

**Association of Adverse Cardiovascular Outcomes
with Weighted Morphologic Variability
following Non-ST-Elevation Acute Coronary Syndromes**

by

Joyatee Mudra Sarker

S.B., Massachusetts Institute of Technology (2009)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

ARCHIVES

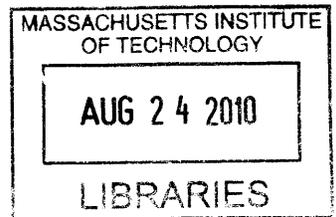
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2010

[June 2010]

© 2010 Massachusetts Institute of Technology. All rights reserved.



Author

Department of Electrical Engineering and Computer Science
May 21, 2010

Certified by

Collin M. Stultz
Associate Professor of Electrical Engineering and Computer Science
and Health Sciences Technology
Thesis Supervisor

Accepted by

Dr. Christopher J. Terman
Chairman, Department Committee on Graduate Theses

**Association of Adverse Cardiovascular Outcomes
with Weighted Morphologic Variability
following Non-ST-Elevation Acute Coronary Syndromes**

by
Joyatee Mudra Sarker

Submitted to the Department of Electrical Engineering and Computer Science

on May 21, 2010

In Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

Patients who have had an acute coronary syndrome (ACS) are at a relatively high risk of having subsequent adverse cardiac events. Several electrocardiographic (ECG) measures such as heart rate variability, heart rate turbulence, deceleration capacity, T-wave alternans, and morphologic variability have been used to identify patients at an increased risk of recurrent myocardial infarctions and cardiovascular death.

In this work, we develop a new ECG-based measure for patient risk stratification called weighted morphologic variability. This measure is based on assessment of beat-to-beat changes in the morphology of consecutive beats. Weighted morphologic variability identifies patients who are at more than four-fold risk for cardiovascular death, which is an improvement in ECG-based risk stratification. The body of this work suggests that prognosticating patients based on electrocardiographic measures is an effective way of identifying those at risk of adverse cardiovascular outcomes.

Thesis Supervisor: Collin M. Stultz
Title: Associate Professor of Electrical Engineering and Computer Science
and Health Sciences Technology

ACKNOWLEDGEMENTS

Above all, I would like to thank Collin for all of the support, advice, and encouragement he has given me through these years. With no irony intended (okay, maybe a little), Collin has the biggest heart of any advisor I have known. He has been such a great listener and his willingness to meet immediately at the drop of a call shows his dedication to his work. I have learned tremendously working with Collin, and I hope to one day be as great of a person as he is.

I would also like to thank Dr. John Guttag and his students for all of their ideas and suggestions. John has a remarkable brain, and it was wonderful to have weekly conversations with him. Thank you to Zeeshan Syed and Phil Sung for laying out the groundwork and answering frantic emails in the beginning. I also want to thank Jenna Wiens and Anima Singh for their input on this body of work.

I would like to thank the soul of the lab, my lab mates. To Dr. Sophie Walker, Dr. Ramon Salsas-Escat, Dr. Paul Nerenberg, Dr. Christine Philips, and the to-be doctors Charles Fischer, Orly Ullman, and Elaine Gee: I thank you for all of the fun and help over the year. A special thanks to Veena Venkataraman and Chelsea Ostrander for their contributions to the project.

Minah Shahbaz, Tia Cartisano, and Carlos Amaya—thank you for putting up with me through my crazies.

Finally, I would like to thank my parents, Bhaba and Lopa, for all of their love, patience, and support. Their own perseverance has inspired me to come this far. To Aneek: thanks for grounding me, and, yes, I will always spoil you.

CONTENTS

ABSTRACT.....	3
ACKNOWLEDGEMENTS.....	5
Chapter I: INTRODUCTION	13
<i>A. Background</i>	15
1. <i>Physiology of the Heart</i>	15
2. <i>Electrocardiograms</i>	15
<i>B. Risk Stratification Measures</i>	18
1. <i>Non ECG-Based Risk-Stratification Measures</i>	18
2. <i>Electrocardiographic Risk-Stratification Measures</i>	19
Chapter II: MORPHOLOGIC VARIABILITY	25
<i>A. Patient Population</i>	25
<i>B. Dynamic Time Warping</i>	26
<i>C. Morphologic Distance</i>	29
<i>D. Morphologic Variability</i>	30
<i>E. Receiver Operating Characteristic</i>	31
Chapter III: WEIGHTED MORPHOLOGIC VARIABILITY.....	33
<i>A. Weighted Morphologic Variability</i>	33
1. <i>WMV with 3 Frequency Bands</i>	36

2. <i>WMV with 600 Frequency Bands</i>	37
<i>B. Simulated Annealing</i>	38
1. <i>Choice of Initial Temperature</i>	39
2. <i>Choice of Temperature Decrement Function</i>	40
3. <i>Number of Perturbations at Each Temperature</i>	41
4. <i>Stop Criterion</i>	42
Chapter IV: EVALUATION OF WEIGHTED MORPHOLOGIC VARIABILITY	45
<i>A. Determining Cutoffs for High and Low Risk</i>	45
1. <i>Cutoffs Using Quartile Values</i>	46
2. <i>Cutoffs Using Hazard Ratios</i>	47
3. <i>Cutoffs Using Sensitivity and Specificity</i>	48
<i>B. Cross-Validation</i>	50
<i>C. Risk Stratifying Cardiovascular Death</i>	52
<i>D. Risk Stratifying Myocardial Infarction</i>	55
Chapter V: SUMMARY AND CONCLUSIONS	59
<i>A. Summary</i>	59
<i>B. Physiological Basis</i>	61
<i>C. Conclusions</i>	62
BIBLIOGRAPHY	65

LIST OF FIGURES

Figure I-1. *Physiology of the heart.* 16

Figure I-2. *Electrophysiology of the heart.* 17

Figure I-3. *RR interval discretization for heart rate turbulence.* 21

Figure I-4. *Heart rate turbulence calculation.* 22

Figure I-5. *RR interval discretization for deceleration capacity.* 23

Figure II-1. *Alignment of heart beats using dynamic time warping.* 27

Figure II-2. *MD time series.* 29

Figure II-3. *Heat map of morphologic variability.* 32

Figure IV-1. *Receiver operating characteristic curve for WMV and cardiovascular death.* 49

Figure IV-2. *Weights for various frequencies in WMV risk stratification of cardiovascular death.*
..... 53

Figure IV-3. *Kaplan Meier hazard curves for MERLIN and DISPERSE populations using WMV
> 1.356.* 54

Figure IV-4. *Weights for various frequencies in WMV risk stratification of myocardial infarction.*
..... 55

Figure IV-5. *Kaplan Meier hazard curves for MERLIN and DISPERSE populations using WMV
> 2.357.* 57

LIST OF TABLES

Table III-1. *C-statistics from weighting various MV values.* 37

Table IV-1. *Hazard ratio, sensitivity, and specificity using quartile cutoff for cardiovascular death.* 47

Table IV-2. *Hazard ratio, sensitivity, and specificity using maximum hazard ratio cutoff for cardiovascular death.* 48

Table IV-3. *Hazard ratio, sensitivity, and specificity using maximum sensitivity and specificity cutoff for cardiovascular death.* 49

Table IV-4. *Hazard ratio, sensitivity, and specificity using maximum sensitivity and specificity cutoff for myocardial infarction.* 49

Table IV-5. *Cross-validation hazard ratios using WMV to risk stratify for cardiovascular death.* 51

Table IV-6. *Cross-validation hazard ratios using WMV to risk stratify for myocardial infarction.* 51

Table IV-7. *Hazard ratio for risk stratification measures of cardiovascular death on MERLIN and DISPERSE patients.* 53

Table IV-8. *Hazard ratio for risk stratification measures of myocardial infarction on MERLIN and DISPERSE patients.* 56

Chapter I:

INTRODUCTION

Approximately 3.5 million people in the United States are hospitalized every year due to an acute coronary syndrome [1]. An acute coronary syndrome (ACS) is a broad term that describes clinical events characterized, in part, by myocardial ischemia, which is characterized by decreased blood flow to the heart. Patients with an ACS often enter the hospital complaining of chest discomfort that may be accompanied with shortness of breath. ACS can be grouped into two categories: unstable angina and myocardial infarction (MI). Unstable angina is due to ischemia without evidence of infarction, or death of the heart muscle (myocardium). A myocardial infarction, more commonly referred to as a heart attack, is the condition when ischemia leads to death of the myocardium.

Patients who are diagnosed with an ACS are more likely to have other adverse cardiovascular events. The GUSTO-IIb trial followed 12,142 patients post-ACS to determine the effects of recurrent ischemia [2]. Patients with a non-ST segment elevation MI (NSTEMI) had a 5.7% 30-day mortality rate and an 11.1% 1-year mortality rate. Patients with an unstable angina had a 2.4% 30-day mortality rate and a 7.0% 1-year mortality rate. 7.5% of NSTEMI patients and 4.8% of unstable angina patients had a myocardial (re)infarction within 30 days. Post-ACS patients are also likely to have congestive heart failure (CHF), a disease in which failure of the heart to adequately pump blood to vital organs leads to increased pressure in the pulmonary vasculature. The increased pulmonary vasculature pressure results in shortness of breath. The accurate identification of high risk patients is of importance because once these patients are

detected, doctors can offer them treatments that may decrease their risk of future adverse outcomes.

Typically, an acute coronary syndrome (ACS) is caused by atherosclerosis, a potentially dangerous disease characterized by the accumulation of plaques in blood vessels. Atherosclerosis is very common—nearly all adults have some plaque amassment in their blood vessels [3]. However, not everyone is aggressively treated for atherosclerosis because they are not at high risk of adverse cardiovascular outcomes. Patients may be at high risk due to hereditary factors or their lifestyle choices. The extent of an ACS may also vary widely. Therefore, doctors are keen on understanding the likelihood and type of adverse cardiovascular event that a post-ACS patient may have.

Currently, there are several techniques that are used to predict adverse cardiovascular outcomes. The Thrombolysis in Myocardial Infarction (TIMI) risk score is computed from various clinical factors observed at the initial admission of a patient. Cardiac catheterization is a helpful invasive test; however, indiscriminate use on all patients may lead to unnecessary complications on patients who are at low-risk. Many non-invasive tests are used to risk stratify patients, such as echocardiography, cardiac magnetic resonance imaging (MRI), and cardiac x-ray computed tomography (CT), but these are expensive and not readily available in many communities. In contrast, electrocardiographic (ECG) -based measures are inexpensive, non-invasive, and are easily obtained. We propose the use of a novel ECG-based metric to predict adverse cardiovascular outcomes in post-ACS patients.

This thesis is organized as follows: the remainder of this chapter is devoted to background information on electrocardiograms and other risk-stratification techniques. Chapter 2 describes a specific electrocardiographic technique in detail, called morphologic variability

(MV). Chapter 3 explains the novel electrocardiographic risk-stratification technique called weighted morphologic variability (WMV) in association with mortality and myocardial infarctions. Chapter 4 evaluates the performance of WMV. Finally, chapter 5 ends with a summary of WMV with a hypothesis about its physiological basis.

A. Background

1. Physiology of the Heart

The heart is a pump that transports blood, carrying essential nutrients and oxygen, throughout the body. The heart consists of four chambers: the left and right upper atria, and the left and right lower ventricles (Figure I-1). The left atrium collects oxygenated blood from the lungs and delivers it to the left ventricle. The left ventricle then pumps this oxygenated blood to the rest of the body. After blood has been dispersed through the body, and oxygen and nutrients have been taken up, the deoxygenated blood flows into the right atrium. The right atrium delivers this blood to the right ventricle, which pumps the blood to the lungs to be oxygenated. A full heart beat consists of a simultaneous contraction of the atria, followed by a simultaneous contraction of the ventricles.

