

# Scalable Noise Mining in Long-Term Electrocardiographic Time-Series to Predict Death Following Heart Attacks

Chih-Chun Chia, Zeeshan Syed  
University of Michigan  
2260 Hayward St.  
Ann Arbor, MI, U.S.A  
{jazzchia,zhs}@eecs.umich.edu

## ABSTRACT

Cardiac disease is the leading cause of death around the world; with ischemic heart disease alone claiming 7 million lives in 2011. This burden can be attributed, in part, to the absence of biomarkers that can reliably identify high risk patients and match them to treatments that are appropriate for them. In recent clinical studies, we have demonstrated the ability of computation to extract information with substantial prognostic utility that is typically disregarded in time-series data collected from cardiac patients. Of particular interest are subtle variations in long-term electrocardiographic (ECG) data that are usually overlooked as noise but provide a useful assessment of myocardial instability. In multiple clinical cohorts, we have developed the pathophysiological basis for studying probabilistic variations in long-term ECG and demonstrated the ability of this information to effectively risk stratify patients at risk of dying following heart attacks. In this paper, we extend this work and focus on the question of how to reduce its computational complexity for scalable use in large datasets or energy constrained embedded devices. Our basic approach to uncovering pathological structure within the ECG focuses on characterizing beat-to-beat time-warped shape deformations of the ECG using a modified dynamic time-warping (DTW) and Lomb-Scargle periodogram-based algorithm. As part of our efforts to scale this work up, we explore a novel approach to address the quadratic runtime of DTW. We achieve this by developing the idea of adaptive downsampling to reduce the size of the inputs presented to DTW, and describe changes to the dynamic programming problem underlying DTW to exploit adaptively downsampled ECG signals. When evaluated on data from 765 patients in the DISPERSE2-TIMI33 trial, our results show that high morphologic variability is associated with an 8- to 9-fold increased risk of death within 90 days of a heart attack. Moreover, the use of adaptive downsampling with a modified DTW formulation achieves a 7- to almost 20-fold reduction in runtime relative to DTW, without a significant change in biomarker discrimination.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

KDD '14, August 24–27, 2014, New York, NY, USA.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-2956-9/14/08 ...\$15.00.

<http://dx.doi.org/10.1145/2623330.2623702>

## Categories and Subject Descriptors

J.3 [Computer Applications]: Life and Medical Sciences; H.4 [Information Systems]: Information System Applications; I.2.1 [Computing Methodologies]: Applications and Expert Systems Medicine and science

## Keywords

cardiovascular disease; electrocardiogram; adaptive downsampling; time-warping; noise mining

## 1. INTRODUCTION

Heart disease is the leading cause of death around the world. The burden of ischemic heart disease alone in 2011 was 7 million lives [23]. In the United States, heart disease claims 830,000 lives each year (34% of all deaths, or one death every 38 seconds) [14]. Nearly 151,000 of these deaths take place in patients under the age of 65, and a third occur before the age of 75 [14]. Much of the difficulty in reducing this burden follows from an inability to match cardiac patients to treatments that are most appropriate for their individual risk. As one example of this, devices such as implantable cardioverter defibrillators (ICDs) can be life-saving for patients who experience fatal arrhythmias (over 300,000 sudden cardiac deaths in the U.S. each year among patients with diagnosed coronary disease) [16]. In most of these cases, the effects of the arrhythmia can be reversed if the victim is treated with an electrical shock within the first few minutes. However, existing decision-making methods fail to prescribe ICDs to the majority of patients who die [4]. Conversely, most of the patients who do presently receive an ICD do not receive any benefit from their device [4], resulting in an unnecessary risk to patients and unnecessary costs to the healthcare system.

Physicians use a variety of biomarkers to estimate patient risk and to match patients to treatments. In the setting of heart disease, these biomarkers are typically limited to information available through blood-based measurements of biochemical substrates (e.g., troponin I, C-reactive protein, and brain natriuretic peptide), or through imaging (e.g., left ventricular ejection fraction obtained through echocardiography) [12, 10]. In both these cases, the focus is on studying information that is present in instantaneous (i.e., ‘snapshot’) data, and where this information can be directly measured with limited or no computational aid. Despite these efforts, however, finding biomarkers that can accurately assess patient risk remains a challenge. For instance, while depressed left ventricular ejection fraction is commonly used to iden-

tify high risk patients following heart attacks, the absolute number of deaths is far greater among patients with relatively preserved left ventricular function [8].

