinolytic therapy using automated versus cardiologist electrocardiographic diagnosis of myocardial infarction: abortion of myocardial infarction and unjustified fibrinolytic therapy.* Am Heart J, 2004. **147**(3): p. 509-15.

27. Bruce, R.A. and J.R. McDonough, *Stress testing in screening for cardiovascular disease.* Bull N Y Acad Med, 1969. **45**(12): p. 1288-305.

28. Lehtinen, R., *ST/HR hysteresis: exercise and recovery phase ST depression/heart rate analysis of the exercise ECG.* J Electrocardiol, 1999. **32 Suppl**: p. 198-204.

29. Conrath, C.E. and T. Opthof, *Ventricular repolarization: an overview of (patho)physiology, sympathetic effects and genetic aspects.* Prog Biophys Mol Biol, 2006. **92**(3): p. 269-307.

30. Bahit, M.C., et al., *Thresholds for the electrocardiographic change range of biochemical markers of acute myocardial infarction (GUSTO-IIa data).* Am J Cardiol, 2002. **90**(3): p. 233-7.

31. Levites, R., et al., *Effects of procainamide on the dispersion of recovery of excitability during coronary occlusion.* Circulation, 1976. **53**(6): p. 982-4.

32. McCallister, L.P., S. Trapukdi, and J.R. Neely, *Morphometric observations on the effects of ischemia in the isolated perfused rat heart.* J Mol Cell Cardiol, 1979. **11**(7): p. 619-30.

33. Michaelides, A.P., et al., *New coronary artery disease index based on exercise-induced QRS changes.* Am Heart J, 1990. **120**(2): p. 292-302.

34. Minisi, A.J. and M.D. Thames, *Activation of cardiac sympathetic afferents during coronary occlusion. Evidence for reflex activation of sympathetic nervous system during transmural myocardial ischemia in the dog.* Circulation, 1991. **84**(1): p. 357-67.

35. O'Connor, R.E., et al., *Part 10: acute coronary syndromes: 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care.* Circulation, 2010. **122**(18 Suppl 3): p. S787-817.

36. Delen, D., G. Walker, and A. Kadam, *Predicting breast cancer survivability: a comparison of three data mining methods.* Artif Intell Med, 2005. **34**(2): p. 113-27.

37. Jonsdottir, T., et al., *The feasibility of constructing a Predictive Outcome Model for breast cancer using the tools of data mining.* Expert Systems with Applications, 2008. **34**(1): p. 108-118.

38. Alizadehsani, R., et al., *A data mining approach for diagnosis of coronary artery disease.* Comput Methods Programs Biomed, 2013. **111**(1): p. 52-61.

39. Intelligence, S.A., *What is Machine Learning?* 2007.

40. Abdo, A.A., et al., *The Fermi Gamma-Ray Space Telescope discovers the pulsar in the young galactic supernova remnant CTA 1.* Science, 2008. **322**(5905): p. 1218-21.

41. Abson, A., et al., *Chemistry of pseudomonic acid. Part 16. Aryl and heteroaryl ketone derivatives of monic acid.* J Antibiot (Tokyo), 1996. **49**(4): p. 390-4.

42. Dangare, H.M., et al., *Cri du chat syndrome: a series of five cases.* Indian J Pathol Microbiol, 2012. **55**(4): p. 501-5.

43. Goldberger, A.L., et al., *PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.* Circulation, 2000. **101**(23): p. E215-20.

44. BIOPAC Systems, I., *Acqknowledge*. 42 Aero Camino, Goleta, CA.

45. Feeny, A., L. Han, and L.G. Tereshchenko, *Repolarization lability measured on 10-second ECG by spatial TT' angle: reproducibility and agreement with QT variability.* J Electrocardiol, 2014. **47**(5): p. 708-15.

46. Abanto, C., et al., *Predictors of functional outcome among stroke patients in Lima, Peru.* J Stroke Cerebrovasc Dis, 2013. **22**(7): p. 1156-62.

47. Darpo, B., et al., *Man versus machine: is there an optimal method for QT measurements in thorough QT studies?* J Clin Pharmacol, 2006. **46**(6): p. 598-612.

48. Darpo, B., et al., *Cardiac Safety Research Consortium: can the thorough QT/QTc study be replaced by early QT assessment in routine clinical pharmacology studies? Scientific update and a research proposal for a path forward.* Am Heart J, 2014. **168**(3): p. 262-72.

49. Dash, Y., *Maintainability Prediction of Object Oriented Soware System by Using Artificial Neural Network Approach.* International Journal of Soft Computing and Engineering, 2012. **2**(2): p. 420-423.

50. Kanters, J.K., et al., *Short- and long-term variations in non-linear dynamics of heart rate variability.* Cardiovasc Res, 1996. **31**(3): p. 400-9.

51. Zebrowski, J.B., R, *Nonlinear Instabilities and Nonstationarity in Human Heart-Rate Variability.* Computing in Science & Engineering, 2004. **6**: p. 78-83.

52. Challis, R.E., Kitney, R.I., *Biomedical signal processing (in four parts).* Med. Biol. Eng. Comput., 1991. **29**: p. 1-17.

53. Kohavi, D., et al., *Adsorption of salivary proteins onto prosthetic titanium components.* J Prosthet Dent, 1995. **74**(5): p. 531-4.

54. IBM Corp., *IBM SPSS Statistics for Windows, Version 19.0.* Armonk, NY.

55. Davis, F., *Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology*. MIS Quarterly, 1989. **13**(3): p. 319-340.

56. Bellazzi, R., *Telemedicine and diabetes management: current challenges and future research directions.* J Diabetes Sci Technol, 2008. **2**(1): p. 98-104.

57. World Health Organization. *Cardiovascular disease Fact sheet N*317*. January 2015 [cited 2015 January 30th]; Available from: http://www.who.int/mediacentre/factsheets/fs317/en/.

VITA

Rahim Ruknudin Sewani, D.O.

Candidate for the Degree of

Doctor of Philosophy

Dissertation: SB-QTVW: AN ALTERNATE INDEX TO DIFFERENTIATE BETWEEN HEALTHY AND ACUTE ISCHEMIC/INFARCTED PATIENTS USING VECTOR MAGNITUDE DERIVED QT VARIABILITY.

Major Field: Biomedical Sciences

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy at Center for Health Sciences, Oklahoma State University, Tulsa, OK in July 2021.

Completed the requirements for the Doctor of Osteopathic Medicine at College of Medicine, Oklahoma State University, Tulsa, Oklahoma in May 2020.

Completed the requirements for the Bachelor of Science in Aerospace Engineering at Iowa State University, Ames, Iowa in August 1995.

Experience:

Currently, PGY-2, Family Medicine Residency at the Cherokee Nations Outpatient Clinic and W. W. Hastings Hospital, Tahlequah, Oklahoma.

Worked as Senior Powerplant Engineer at American Airlines from January 2006 to August 2010.

Worked as Crew Chief/Aircraft Maintenance Technician at American Eagle Airlines from March 2004 to December 2004.