2. Electrocardiograms

Electrocardiograms (ECGs) measure surface electrochemical potential differences that result from signals that originate in the heart (Figure I-2a). Portable Holter monitors can collect ECGs from patients for at least 24 hours. Between 3 to 12 leads are placed in various locations on the body to measure surface potential differences arising from the heart. Many electrocardiographic risk stratification metrics only require information from a single channel

(consisting of three leads) to predict a patient's cardiovascular health. The work in this thesis is developed on the information from a single channel.

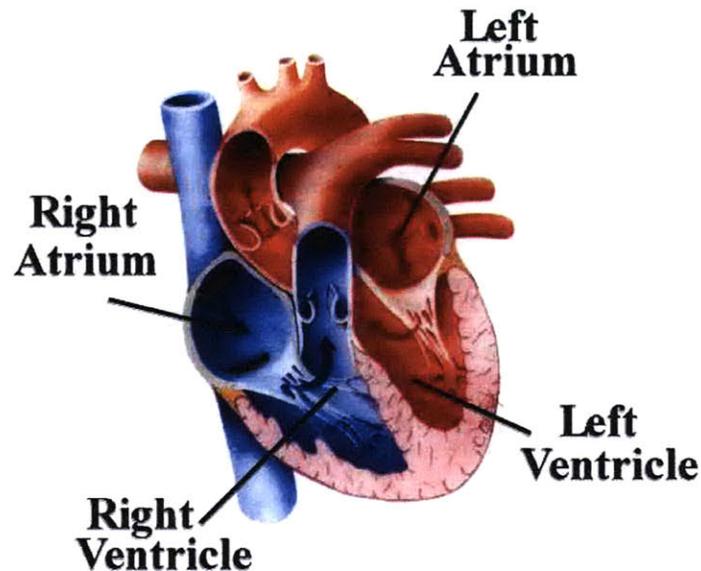


Figure I-1. *Physiology of the heart.* The heart has four chambers: the left and right atria, and left and right ventricles. Image courtesy of Hartzog [4].

At rest, a myocardial cell, or myocyte, remains at a negative potential relative to the outside of the cell. If the myocyte is stimulated, it becomes depolarized as positive ions flow into the cell. The cell then repolarizes and returns to its normal, resting state. In the heart, the electrical stimulation begins at the sinoatrial node (Figure I-2a). This stimulation spreads through the atria, then through the septum that separates the two ventricles, and finally along the outer sides of the ventricles. After the initial electrical stimulation, the myocytes all repolarize in order to return to their resting potential.

An electrocardiogram measures the sum of the potential differences of all of the myocytes of the heart. Several leads are placed in different locations on the thorax. Each pair of

leads records the morphology of the voltage changes along a particular axis in the heart. By analyzing all of the recordings from the leads together, one can determine how the electrical impulse travels through the heart. Physicians typically identify different parts of the conduction path with the names “P wave,” “PR segment,” “QRS complex,” “ST segment,” and “T wave” (Figure I-2b). The P wave represents right and left atrial depolarization, the QRS complex represent ventricular depolarization, and the T wave represents ventricular repolarization.

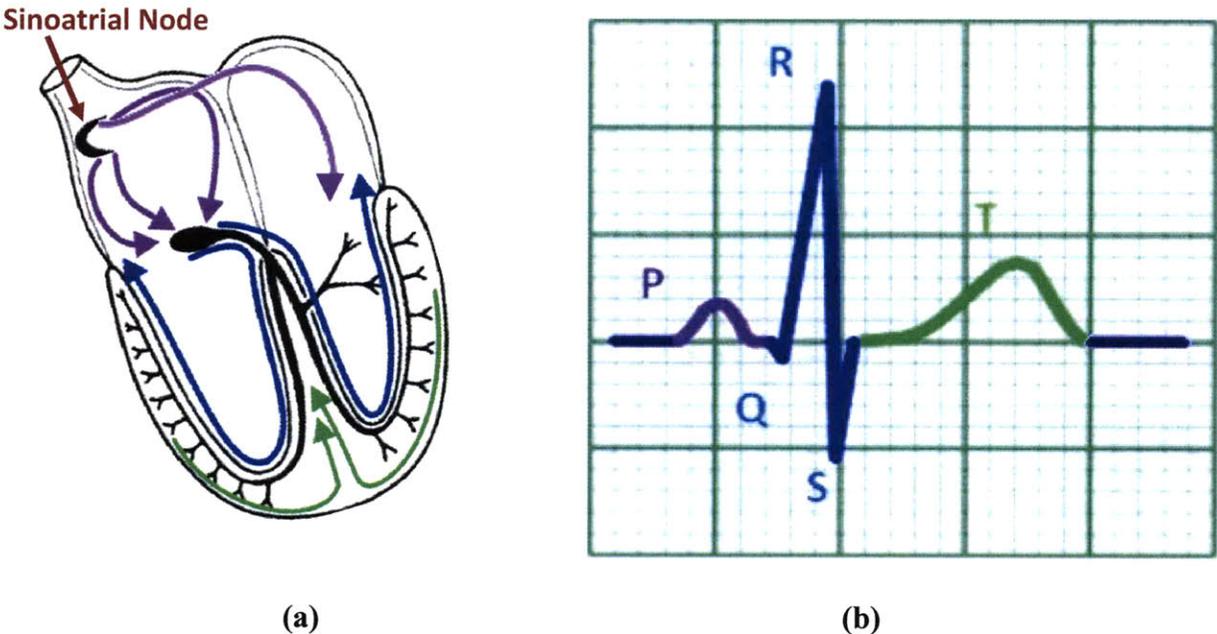


Figure I-2. Electrophysiology of the heart. Figure I-2a depicts the electrical impulses travelling through the heart, and Figure I-2b shows the corresponding ECG signal recorded. Images courtesy of Lemkin [5].

B. Risk Stratification Measures

Various techniques are used to risk stratify patients for adverse cardiovascular outcomes. The following section will describe a few methods in detail. In particular, the TIMI risk score and echocardiograms are used to noninvasively risk-stratify patients. Also, various electrocardiographic techniques can be used to predict if post-ACS patients will have a future adverse cardiovascular outcome.

1. Non ECG-Based Risk-Stratification Measures

A TIMI risk score was developed to assess a patient's risk of death, recurrent MI, or ischemia, thereby allowing doctors to make better therapeutic decisions. The score is tabulated based on seven independent factors, which are all associated with poor cardiac outcomes: age \geq 65 years, the presence of at least three risk factors for cardiac heart disease, prior coronary stenosis of \geq 50%, the presence of ST segment deviation on admission ECG, at least two angina episodes in the prior 24 hours, elevated serum cardiac biomarkers, and the use of aspirin in the prior seven days [6]. While the TIMI risk score is a useful predictor of future cardiac events, some patients who do not have a high risk score may still suffer from cardiovascular events and other patients with a low risk score may have adverse cardiovascular sequelae. Therefore, additional methods for accurate risk stratification are needed.

An echocardiogram is an ultrasound study of the heart, and can reveal information about blockages and blood flow within the heart. Other imaging techniques can provide similar information. Clinicians often use the left ventricular ejection fraction (LVEF) as a measure of the health of the heart, which can be calculated using an echocardiogram. The LVEF is the fraction of blood in the left ventricle that is pumped out of the heart during systole, when the left

ventricle contracts. Typically, a healthy patient has an LVEF of 55-75% [7]. The Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM) Program found that adverse cardiovascular outcomes were minimal in patients with LVEF higher than 45% [8]. Clinicians typically aggressively treat patients with LVEF lower than 40%.

2. Electrocardiographic Risk-Stratification Measures

Damaged myocardium conducts electrochemical impulses differently than normal myocardium. As a result, the morphology of ECG signals may look very different in diseased hearts. The morphology of the ECG can therefore give insight to the health of the cardiovascular system. By observing the shape of a patient's ECG, physicians can gain insight into the problems a patient may have. Heart rate variability (HRV), deceleration capacity (DC), and T-wave alternans (TWA)—quantities that are calculated from the ECG—can provide information on the “electrical health” of the heart, thereby identifying patients at potential risk.

Heart rate variability (HRV) is a measure of small differences in the heart rate. Since the heart rate is largely determined by the autonomic nervous system, HRV is an indirect measure of the health of the cardiovascular and autonomic nervous systems. In some people, either the autonomic nervous system does not function properly or the heart does not respond appropriately to the autonomic stimuli. An unhealthy patient's heart rate exhibits relatively small beat-to-beat variations in measured Holter recordings.

HRV is based on an analysis of the RR time series. The RR series is a sequence of times between successive R-waves in an ECG. The NN series (for normal-to-normal) is defined as the RR time series for “normal beats; i.e., beats representing normal conduction through the myocardium. Time-domain metrics commonly include the SDNN (standard deviation of the NN

series) and the SDANN (standard deviation of the mean NN series in a five-minute window). Frequency-domain metrics use the power spectral density (PSD) of the RR interval time series. One frequency metric includes LF/HF, which is the ratio of the power in the low frequency band (0.04 – 0.15 Hz) to the power in the high frequency band (0.15 – 0.40 Hz). The LF/HF ratio is computed for 5-minute windows. Patients with a LF/HF ratio < 0.95 are considered to be at risk [9]. Frequency-domain metrics correlate well with in-hospital mortality independently of other commonly-used predictors such as left ventricular ejection fraction [10]. Low HRV also correlated with death within 30 days and death within a year after admission to a hospital for an unstable angina or acute MI [11]. Combined with other risk factors, HRV increases the predictability of cardiac events [10].

Heart rate turbulence (HRT) is an electrocardiographic measure that is also based on changes in the heart rate. In particular, HRT assesses the response of the heart rate following premature ventricular complexes (PVCs). The premature beat causes less blood than normal to eject from the heart, resulting in a lower blood pressure than expected. The autonomic nervous system maintains homeostasis by increasing the heart rate in order to increase the blood pressure. Usually, the autonomic nervous system over-regulates and the blood pressure increases too much. The heart rate is then decreased in order to bring the blood pressure back to homeostatic levels.

HRT quantifies the heart rate fluctuations with two ECG metrics: the turbulence onset (TO) and the turbulence slope (TS). TO is the percentage change, expressed as a ratio, in the first RR interval after PVC to the last RR interval before PVC, to quantify the increasing heart rate changes. This is shown in (Equation I-1), where RR 1 and RR 2 are the first two normal intervals

following a PVC, and RR -2 and RR -1 are the first two normal intervals preceding a PVC (Figure I-3, Figure I-4).

$$TO = \frac{(RR\ 1 + RR\ 2) - (RR\ -2 + RR\ -1)}{(RR\ -2 + RR\ -1)} \quad \text{(Equation I-1)}$$

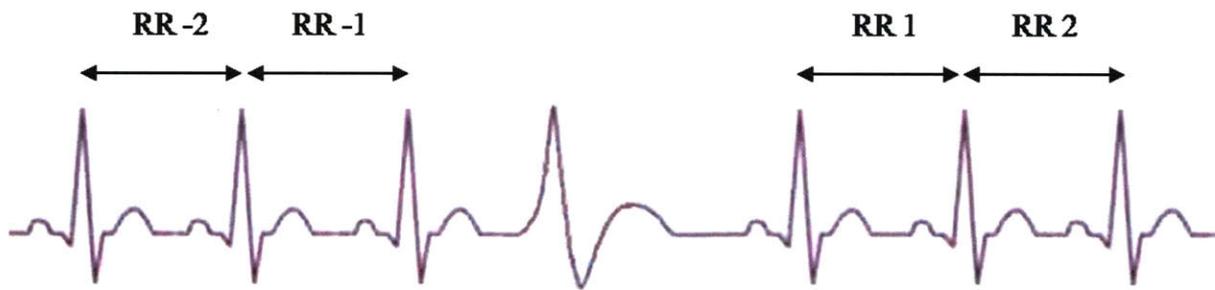


Figure I-3. RR interval discretization for heart rate turbulence. RR -2 and RR -1 are the first two normal intervals preceding a PVC and RR 1 and RR 2 are the first two normal intervals following a PVC. Figure adapted from Biocom Technologies [12].