Our recent research aims to address this situation through novel biomarkers that are computationally derived from physiological time-series, and that are designed to offer complementary information to existing blood- or imaging-based markers. Of particular interest is information potentially available in long-term electrocardiographic (ECG) signals. The ECG provides a continuous assessment of the electrical activity of the heart, and is routinely collected from patients during hospitalization to determine heart rate and detect arrhythmias. ECG has the advantage of being easy to acquire; the electrical activity of the heart can be measured on the surface of the body in an inexpensive and non-invasive manner over long periods. In an in-patient setting, the ECG is typically captured by bedside monitors. In an out-patient setting, a Holter monitor (a portable ECG device worn by patients) can record data continuously over multiple days. Since the ECG is routinely collected from patients in a wide variety of clinical settings, computational biomarkers deriving from long-term ECG time-series can be incorporated broadly into clinical practice without the need for any new hardware or without creating any additional burden on patients or caregivers.

In this paper, we focus on the question of how we can leverage long-term recordings of ECG activity (over days) to discover information that is unrecognized in previous studies centered on small snippets of data (over seconds). Our recent experiments, reported in the clinical literature [21], have approached this question by exploring the hypothesis that much of what is commonly perceived as noise in ECG signals may contain subtle but useful information about the health of the heart. The theory underlying this work is that increased variability in the morphology of ECG time-series is likely associated with a lack of consistency and repeatability in the electrical function of the heart. In other words, persistent fluctuations in the shape of the ECG waveform may suggest electrical instability in the heart muscle predisposing patients to fatal arrhythmias. The challenge in detecting this variability, however, is in being able to distinguish between shape deformations associated with pathological phenomena reflecting the health of the underlying heart muscle, and changes associated with artifacts that represent true noise. Making this distinction is difficult in short ECG recordings, but with the availability of long-term ECG time-series, pathological variations can be distinguished from true noise as structure that is persistent over long periods of time.

In preliminary clinical studies, we have developed the computationally generated ECG biomarker of pathological morphologic variability (which we subsequently refer to only as morphologic variability or MV) in long-term ECG to predict death following heart attacks. As a first step, our approach to measuring MV makes use of a modified dynamic time-warping (DTW)-based algorithm to quantify time-warped shape deformations in ECG time-series over long periods of time. As a second step, we draw upon a Lomb-Scargle periodogram approach to analyze the resulting non-uniformly sampled time-series representation of aggregate noise in the ECG for pathological structure. While the use of this methodology has shown promise in clinical cohorts [21] and provides information that is complementary to existing markers based on patient history and physical exam findings, echocardi-

graphy, and blood-based laboratory reports, our prior work has demonstrated the clinical utility of MV and not focused on the question of how to scale this basic approach to large databases of long-term ECG time-series or use in embedded devices (e.g., pacemakers and ICDs). To achieve this, we investigate here a novel approach that reduces the computational runtime of DTW through an adaptive downsampling of time-series inputs. The use of adaptive downsampling significantly reduces the ECG data presented to DTW while preserving rapidly changing waves (e.g., the QRS complex) smoothed out by existing downsampling approaches. However, due to adaptively downsampling rapidly changing parts of the ECG time-series less than more slowly changing parts of the signals, this approach also requires changes to the dynamic programming problem underlying DTW. In this paper, we present solutions for both the goal of adaptively downsampling ECG time-series, and for modifying the DTW dynamic programming formulation to leverage adaptively downsampled inputs.

We evaluate our ideas on data from 765 patients presenting with acute coronary syndrome in the DISPERSE2-TIMI33 trial. Our results show that high MV is associated with an 8- to 9-fold increased risk of death within 90 days of a heart attack. Moreover, the use of adaptive downsampling with a modified DTW formulation achieves an almost 4-fold reduction in runtime relative to DTW, without a significant change in biomarker discrimination. In contrast, existing downsampling approaches obtain a similar reduction in runtime but with noticeably worse performance for risk prediction.