The turbulence slope (TS) measures the slowing of the heart rate as it returns to its baseline value. TS is the maximum regression slope of any five consecutive RR interval values following a PVC (Figure I-4). A lower slope means the heart takes longer to return to the baseline value, indicating an unhealthy heart or an unhealthy autonomic nervous system.

Patients are scored 0 for normal HRT values, 1 for either an abnormal TS or abnormal TO, and 2 for abnormal TS and TO. Patients who are in the HRT category 2 are more likely to die following a myocardial infarction (MI) [13]. The baroreflex function activates the parasympathetic nervous system, which lowers blood pressure within the body and is found to correlate with HRT [14].

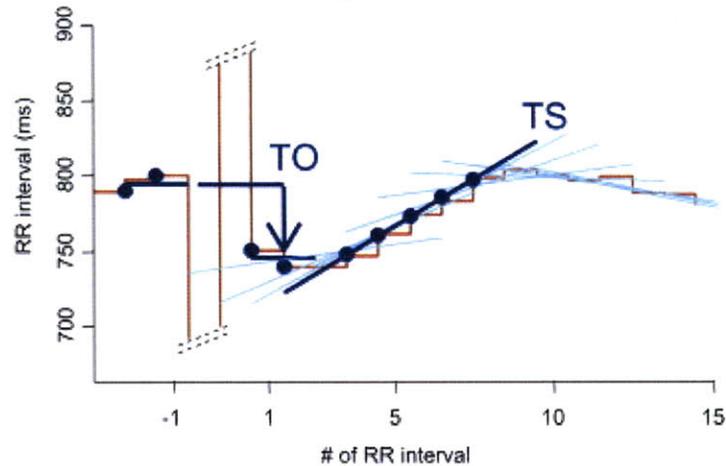


Figure I-4. Heart rate turbulence calculation. Heart rate turbulence (HRT) is calculated by the turbulence onset (TO) metric and the turbulence slope (TS) metric. The PVC is identified as beat ‘0’. TO is the ratio between the average of the two RR intervals preceding a PVC to the average of the two RR intervals following a PVC. TS is the maximum regression slope of 5 consecutive RR intervals within the first 15 RR intervals after the PVC. The light blue lines indicate all of the possible regression slopes, and the dark blue line indicates the maximum regression slope. Figure courtesy of Bauer [15].

Deceleration capacity (DC) has also been associated with undesired cardiac events [16]. This work has been developed as an extension to the work on HRT, but no PVC is required for the calculations. First, anchors are identified in which one long RR interval follows a shorter RR interval. The anchor is defined as the longer RR interval. The computation for DC is similar to that of TO calculated in HRT. The lengths of the two RR intervals before (RR -2 and RR -1) the anchor, the length of the anchor (RR 0), and the length of the interval after the anchor (RR 1) are recorded (Figure I-5). The average of all of the lengths of the RR -2 intervals, RR -1 intervals,

RR 0 intervals, and RR 1 intervals are denoted by $X[-2]$, $X[-1]$, $X[0]$, and $X[1]$, respectively.

Then, the DC calculation is as follows:

$$DC = \frac{(X[0] + X[1]) - (X[-2] + X[-1])}{4} \quad \text{(Equation I-2)}$$

Low DC values were found to be more predictive of mortality than LVEF in post-myocardial infarction patients [16].

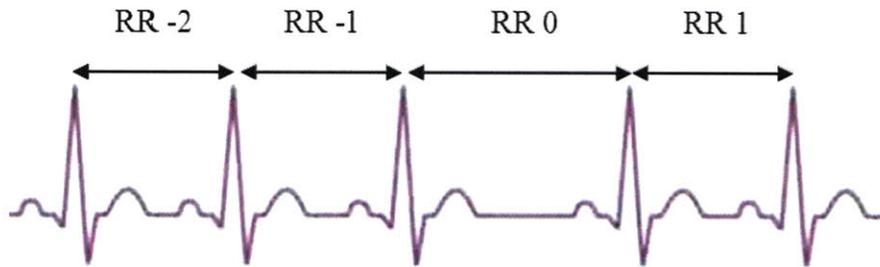


Figure I-5. RR interval discretization for deceleration capacity. The anchor is the identified as the long RR intervals where one long RR interval precedes a shorter one. RR -2 and RR -1 are the first two RR intervals preceding the anchor, RR 0 is the long RR interval anchor, and RR 1 is the first normal interval following the anchor. Figure adapted from Biocom Technologies [12].

Finally, T-wave alternans (TWA) is another ECG-based analytical technique used to risk stratify patients. In some patients, the T-wave varies in magnitude from beat-to-beat, usually alternating between two morphologies. TWA measures these variations. TWA has been associated with ventricular arrhythmias [17] and may be correlated with sudden cardiac death (SCD) [14]. However, TWA is often difficult to measure from standard ECGs because it operates in the range of microvolts while standard ECG recordings are on a millivolt scale. Therefore, sensitive equipment is required to measure TWA.

HRV, HRT, DC, and TWA all correlate with poor outcomes such as death and ventricular arrhythmias. However, none of these measures are able to identify all patients who are at risk of adverse cardiac events. Therefore, improved risk-stratification measures are still required.

Chapter II: MORPHOLOGIC VARIABILITY

Morphologic variability (MV) measures the underlying beat-to-beat changes in the shape of an ECG. To calculate the MV of a patient, the ECG first needs to be segmented into the individual beats. Then, an intermediate time series called the morphologic distance (MD) is calculated using a technique called dynamic time warping (DTW) that measures the difference between two successive beats. Finally, the MD is analyzed in the frequency domain to determine the MV value for a patient.

A. Patient Population

The goal of this work is to identify patients who are high risk of cardiovascular death and recurrent myocardial infarction after an ACS. To accomplish this, we developed and tested the WMV method using data from the DISPERSE-2 (TIMI 33) trial and MERLIN (TIMI 36) trial.

The MERLIN trial followed 6560 post-non-ST segment elevation ACS (NSTEMI) patients for a median follow up period of 348 days, and the DISPERSE-2 (TIMI33) trial followed 990 post-NSTEMI ACS patients for 90 days for clinical drug studies [18, 19]. ECG data were sampled at 128 Hz for a portion of those patients, and data with too much noise were discarded.

The ECGs from these patients were pre-processed for calculating MV [20]. Patients with less than 24 hours of continuous ECG data were excluded from the study. Noise was removed from the signals in two steps. First, the ECG signal was median filtered to estimate the amount of

baseline wander. This estimated baseline wander was subtracted from the ECG signal. Second, additional noise was removed by applying a wavelet denoising filter with a soft threshold [21]. Afterwards, parts of the ECG signal were rejected if the signal-to-noise ratio was still small post-noise-removal. The Physionet Signal Quality Index (SQI) package removed portions of the ECG with a high signal-to-noise ratio and ectopic beats [22]. This was to ensure that the actual morphology of the beats was being analyzed instead of the length of the beats. Finally, the ECG signal was divided into 30-minute segments. For each segment, if the standard deviation of the R-wave amplitude was greater than the standard deviation for a uniform distribution between 0.5 and 1.5, the segment was removed [20]. The reasoning behind this is that, physiologically, it is unlikely for the R-wave to change uniformly by more than 50% of its mean value. Therefore, a maximum standard deviation of 0.2887 was chosen for the R-wave amplitudes in each 30-minute interval.

Data from a total of 764 DISPERSE patients (14 cardiovascular deaths, 28 post-ACS myocardial infarctions) and 4555 MERLIN patients (230 cardiovascular deaths, 347 myocardial infarctions) were observed. 2254 of the MERLIN patients were given a trial drug called ranolazine. A placebo was administered on the rest of the MERLIN patients. The patients were given Holter monitors to continuously record their ECGs for 7 days after a hospitalization occurred within 48 hours of an ACS.

B. Dynamic Time Warping

To calculate morphologic variability, the energy difference between each pair of consecutive beats first needs to be found. If the beats were strictly subtracted from each other however, samples from one beat may be compared with samples from another beat that are part

of different waves or intervals. For example, in Figure II-1, the end of the T-wave in the blue ECG corresponds to the middle of the T-wave in the red ECG. To determine the difference between two beats based on morphology, we wanted to ensure that appropriate segments of each beat were being compared. Dynamic time warping (DTW) measures the morphologic similarity between two heart beats (Figure II-1). The two beats are “warped” in the time domain to reveal the underlying morphological differences in the beats. One sample in a beat can match with several samples in another beat.

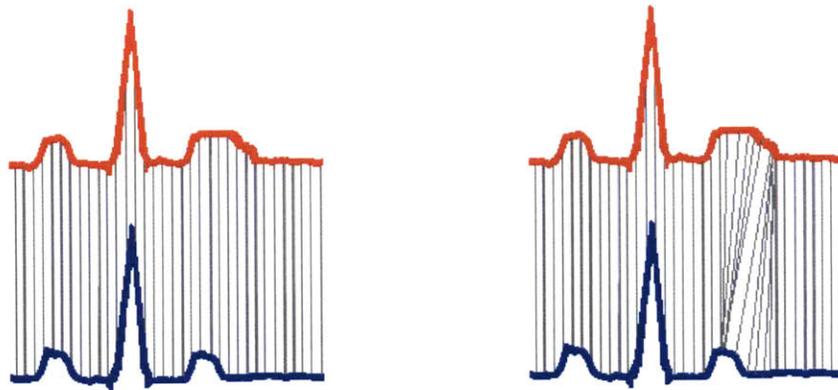


Figure II-1. Alignment of heart beats using dynamic time warping. The two beats on the left depict a comparison of the beats using strictly the Euclidean distance. The vertical lines show that the sample from a given time in the red beat is compared to a sample at the same time from the blue beat, though the beats may be at different phases on conduction at that time. The end of the T-wave in the blue beat is being compared to the middle of the T-wave in the red beat. The figure on the right shows how DTW can align two beats so that relevant conduction phases are compared. Relevant points in the conduction path of the two beats are evaluated. This figure depicts the variability in the T-wave. However, DTW can capture differences along the entire conduction cycle. Image courtesy of Syed, et al. [20].

DTW is an alignment of two time series, A and B , of length m and n , respectively. An $m \times n$ distance matrix is constructed, where the (i, j) element of the matrix is the distance between the i^{th} element of A and the j^{th} element of B . The distance metric used is the Euclidean distance, $(A(i) - B(j))^2$. A warping path, φ , of length K is found from this distance matrix. An alignment of the two beats is simply the warping path. $\varphi(k)$ is defined as:

$$\varphi(k) = (\varphi_A(k), \varphi_B(k)), 1 \leq k \leq K, \quad (\text{Equation II-1})$$

where φ_A represents the row index of the distance matrix and φ_B represents the column index of the distance matrix.

The *dynamic time warping alignment* is the path that has the minimum cost associated with it. The cost, C , for a particular path is defined as the sum of the squares of the differences between the pairs of matched elements (Equation II-2).

$$C_\varphi(A, B) = \sum_{k=1}^K (A[\varphi_A(k)] - B[\varphi_B(k)])^2 \quad (\text{Equation II-2})$$

The cost to align two successive beats is the energy difference between the two beats. Therefore, the dynamic time warping alignment between the two beats can be used as a metric for how much the two beats differ (Equation II-3).

$$DTW(A, B) = \min_{\varphi} C_\varphi(A, B) \quad (\text{Equation II-3})$$

Thus, the final alignment of the two beats depends both on the amplitude alignment of the beats, and temporal aspect of the beats, as reflected in the length of the alignment, K .

C. Morphologic Distance

The morphologic distance (MD) is a measure of the difference in the morphology of two heart beats. DTW aligns two heart beats by their morphology. The MD value for two beats is defined as the energy difference, or cost, for the DTW alignment of those two beats. An ECG signal is converted to an MD time series by comparing each beat with its previous beat in a pairwise manner (Figure II-2). This MD series is then smoothed using a median filter of length 8 for further analysis [20].

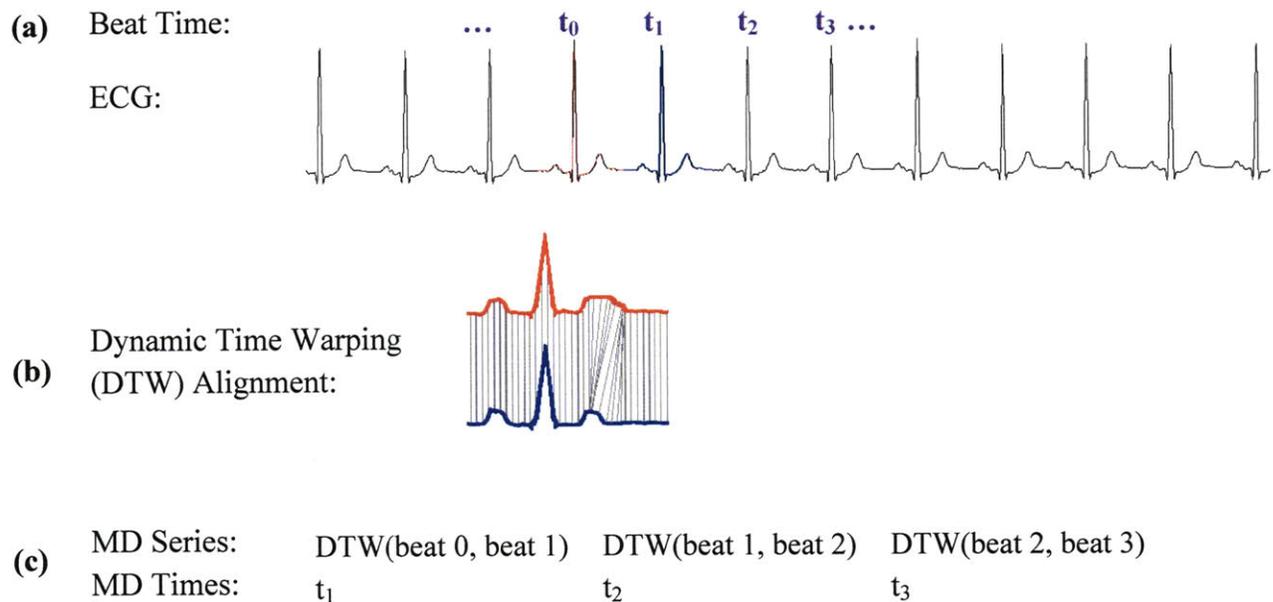


Figure II-2. MD time series. The successive beats of an ECG (a) are aligned using dynamic time warping (DTW) (b). Vertical lines connect data that correspond to each other in the red and blue heart beats. After using DTW, related portions of a beat are mapped to correspond to each other. The morphologic distance (MD) is the sum-of-squares energy difference between aligned beats, or the DTW cost (c). DTW image courtesy of Syed, et al. [20].

D. Morphologic Variability

Commonly used electrocardiographic risk-stratification measures observe variations in the heart rate of a patient. Morphologic variability (MV) observes changes in the shapes of the heart beats in an ECG [20]. These changes in the morphology of the heart beat may indicate some underlying problem with the conduction path of the electrical signal through the heart. The problem in the conduction pathway could be due to ischemia or myocardial infarction.

MV is computed from the morphologic distance time series. First, the ECG signal for a patient is divided into 5-minute intervals. Only the first 24 hours of the signal are used. Then, the DTW distance between each pair of consecutive beats is computed and recorded as the MD time series. Using the rationale from HRV, the frequency domain of the MD time series can also characterize the variability in the ECG morphology. Therefore, the power spectral density of the MD series is computed. To compute the MV, the sum of the power over a particular diagnostic frequency band is found for all five-minute intervals for each patient (Equation II-4).

$$MV_{\theta} = \sum_{v=LF}^{HF} Power_{\theta}(v) \quad (\text{Equation II-4})$$

where θ represents a particular 5 minute interval, v is the frequency, LF is the low frequency bound, and HF is the high frequency bound. The final MV value is the 90th percentile value of all of the sums, or 90th percentile value of the MV_{θ} 's.

The choice of the optimal frequency band in the computation of MV is a key component of the analysis. To find an optimal diagnostic frequency band, the subsequent procedure is followed. For all combinations of low and high threshold frequencies between 0.10 Hz and 0.60 Hz in intervals of 0.01 Hz, MV values are calculated for each patient. The correlation of the MV values and certain cardiovascular outcomes are then measured based on a receiver operating

characteristic (ROC) curve. The optimal frequency band is found from this ROC analysis, as described in the following section.

E. Receiver Operating Characteristic

The receiver operating characteristic (ROC) is a measure of the predictive capability of a metric. Specifically, the ROC measures the true positive, hit rate of correctly identifying patients at high risk of adverse cardiovascular outcomes versus the false positive rate of incorrectly identifying patients at high risk. The area under the ROC curve is the c-statistic. A c-statistic of 0.5 indicates that the true positive rate is equivalent to the false positive rate, meaning the predictor cannot distinguish between high risk and low risk patients. The larger the area under the ROC curve, the larger the true positive rate. Generally, a c-statistic of 0.7 or greater is considered to adequately distinguish between two outcomes.

By computing ROC curves for many different frequency bands, we identify the frequency range that is best able to differentiate between low risk and high risk patients. Based on the MV values, a frequency band that correlates the most with a certain outcome is identified. For each combination of frequency bands from 0.10 Hz to 0.60 Hz, with a resolution of 0.01 Hz, a c-statistic is computed. These c-statistics are then combined graphically to form a ‘heat map.’ A heat map shows which frequency ranges provide the highest c-statistic, or the best predictive value (Figure II-3).

Previous work suggests that patients in the upper quartile of MV values (> 52.5) had a c-statistic of 0.72 for a frequency range between 0.30-0.55 Hz when calculated to predict death. This diagnostic frequency of MV values was found to have a hazard ratio for death of 8.46 post-ACS, indicating that MV was predictive of death [20].

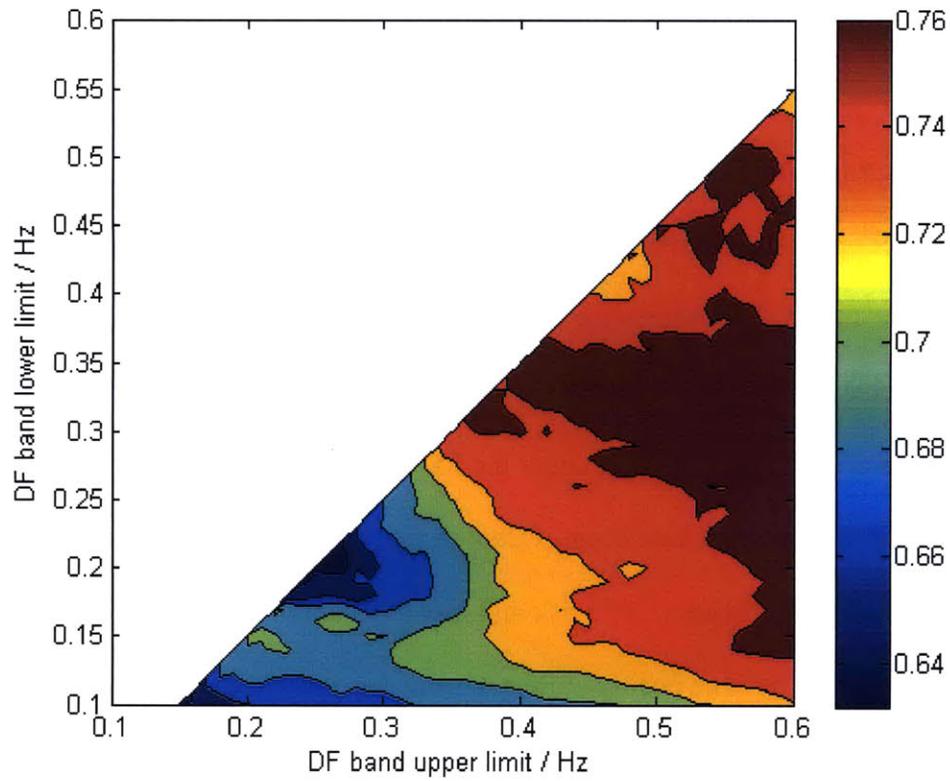


Figure II-3. Heat map of morphologic variability. The best c-statistic for MV was 0.774 in the 0.35-0.36 Hz frequency range. However, the diagnostic frequency was generalized to 0.30-0.55 Hz.

Chapter III:

WEIGHTED MORPHOLOGIC VARIABILITY

In this chapter, a new method, called weighted morphologic variability (WMV), will be introduced that uses electrocardiographic morphologic information to predict post-ACS adverse cardiovascular outcomes. This chapter is organized as follows. First, we will describe the patient populations we used to evaluate WMV. Next, the calculation of WMV will be presented. WMV weights various frequencies according to their importance in risk stratification. We end by describing the simulated annealing algorithm that was used to find the best set of weights for risk stratifying patients.

A. Weighted Morphologic Variability

Weighted morphologic variability (WMV) is an extension of the work on morphologic variability. Morphologic variability uses one frequency band to estimate a given patient's risk of subsequent adverse cardiac event. WMV uses multiple frequency bands to estimate risk factors.

Recall that the MV was found by initially calculating the power spectral density of the MD series, divided into 5-minute intervals. The MV for a particular 5 minute interval was found as the sum of the power over a particular frequency band (Equation III-1). A patient's MV was then the 90th percentile value of the entire individual 5-minute interval MVs. The 90th percentile value was chosen maximize information, but minimize noise [23]. Thus, the steps for finding MV are:

1. Compute the MD time series and divide the MD time series into 5 minute intervals.

2. Compute the MV_{θ} for each 5-minute interval over a particular frequency band (Equation III-1).

$$MV_{\theta} = \sum_{v=LF}^{HF} Power_{\theta}(v)$$

(Equation III-1)

3. MV is the 90th percentile MV_{θ} value.

Weighted morphologic variability similarly uses the MD time series, divided into 5-minute intervals. However, WMV associates a weight with each frequency interval. WMV is computed as following:

1. The MD series is calculated and divided into 5 minute intervals.
2. The maximum frequency band is divided into n intervals.
3. For each interval, WMV_i is the 90th percentile value of the sum of the power in that particular frequency band (Equation III-2).

$$WMV_i = \sum_{v=LF_i}^{HF_i} Power_{\theta}(v)$$

(Equation III-2)

4. WMV is the sum of the weighted WMV_{θ} 's (Equation III-3).

$$\begin{aligned} WMV &= \sum_{i=1}^n \alpha_i \cdot WMV_i \\ &= \sum_{i=1}^n \alpha_i \sum_{v=LF_i}^{HF_i} Power_{\theta}(v) \end{aligned}$$

(Equation III-3)

where n is the number of intervals into which the complete frequency range is divided into, α_i is the weight associated with a particular frequency band, LF_n is the lower bound frequency of a particular frequency band, and HF_n is the upper bound frequency of the frequency band. $Power_{\theta}$

is the sum of the power over the specified frequency range for a particular 5-minute interval. To give the weights a bound, the sum of α_i 's are set to be 1 (Equation III-4).

$$\sum_{i=1}^n \alpha_i = 1$$

(Equation III-4)

MV is a special case of WMV. If $n = 1$ and $\alpha_i = 1$, (or if $\alpha_i = 1$, and all other $\alpha_x = 0$) then:

$$\begin{aligned} WMV &= \sum_{i=1}^1 1 \sum_{v=LF_i}^{HF_i} Power_{\theta}(v) \\ &= \sum_{v=LF}^{HF} Power_{\theta}(v) = MV \end{aligned}$$

(Equation III-5)

However, it is important to note that extracting the 90% value when computing MV and WMV occur at different points in the calculation. When calculating MV, the 90th percentile value of the sums is found last. In WMV calculations, the 90th percentile values are found first, before weighting the WMV_i's.

Two tests were performed on the various datasets to determine the best set of weights, or α_i 's, and the robustness of WMV in predicting death and myocardial infarction. First, the maximum frequency range was divided into three intervals, where $n = 3$. Then several sets of weights were tested to determine the importance of WMV. Afterwards, simulated annealing was used to find the best set of alphas with $n = 600$, the largest possible n , given the data. The following two subsections detail these two processes.

1. *WMV with 3 Frequency Bands*

WMV was first tested using three frequency bands. The power spectral density of the MD time series was computed from 0.00 Hz to 0.60 Hz. First, the 0.001Hz to 0.600 Hz frequency spectrum was divided into three segments. The spectrum had two variable cutoff points, β_1 and β_2 . The β 's were iterated through all possible values from 0.00 to 0.60 Hz in 0.01 Hz intervals. A weighting, α_i , was assigned to each band. The data were iterated through all of the possible weights from 0 to 1 in intervals of 0.1 for the three segments, and through all of the possible cutoffs for the division into the three segments. The WMV over these ranges were calculated, and weighted according to the possible α 's (Equation III-6).

$$WMV = \alpha_1 \sum_{v=0 \text{ Hz}}^{\beta_1} Power_{\theta}(v) + \alpha_2 \sum_{v=\beta_1}^{\beta_2} Power_{\theta}(v) + \alpha_3 \sum_{v=\beta_2}^{0.6 \text{ Hz}} Power_{\theta}(v) \quad (\text{Equation III-6})$$

Again, the sum of the α_i 's was set to 1. The sum of these weighted MVs was then used to find a c-statistic. The combination of the best cutoffs and best weights was subsequently analyzed. The best c-statistic would have to perform at least as well as the value determined from MV because if β_1 and β_2 were set to 0.30 Hz and 0.55 Hz, respectively, and only $\alpha_2 = 1$, the resulting c-statistic would be exactly that computed from MV in previous studies. The results from this experiment are shown in Table III-1. Indeed, the best c-statistic using WMV was 0.785 for death, compared to that of 0.771 using MV. For myocardial infarctions, the best c-statistic using WMV was 0.610, an improvement from a c-statistic of 0.488 using MV.

The results from this experiment indicated that data from multiple frequencies contain predictive value. Thus, we performed an annealing algorithm to determine which set of weights for the various frequencies would best predict various cardiac outcomes.

DISPERSE-2	α_1	α_2	α_3	β_1	β_2	c-statistic
Death (c-statistic for 0.3-0.55 Hz = 0.771)	0.0	0.9	0.1	0.350 Hz	0.360 Hz	0.785
MI (c-statistic for 0.3-0.55 Hz = 0.488)	1	0	0	0.100 Hz	0.110 Hz	0.610

Table III-1. C-statistics from weighting various MV values. MV values derived from three frequency bands (0.1Hz – β_1 Hz, β_1 Hz – β_2 Hz, β_2 Hz – 0.6 Hz) were weighted according to various α 's. The combination of cutoffs (β 's) and weights (α 's) that produced the best c-statistics are shown. The best c-statistic for predicting cardiovascular death, given MV in the 0.3-0.55 Hz frequency range was 0.771. The best c-statistic for predicting myocardial infarction (MI) previously was 0.488.

2. WMV with 600 Frequency Bands

We further investigated optimal weighting by finding the best possible combination of weights for the smallest frequency bands allowed. The power spectral density of the MD time series was created from 0.001 to 0.600 Hz, in 0.001 Hz intervals. Therefore, 600 weights for each 'band' (0-0.100 Hz, 0.001-0.002 Hz, etc.) needed to be determined (Equation III-7). The morphologic variability for a single frequency (MVSF) is the 90th percentile power from the MD time series for a particular frequency (Equation III-8).

$$WMV = \sum_{i=1}^{600} \alpha_i \cdot MVSF_i \quad (\text{Equation III-7})$$

$$MVSF_i = Power_{\theta}(v), \quad \text{where } v = i \cdot 10^{-3} \quad (\text{Equation III-8})$$

After determining the MVSF's, a set of weights for the various frequencies that best predict various cardiac outcomes needed to be found. Finding the optimal alphas in a high dimensional space is an NP-hard optimization problem. For such problems, one can use a commonly-used method to find the solutions. We used simulated annealing to determine the set of weights that best predict adverse cardiovascular outcomes. A simulated annealing run randomly samples various sets of weights and finds the lowest energy point in an energy landscape. For our purpose, the energy landscape was the set of possible negative of the c-statistics found from calculating the WMV of all possible sets of weights. Therefore, the minimum in the negative c-statistic landscape actually found the set of weights that produced the best c-statistic.

The simulated annealing algorithm, as described in the following section, is performed to find a set of weights that produce the best c-statistic for a particular set of training data on a particular outcome.

B. Simulated Annealing

Simulated annealing is a popular method of optimizing complex functions [24]. The formulation for this method is analogous to crystal optimization. One way to perfect crystallization is to heat the crystals and allow them to slowly cool. Similarly, the simulated annealing algorithm searches through space, randomly sampling, in order to find the global minimum. A control parameter, 'temperature', is used to determine the annealing schedule. This term does not have any physical meaning, and should not be confused with our standard notion of temperature. The temperature begins at a very high level. At each temperature, several points in global space are randomly sampled, and their energies are compared. Some samples are

accepted, based on a Metropolis criterion. As the temperature decreases, samples are accepted only if their energies are lower. The temperature is decreased until some stopping criterion is reached. For our purpose, the energy space for the simulated annealing algorithm was the set of c-statistics produced from computing the WMV for all possible sets of weights. In actuality, because we want to optimize the maximum c-statistic, and simulated annealing finds the minimum of a global space, we use the negative of the c-statistics to determine the global energy space.

An annealing schedule is fully described by the following steps:

1. Choice of the initial temperature
2. Choice of the temperature decrement function
3. Number of perturbations at each temperature
4. Stopping criterion

The remainder of this chapter details the annealing schedule selected to determine the best set of weights for WMV.

1. Choice of Initial Temperature

The initial temperature is set high, such that most perturbations of the weights will be accepted. This ensures that a large energy space is sampled. For this simulation, a temperature was selected that had $> 97\%$ acceptance. This ensured that c-statistics had a large variation in the first step (between 0.3 and 0.8). As temperature decreases, the percentage of acceptances decreases.

2. Choice of Temperature Decrement Function

The temperature is decremented slowly to ensure that energy space is sampled adequately. However, the annealing run needs to complete in finite time. Therefore, the temperature decrementing function needs to be set to optimize the annealing schedule.

The decrementing function used in this simulation is a modified, slower version of the function described in Stultz, et al. [25]. For the annealing step, T_{n+1} , the temperature is decremented by the minimum of ΔT_n and $T_n/8$ (Equation III-9).

$$T_{n+1} = T_n - \min(\Delta T_n, T_n/8), \quad (\text{Equation III-9})$$

where

$$\Delta T_n = \begin{cases} \frac{\mathbf{v}T_n^2}{\tau_n \cdot \sigma(E_0)} & \text{if } \text{lstsq}(T_n) < \textit{tolerance} \\ \frac{T_n^2}{4 \cdot \sigma(E_0)} & \text{if } \text{lstsq}(T_n) > \textit{tolerance} \end{cases}$$

The relaxation time, τ_n , is the time required for the energy perturbations to reach equilibrium at a given temperature. \mathbf{v} is the ‘thermodynamic distance,’ or how close the set of perturbations are to equilibrium. E_0 is the mean energy, or the mean of the negative of the c-statistics, for a set of perturbations at a given temperature. $\sigma(E_0)$ is the standard deviation of the fluctuations in the mean energy over the perturbations.

This annealing schedule is desirable because variable temperature decrements can be made. When the temperature is low, relaxation is very long and small temperature decrements are taken to ensure the global minimum is reached. However, once the system has reached equilibrium, larger temperature decrements may be taken in order to complete the simulation.

3. *Number of Perturbations at Each Temperature*

In order to anneal to the best set of weights, a random set of weights that varied between 0 and 1, with the sum equal to 1, was created. First, all weights were set to 0. Then, a random alpha, α_k , was selected to be set between 0 and 1. Another random zeroed alpha was then set to between 0 and $(1 - \alpha_k)$. This was repeated so that each subsequent alpha was set to some random value between 0 and $(1 - \sum_{i=1}^{600} \alpha_i)$. Thus, the initial set of alphas varied from 0 to 1, summing to 1.

A perturbation was made to each set of alphas. A random alpha, α_k , was chosen to be perturbed. This α_k was set to a random number between a lower bound and an upper bound. The bounds were set as following:

$$\text{upper bound} = \text{minimum}(\alpha_k + \text{rand}, 1)$$

$$\text{lower bound} = \text{maximum}(\alpha_k - \text{rand}, 0)$$

The new alpha, α_k' , then determined the values of the remaining alphas. All of the alphas were renormalized, such that the sum of the weights equals to 1, as shown in (Equation III-10).

$$\alpha'_i = \alpha_i \cdot \frac{1 - \alpha'_k}{1 - \alpha_k}, \quad \text{where } \alpha'_i = \text{new alpha} \tag{Equation III-10}$$

After each perturbing a single alpha value, and renormalizing the rest of the weights on the frequencies, the Metropolis criterion is tested. Each weight, α_i , is used to scale the MVSF value at each frequency for each patient. The total WMV value is the sum of the weighted MVSF values. A c-statistic is calculated from the sensitivity and specificity of the WMV values of each patient. The Metropolis criterion uses the negative of the c-statistic to decide whether to accept or reject the set of weights. The Metropolis criterion accepts the weights that meet one of two requirements:

1. The current c-statistic is larger than the previous perturbation's c-statistic. Equivalently, the negative of the current c-statistic is less than the negative of the previous perturbation's c-statistic.
2. A proposal Gaussian density, centered at the negative c-statistic of the current set of weights, is greater than some random number in (0, 1). Formally, this is written as

$$e^{-\frac{E-E'}{T}} > \rho$$

where E is the old c-statistic, E' is the new c-statistic, T is the temperature, and ρ is a random value between 0 and 1.

All of the accepted c-statistics for the current temperature are then fit to an exponential curve. The time constant, τ , of the curve is determined, and is defined as the relaxation time. The temperature is decremented either if more than 3τ perturbations have been made or if the maximum number of perturbations, set by the user, has been reached. For our purposes, the maximum number of perturbations was 100x the degrees of freedom, or 60,000 maximum perturbations.

4. Stop Criterion

The entire annealing is considered complete when the c-statistic value is not changing significantly. At high temperatures, perturbations occur far way from each other on an energy landscape. However, at lower temperatures, accepted α perturbations have c-statistics near each other. At very low temperatures, the c-statistics of the α perturbations should occur right next to each other on an energy landscape. Therefore, when the c-statistic does not vary much, the annealing run is thought to have found the global energy minimum. To determine if the annealing run has reached this minimum, the standard deviation of the c-statistics calculated at

the end of each temperature. If the standard deviation of the last 10% of the c-statistics is smaller than some ϵ , the run is considered complete. ϵ was set to 5×10^{-5} .

The final product of the simulated annealing run is a set of weights that produce the highest c-statistic, or the lowest negative of the c-statistic, for a particular adverse cardiovascular outcome. The next chapter will evaluate the performance of the weights found for WMV using simulated annealing. The results will be assessed for both cardiovascular death and myocardial infarctions post-ACS.

Chapter IV:

EVALUATION OF WEIGHTED MORPHOLOGIC VARIABILITY

We assessed the performance of weighted morphologic variability (WMV) using patients from both the MERLIN and DISPERSE-2 data sets.

To ensure that small variations in the sizes of the weights would not drastically affect the performance of WMV, we performed a smoothing of the weights from the simulated annealing run. A running average of various window sizes was performed on the alphas. A window size of 0.007 Hz was chosen to compute a running average on the weights, when computing WMV. These smoothed weights were used for the remainder of this work.

This chapter outlines how to separate individuals at low risk of adverse cardiovascular outcomes from high risk individuals, how we validated our data, and the performance of WMV in prognosticating.

A. Determining Cutoffs for High and Low Risk

WMVs for all of the patients were calculated using the method described in chapter 3. Our aim was to determine a low risk and a high risk group of patients for complete risk stratification. The hazard ratio, sensitivity, and specificity of the test could thus be calculated. The hazard ratio is a measure of risk. We used Cox proportional hazards model to estimate the hazard ratio, which compares the rates of death for the high risk and low risk patients [26]. Sensitivity is the true positive rate of the test, or the ratio of true positives to the total number of

positives in a test. Specificity is the true negative rate. A high sensitivity and specificity suggests that a test is accurate at predicting risk. In order to determine the risk, the WMV values for all of the patients are sorted, and a cutoff was determined that separated a low risk group from a high risk group. This cutoff was determined in one of three ways: using the quartile as a cutoff, maximizing the hazard ratio, and maximizing sensitivity and specificity values. These three methods are described in detail below.

1. Cutoffs Using Quartile Values

The first method of determining a risk cutoff was using the quartile as a marker. In Syed, et al. [20], the upper quartile of MV values was considered high risk, and the lower three quartiles of MV values were considered low risk. Similarly, the upper quartile of WMV values was determined to be high risk. The hazard ratio, sensitivity, and specificity for this cutoff were calculated using the training and testing sets, as shown in Table IV-1. Using quartiles as the difference between low and high risk groups was typically used for research in MV; however, a more data-driven approach to determining a cutoff was desired.

	Dataset	Hazard Ratios	Sensitivity	Specificity
<i>Best Fit Weights</i> <i>(Cutoff =1.3794)</i>	MERLIN, placebo	6.61	0.559	0.764
	DISPERSE	6.37	0.714	0.717
	MERLIN, ranolazine	3.68	0.429	0.806
<i>Running Average</i> <i>(Cutoff =1.3778)</i>	MERLIN, placebo	6.03	0.539	0.764
	DISPERSE	6.21	0.714	0.711
	MERLIN, ranolazine	3.70	0.429	0.791

Table IV-1. Hazard ratio, sensitivity, and specificity using quartile cutoff for cardiovascular death. The DISPERSE testing dataset contained 764 patients (14 cardiovascular deaths), the MERLIN placebo training group included 2301 patients (102 cardiovascular deaths), and the MERLIN ranolazine testing group had 2254 patients (91 cardiovascular deaths).

2. Cutoffs Using Hazard Ratios

The second method of determining a risk cutoff was using hazard ratios. The WMV values were sorted from lowest to highest value. Then, the hazard ratios for all of the possible cutoffs between two WMV values were calculated. The cutoff that maximized the hazard ratio was used to determine the difference between high and low risk. These results are shown in Table IV-2. However, these results were not used for a final cutoff because the results over-fit for a high hazard ratio. The high hazard ratio used for one set of data performed poorly on other sets of data. Therefore, this method of finding a cutoff was not used for the final results.

	Dataset	Hazard Ratios	Sensitivity	Specificity
Best Fit Alphas (Cutoff = 1.1816)	MERLIN, placebo	1366302.55	0.990	0.054
	DISPERSE	1347131.55	1.000	0.039
	MERLIN, ranolazine	1.09	0.967	0.048
Running Average (Cutoff = 1.2448)	MERLIN, placebo	7.98	0.912	0.254
	DISPERSE	3.14	0.929	0.196
	MERLIN, ranolazine	3.47	0.846	0.250

Table IV-2. Hazard ratio, sensitivity, and specificity using maximum hazard ratio cutoff for cardiovascular death. This technique of determining cutoffs tends to produce high hazard ratios for the testing data, seemingly over-fitting the data. The sensitivity (true positive rate) is nearly 100%. However, the specificity (true negative rate) is very poor. This suggests that most of the patients are classified as high risk, and few are classified as low risk. This method of determining a cutoff does not discriminate between high and low risk well, and, therefore, a better method is required.

3. Cutoffs Using Sensitivity and Specificity

Sensitivity and specificity are measures of the predictive ability of a particular test. Sensitivity is the true positive, hit rate of correctly identifying patients at high risk. Specificity is the true negative rate, the rate at correctly identifying low-risk patients. Our final metric of determining the cutoff was finding a point which corresponded to the maximum specificity and sensitivity of the training set. Typically, the apex of the receiver operating characteristic (ROC) curve is chosen. However, the ROC curve for this particular dataset does not have a natural apex point, as shown in Figure IV-1. Therefore, we chose the cutoff point that maximized the sum of the sensitivity and specificity. The results for this chosen cutoff are shown in Table IV-3 and Table IV-4.

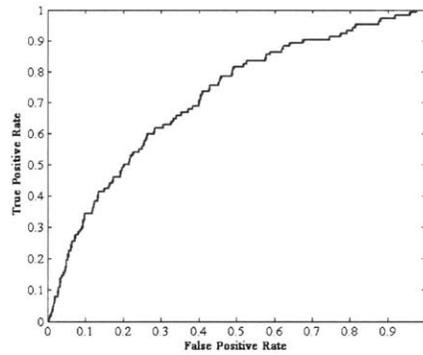


Figure IV-1. Receiver operating characteristic curve for WMV and cardiovascular death.

	Dataset	Hazard Ratios	Sensitivity	Specificity
Best Fit Alphas (Cutoff = 1.3037)	MERLIN, placebo	8.91	0.833	0.560
	DISPERSE	3.59	0.786	0.492
	MERLIN, ranolazine	3.78	0.648	0.577
Running Average (Cutoff = 1.3564)	MERLIN, placebo	6.24	0.618	0.717
	DISPERSE	7.15	0.786	0.659
	MERLIN, ranolazine	4.32	0.516	0.741

Table IV-3. Hazard ratio, sensitivity, and specificity using maximum sensitivity and specificity cutoff for cardiovascular death. Based on these results, the highlighted cutoff used for further studies was found by maximizing sensitivity and specificity, using a running window size of 0.007 Hz.

	Dataset	Hazard Ratios	Sensitivity	Specificity
Best Fit Alphas (Cutoff = 2.392)	MERLIN, placebo	2.35	0.489	0.696
	DISPERSE	0.54	0.250	0.602
	MERLIN, ranolazine	1.16	0.360	0.673
Running Average (Cutoff = 2.357)	MERLIN, placebo	2.37	0.532	0.620
	DISPERSE	0.73	0.392	0.510
	MERLIN, ranolazine	1.11	0.453	0.585

Table IV-4. Hazard ratio, sensitivity, and specificity using maximum sensitivity and specificity cutoff for myocardial infarction. The highlighted cutoff was used for further studies.

As shown in these results, the sensitivity and specificity cutoff perform better, based on hazard ratios, than using a cutoff based on hazard ratios or on quartile values. Using sensitivity and specificity to determine cutoffs is standard practice [27]. Therefore, the remainder of this chapter will use the sensitivity and specificity cutoff for determining high or low risk.

B. Cross-Validation

In order to show that the hazard ratios produced from our WMV values are relatively consistent, we performed a cross-validation. The MERLIN placebo group was randomly divided into two evenly-sized groups, with approximately equal numbers of patients who had a cardiovascular death in each group. Simulated annealing was performed on one of the groups, in order to find a set of weights and a cutoff. These weights and cutoffs were then tested on the second group, as well as the DISPERSE patients and MERLIN ranolazine patients to see if the hazard ratios were consistently high. This validation was performed 10 times. The results for the cross-validation on cardiovascular death are shown in Table IV-5, and the results for myocardial infarction are shown in Table IV-6. For posterity, only the MERLIN placebo training and testing group results were considered, so our final testing groups would truly be testing groups. The performance of WMV of the DISPERSE and MERLIN ranolazine groups were included in Table IV-5 and Table IV-6 for completeness, but were not considered when evaluating the WMV cross-validation results.

Dataset	Mean	Maximum	Minimum
MERLIN placebo, training set ($n = 1146 \pm 22$, $cvd = 52 \pm 5$)	11.01 ± 6.10	21.43	5.02
MERLIN placebo, testing set ($n = 1155 \pm 22$, $cvd = 50 \pm 5$)	4.47 ± 2.22	10.12	2.04
DISPERSE ($n = 764$, $cvd = 14$)	5.54 ± 1.81	8.12	2.77
MERLIN ranolazine ($n = 2254$, $cvd = 91$)	3.66 ± 0.81	5.37	2.04

Table IV-5. Cross-validation hazard ratios using WMV to risk stratify for cardiovascular

death. The average number of patients in the MERLIN placebo training group was 1146, with a standard deviation of 22. The group had an average of 52 ± 5 cardiovascular deaths. The MERLIN placebo test group had, on average, a patient population of size 1155 ± 22 , with 50 ± 5 cardiovascular deaths. The standard deviation of the hazard ratios are shown next to the mean hazard ratios. The maximum and minimum values of the hazard ratios found during the cross-validation are also shown for comparison.

Dataset	Mean	Maximum	Minimum
MERLIN placebo, training set ($n = 1146 \pm 22$, $MI = 92 \pm 6$)	2.61 ± 0.59	4.11	1.76
MERLIN placebo, testing set ($n = 1155 \pm 22$, $MI = 94 \pm 6$)	1.93 ± 0.33	2.60	1.32
DISPERSE ($n = 764$, $MI = 28$)	0.72 ± 0.23	1.15	0.44
MERLIN ranolazine ($n = 2254$, $MI = 161$)	1.15 ± 0.19	1.51	0.85

Table IV-6. Cross-validation hazard ratios using WMV to risk stratify for myocardial

infarction. The same groups of patients were sampled for cross-validation for myocardial infarctions as for cardiovascular death.

Based on the cross-validation results of Table IV-5 and Table IV-6, we used the entire MERLIN placebo group as a training set. For cardiovascular death, the variation in the hazard ratios was high in both the training and test sets. However, since the minimum hazard ratios were greater than 1, some useful information remained when risk-stratifying cardiovascular patients. For myocardial infarction, the variations in the hazard ratios were small, but the hazard ratios were also close to 1. We found a set of hazard ratios based on training WMV on all of the MERLIN placebo patients to risk stratify for myocardial infarctions, as well.

C. Risk Stratifying Cardiovascular Death

The variables that characterize WMV in its role of risk stratification are the weights for associated frequencies and the cutoff point that determines low versus high risk. The final WMV cutoff point determined from the sensitivity and specificity for cardiovascular death was found to be 1.356. Patients with $WMV > 1.356$ are at a higher risk of cardiovascular death. The weights that are associated with high risk are shown in Figure IV-2.

The hazard ratios for various risk stratifying measures for cardiovascular death are shown Table IV-7. The morphologic methods do not seem to perform as well on patients who were given ranolazine. HRV performs approximately equally well on all three groups. However, MV and WMV perform better on placebo groups than patients given ranolazine. The weights and cutoff found for WMV were trained on the MERLIN placebo group and tested on the DISPERSE and MERLIN ranolazine groups. MV was trained on DISPERSE patients [20]. Therefore, the true group used for comparison of performance is the MERLIN ranolazine group, in which WMV has a higher risk stratification capability. The associated Kaplan-Meier hazard curves are shown in Figure IV-3.

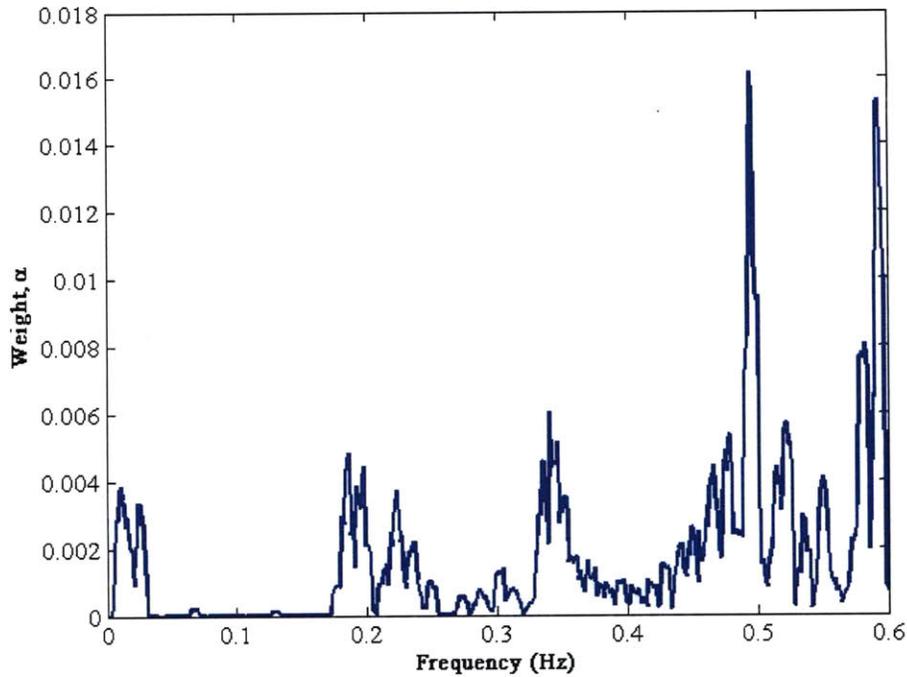


Figure IV-2. Weights for various frequencies in WMV risk stratification of cardiovascular death. This figure depicts the smoothed version of the weights, using a window size of 0.007 Hz.

Risk Stratification Metric	MERLIN placebo group <i>(n = 2301, cvd = 102)</i>	DISPERSE <i>(n = 764, cvd = 14)</i>	MERLIN ranolazine group <i>(n = 2254, cvd = 91)</i>
HRV (LF/HF) < 1.2	3.03	2.83	2.91
MV > 52.5	5.12	8.46	3.85
WMV > 1.356	6.24	7.15	4.32

Table IV-7. Hazard ratio for risk stratification measures of cardiovascular death on MERLIN and DISPERSE patients. For each column, the number of patients (n) and the number of cardiovascular deaths (cvd) are listed for the populations.

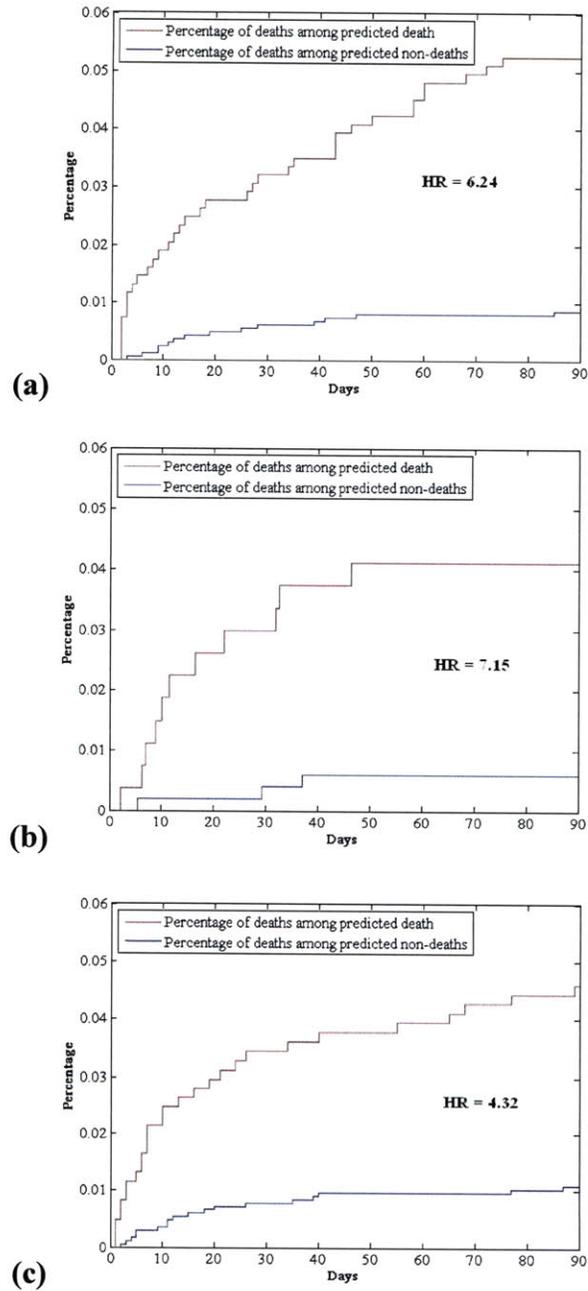


Figure IV-3. Kaplan Meier hazard curves for MERLIN and DISPERSE populations using $WMV > 1.356$. Figure IV-3a is the Kaplan Meier curve based on the MERLIN placebo population, Figure IV-3b is based on the DISPERSE population, and Figure IV-3c is based on the MERLIN ranolazine population.

D. Risk Stratifying Myocardial Infarction

Though the cross-validation results for WMV and myocardial infarction (MI) did not show good risk stratification, the results of risk stratification of MI using WMV are shown below. Patients with WMV > 2.357 are at higher risk for myocardial infarctions. The associated weights are shown in Figure IV-4.

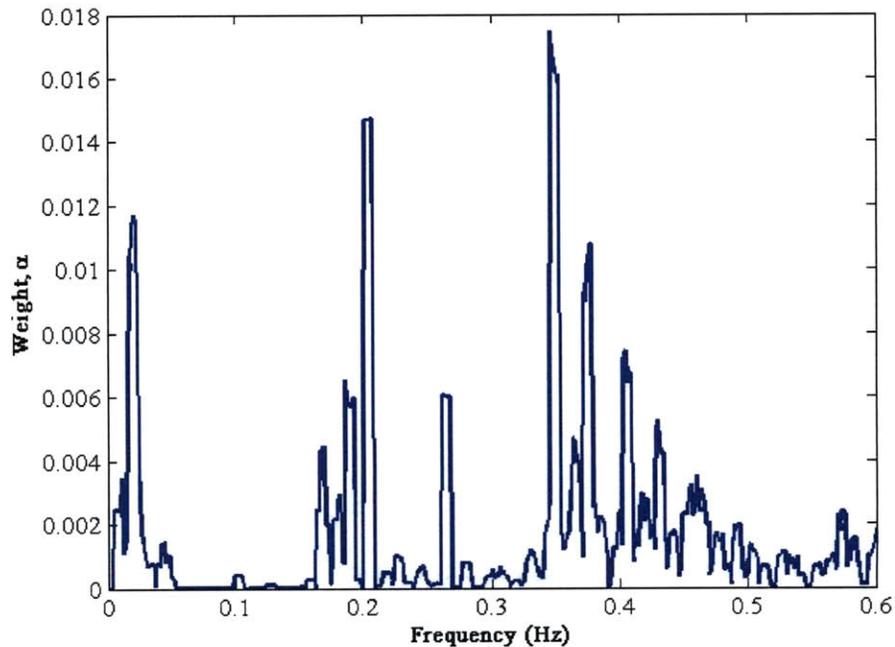


Figure IV-4. *Weights for various frequencies in WMV risk stratification of myocardial infarction.* This figure depicts the smoothed version of the weights, using a window size of 0.007 Hz.

Table IV-8 shows the hazard ratios associated with various electrocardiographic risk stratification metrics for MIs. Most of these electrocardiographic metrics do not predict risk of MI. HRV predicts MI risk with a hazard ratio of about 1.5. However, MV does not perform as well. WMV was trained on the MERLIN placebo group. WMV predicts the DISPERSE and

MERLIN ranolazine testing groups worse than MV predicts. The associated Kaplan-Meier hazard curves are shown in Figure IV-5.

Risk Stratification Metric	MERLIN placebo group (<i>n</i> = 2301, <i>MI</i> = 186)	DISPERSE (<i>n</i> = 764, <i>MI</i> = 28)	MERLIN ranolazine group (<i>n</i> = 2254, <i>MI</i> = 161)
HRV (LF/HF) < 1.2	1.58	1.48	1.67
MV > 52.5	1.28	1.02	1.17
WMV > 2.357	2.37	0.73	1.11

Table IV-8. Hazard ratio for risk stratification measures of myocardial infarction on

MERLIN and DISPERSE patients. For each column, the number of patients (*n*) and the number of myocardial infarctions (*MI*) are listed for the populations. It is important to note that the MV cutoff of 52.5 was found based on risk stratification of cardiovascular death. Therefore, an MV cutoff of 52.5 is not appropriate. For the purposes of this table, however, the 52.5 cutoff is used.

Based on these results, weighted morphologic variability is a new electrocardiographic method that is able to risk stratify patients for cardiovascular death better than current electrocardiographic risk-stratification methods. Weighted morphologic variability is not a good risk stratification method for myocardial infarctions, however. We will discuss the implications of the results and derive a physiological explanation for this work in Chapter 5.

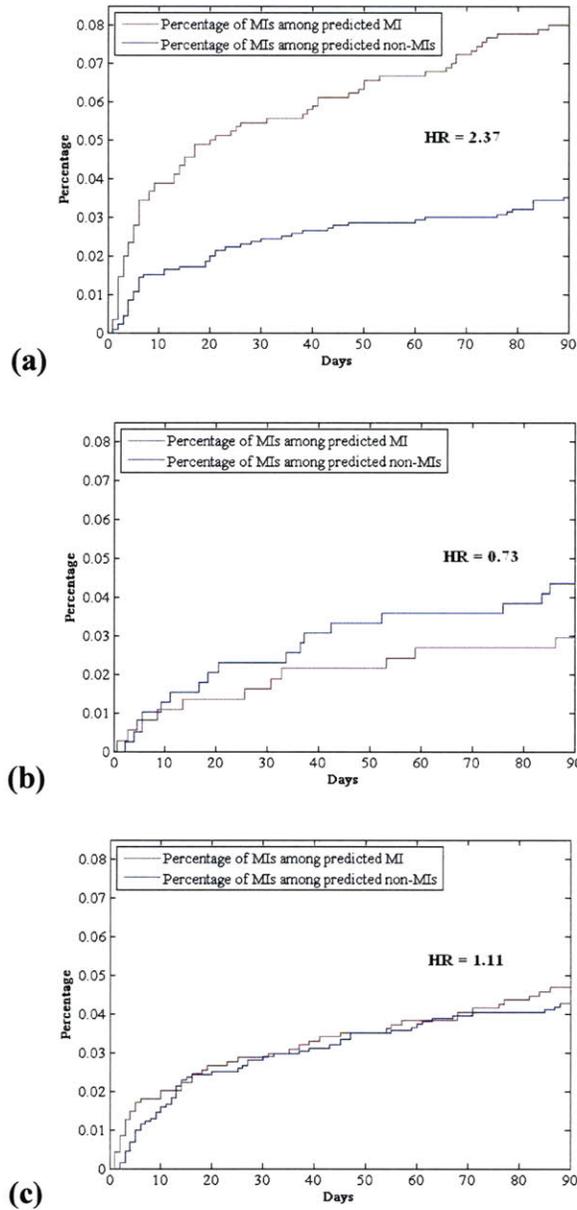


Figure IV-5. Kaplan Meier hazard curves for MERLIN and DISPERSE populations using $WMV > 2.357$. Figure IV-5a is based on the MERLIN placebo training population, Figure IV-5b is based on the DISPERSE population, and Figure IV-5c is based on the MERLIN ranolazine population.

Chapter V:

SUMMARY AND CONCLUSIONS

A. Summary

Patients who have had an acute coronary syndrome remain at risk for future adverse cardiovascular events. Electrocardiographic and other types of risk stratification metrics may aid in the identification of patients who are at high risk of cardiovascular death. However, better risk stratification metrics are needed. Weighted morphologic variability is an inexpensive and accurate way of risk stratifying patients for cardiovascular death that outperforms other ECG-based risk stratification methods.

This thesis derived the calculation of weighted morphologic variability (WMV) and evaluated its performance on risk stratifying for adverse cardiovascular events. WMV is based on an electrocardiographic technique called morphologic variability (MV), which risk stratifies based on subtle changes in the beat-to-beat morphology of an ECG-signal. MV uses a technique called dynamic time warping to quantify beat-to-beat morphologic differences. These beat-to-beat differences yield a morphologic distance (MD) time series. MV is the sum of the power over a particular frequency band within the power spectral density of the MD time series.

Weighted morphologic variability uses the MD time series, as well, to predict cardiovascular death. Instead of using the power of one particular frequency band, WMV uses a combination of frequency bands, where a weight is associated with the various frequency bands.

Each weight reflects the importance of that particular frequency band on the predictive ability of WMV. WMV is the sum of the weighted sums of the power found in various frequency bands.

In order to find the optimum set of weights and frequency bands to compute weighted morphologic variability, we used a technique called simulated annealing. First, we discretized the frequency band 0.001 Hz-0.600 Hz into 0.001 Hz intervals. Then, simulated annealing was performed to find the optimal weights that accurately distinguish high risk from low risk patients. Simulated annealing randomly samples an energy space to find the global minimum of the energy. For our purposes, the negative of the c-statistic was used as a measure of energy. The c-statistic describes the probability of correctly identifying the patient with a disease given two patients, where one has the disease and the other does not. Therefore, a c-statistic of 0.5 indicates that the test is no better than random guessing. For each simulated annealing random sampling, a set of weights were generated. The WMV for all training patients were computed, and a c-statistic was produced. Simulated annealing converged on a solution of weights that produced the highest c-statistic, or the lowest negative of the c-statistic. These weights were then smoothed and used to evaluate the risk stratification performance of WMV on various adverse cardiovascular events.

In order to evaluate cardiovascular performance, we first created a cutoff to determine high and low risk individuals. The maximum sum of sensitivity and specificity was found to determine the difference between high and low risk individuals. A hazard ratio is a ratio of the death rates of high risk individuals compared to low risk individuals and was calculated comparing the two groups based on the WMV cutoff. A cross-validation was performed to ensure that the hazard ratio for the training group was valid. The WMV hazard ratio for risk stratifying cardiovascular death was found to be 7.15 on the test group of 764 patients, including

14 cardiovascular deaths. WMV produced a higher hazard ratio for cardiovascular death than morphologic variability and heart rate variability hazard ratios. However, the hazard ratio for myocardial infarctions using WMV was 0.73 on the same group, with 28 myocardial infarctions. Thus, WMV is a novel electrocardiographic metric that risk stratifies patients for cardiovascular death, but not myocardial infarctions.

B. Physiological Basis

In this section, we propose a physiological interpretation of the weights in WMV associated with high risk of cardiovascular death. The peaks in the weights for weighted morphologic variability may coincide with particular changes in the conduction path of the myocardium. We propose evidence for these conduction path changes, and future work that may improve the prognostic ability of WMV.

In certain types of arrhythmias, the heart conducts in an unusual path in fixed intervals. T-wave alternans is a phenomenon where the T-wave of the ECG exhibits beat-to-beat changes in amplitude. More precisely, in T-wave alternans, a beat with a relatively high amplitude T-wave is followed by a beat with relatively low amplitude; e.g., in an *ABABAB* fashion. Ventricular bigeminy and trigeminy are arrhythmias characterized by a premature ventricular contraction (PVC) every second or third beat, respectively, thus following an *ABABAB* or *AABAABAAB* pattern. In Sung [23], a physiological rationale is discussed for why this may occur. A myocardial cell or a group of cells may take longer to depolarize or repolarize compared to the rest of the heart, and may only depolarize every N heartbeats. The heart depolarizes along a typical conduction path, but when a group of cells takes longer to depolarize, the heart may depolarize along an alternate conduction path. Thus, the ECG may depict two

morphologies: one morphology associated with the typical conduction path, and another morphology associated with the abnormal conduction path. Thus, a morphologic pattern may occur.

WMV weights different frequencies of the ECG based on its importance in risk stratification. The WMV weights that are most associated with cardiovascular death were described in this thesis. We propose that the peaks in the weights are associated with frequencies that may be typical of some sort of alternans, where the ECG morphology follows a pattern of switching between two morphologies. A typical heart rate is somewhere between 60-100 beats per minute. 100 beats per minute is equivalent to 0.6 Hz, which was the maximum frequency we tested. A true beat-to-beat alternans, following an *ABABAB* pattern, would show morphology differences once every two beats, or between 0.3-0.5 Hz. An anomaly that occurs every third beat, or following an *AABAABAAB* pattern, would show morphology differences in the 0.2-0.33 Hz range. As shown in Figure IV-2 on page 53, peaks in the weights occur at 0.2 Hz, 0.34 Hz, and 0.5 Hz, suggesting that the above physiological explanation is plausible. Based on this concept, one can extend the idea of WMV by weighting frequencies that correspond to half or a third of the heart rate more heavily as a risk-stratification tool.

C. Conclusions

Weighted morphologic variability is a new electrocardiographic prognostic tool that outperforms other electrocardiographic risk-stratification methods. WMV is better at identifying high risk patients than MV and HRV. These electrocardiographic tools are helpful to physicians to identify which patients to aggressively treat to prevent adverse cardiovascular outcomes. Along with the results from imaging tests, the presence of certain biomarkers, and catheterization

tests, doctors can use WMV and other ECG methods to gather a holistic view of the cardiac health of a patient. Physicians can then use these results to perfect a patient's initial prognosis and treatment options.

Electrocardiographic risk-stratification metrics present an inexpensive, effective, and efficient way of prognosticating patients. ECGs are easily recorded from a patient, and are essentially free and minimally invasive. The WMV can be quickly calculated from the ECG at essentially no additional cost. In contrast, biomarker tests require time for a lab to process the results. Imaging tests require an expert for analysis. Catheterization is invasive, can pose a risk to the patient, and requires a skilled clinician to perform the procedure. WMV and other electrocardiographic metrics, on the other hand, only have a numerical cutoff that determines whether a patient is at high risk or low risk for a cardiovascular event. Based on ECG metrics, patients who are determined to be low risk for cardiovascular events can be spared from invasive and more expensive tests, and patients who are at high risk for cardiovascular events can undergo further risk stratification that may not have been considered otherwise.

WMV is a powerful prognostic tool that can be integrated into hospitals to risk stratify patients for cardiovascular death. WMV can be easily implemented in a computer or an ECG-recording device. The results can also be readily interpreted by a clinician. This simple, cost-effective metric should help improve the care for the millions of people who are admitted to a hospital with an acute coronary syndrome each year. This body of work demonstrates why ECG-based risk-stratification metrics should be widely and commonly implemented in clinical practice.

BIBLIOGRAPHY

1. DeFrances, C.J., et al., 2006 National Hospital Discharge Survey. Natl Health Stat Report, 2008(5): p. 1-20.
2. Armstrong, P.W., et al., Acute coronary syndromes in the GUSTO-IIb trial: prognostic insights and impact of recurrent ischemia. The GUSTO-IIb Investigators. *Circulation*, 1998. 98(18): p. 1860-8.
3. St Goar, F.G., et al., Detection of coronary atherosclerosis in young adult hearts using intravascular ultrasound. *Circulation*, 1992. 86(3): p. 756-63.
4. Hartzog, K.C. Cardiovascular System. [cited 2010 May 18, 2010]; Available from: <http://www.starsandseas.com/SAS%20Physiology/Cardiovascular/Cardiovascular.htm>.
5. Lemkin, D.P., Gary. Electrophysiology. Introduction to EKGs [cited 2010 May 18, 2010]; Available from: <http://davidge2.umaryland.edu/~emig/ekgtu02.html>.
6. Antman, E.M., et al., The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. *JAMA*, 2000. 284(7): p. 835-42.
7. Lilly, L.S. and Harvard Medical School., Pathophysiology of heart disease : A collaborative project of medical students and faculty. 3rd ed. 2003, Philadelphia: Lippincott Williams & Wilkins. xiii, 445 p.
8. Solomon, S.D., et al., Influence of ejection fraction on cardiovascular outcomes in a broad spectrum of heart failure patients. *Circulation*, 2005. 112(24): p. 3738-44.

9. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, 1996. 93(5): p. 1043-65.
10. Kleiger, R.E., et al., Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am J Cardiol*, 1987. 59(4): p. 256-62.
11. Singh, N., et al., Heart rate variability assessment early after acute myocardial infarction. Pathophysiological and prognostic correlates. GUSTO ECG Substudy Investigators. Global Utilization of Streptokinase and TPA for Occluded Arteries. *Circulation*, 1996. 93(7): p. 1388-95.
12. Heart Rate Variability Analysis Scientific Background. *Heart Rate Variability 2009* [cited 2010 May 19, 2010]; Available from: <http://www.biocomtech.com/hrvscientific>.
13. Barthel, P., et al., Risk stratification after acute myocardial infarction by heart rate turbulence. *Circulation*, 2003. 108(10): p. 1221-6.
14. Goldberger, J.J., et al., American Heart Association/American College of Cardiology Foundation/Heart Rhythm Society Scientific Statement on Noninvasive Risk Stratification Techniques for Identifying Patients at Risk for Sudden Cardiac Death. A scientific statement from the American Heart Association Council on Clinical Cardiology Committee on Electrocardiography and Arrhythmias and Council on Epidemiology and Prevention. *J Am Coll Cardiol*, 2008. 52(14): p. 1179-99.
15. Bauer, A., et al., Heart rate turbulence: standards of measurement, physiological interpretation, and clinical use: International Society for Holter and Noninvasive Electrophysiology Consensus. *J Am Coll Cardiol*, 2008. 52(17): p. 1353-65.

16. Bauer, A., et al., Deceleration capacity of heart rate as a predictor of mortality after myocardial infarction: cohort study. *Lancet*, 2006. 367(9523): p. 1674-81.
17. Rosenbaum, D.S., et al., Electrical alternans and vulnerability to ventricular arrhythmias. *N Engl J Med*, 1994. 330(4): p. 235-41.
18. Cannon, C.P., et al., Safety, tolerability, and initial efficacy of AZD6140, the first reversible oral adenosine diphosphate receptor antagonist, compared with clopidogrel, in patients with non-ST-segment elevation acute coronary syndrome: primary results of the DISPERSE-2 trial. *J Am Coll Cardiol*, 2007. 50(19): p. 1844-51.
19. Morrow, D.A., et al., Effects of ranolazine on recurrent cardiovascular events in patients with non-ST-elevation acute coronary syndromes: the MERLIN-TIMI 36 randomized trial. *JAMA*, 2007. 297(16): p. 1775-83.
20. Syed, Z., et al., Relation of death within 90 days of non-ST-elevation acute coronary syndromes to variability in electrocardiographic morphology. *Am J Cardiol*, 2009. 103(3): p. 307-11.
21. Donoho, D.L., De-Noising by Soft-Thresholding. *IEEE Transactions on Information Theory*, 1995. 41(3): p. 613-627.
22. Li, Q., R.G. Mark, and G.D. Clifford, Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiol Meas*, 2008. 29(1): p. 15-32.
23. Sung, P.S., Risk Stratification By Analysis of Electrocardiographic Morphology Following Acute Coronary Syndromes, in *Electrical Engineering and Computer Science*. 2007, Massachusetts Institute of Technology: Cambridge. p. 100.

24. Kirkpatrick, S., C.D. Gelatt, Jr., and M.P. Vecchi, Optimization by Simulated Annealing. *Science*, 1983. 220(4598): p. 671-680.
25. Stultz, C.M. and M. Karplus, Dynamic ligand design and combinatorial optimization: designing inhibitors to endothiapepsin. *Proteins*, 2000. 40(2): p. 258-89.
26. Cox, D.R., Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1972. 34(2): p. 187-220.
27. Florkowski, C.M., Sensitivity, Specificity, Receiver-Operating Characteristic (ROC) Curves and Likelihood Ratios: Communicating the Performance of Diagnostic Tests. *Clin Biochem Rev*, 2008. 29 Suppl 1: p. S83-7.