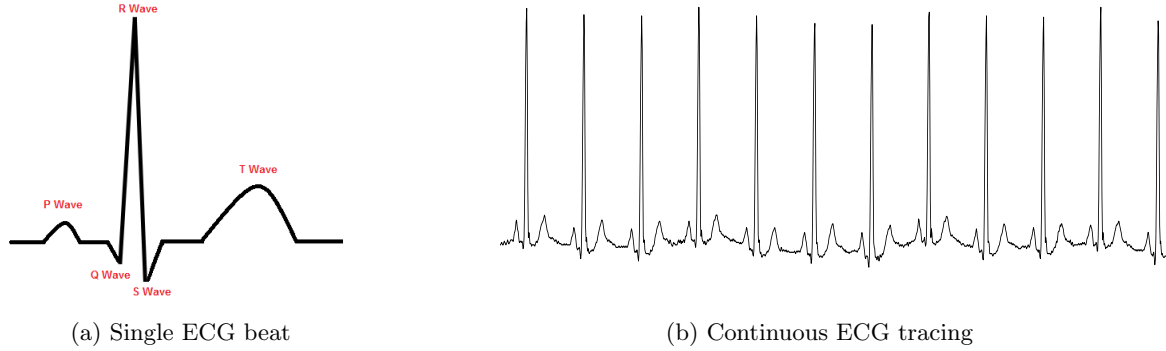
The remainder of this paper is organized as follows. Section 2 presents background on ECG time-series and the pathophysiological basis for our efforts to mine noise in ECG time-series for information relevant to cardiac risk prediction. Section 3 describes the methodology for measuring MV using a modified DTW- and Lomb-Scargle periodogram-based approach. Section 4 proposes the concept of adaptive downsampling, and details how adaptive downsampling can be incorporated within the measurement of MV to scale it up to large amount of long-term time-series data. Section 5 presents the datasets and evaluation methodology for our study. Section 6 discusses the results of our experiments. Section 7 offers a summary and conclusions.

## 2. BACKGROUND

We begin with a brief background on ECG time-series and on the pathophysiological basis for mining noise in ECG time-series for information useful in predicting future cardiac risk.

### 2.1 Electrocardiogram

The ECG is a continuous recording of the electrical activity of the heart muscle or myocardium[13]. At rest, each cardiac muscle cell maintains a voltage difference across its cell membrane. During depolarization (i.e., the ‘firing’ of the heart muscle), this voltage increases. Consequently, when depolarization is propagating through a cell, there exists a potential difference on the membrane between the part of the cell that has been depolarized and the part of the cell at resting potential. After the cell is completely depolarized, its membrane is uniformly charged again, but at a more positive voltage than initially. The reverse situation takes place during repolarization, which returns the cell to base-



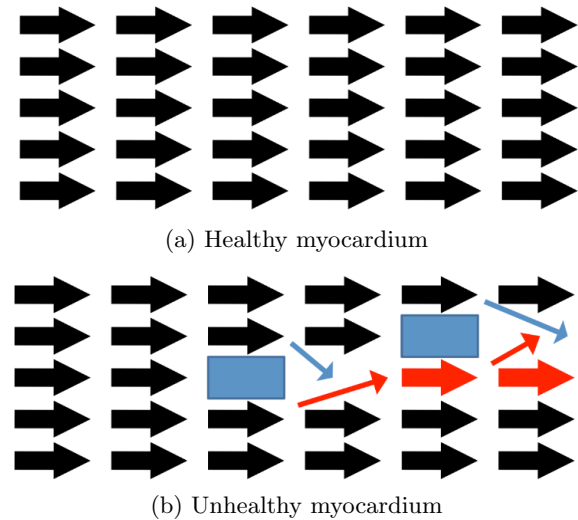
**Figure 1:** (a) Schematic representation of the normal ECG for a single heart beat, and (b) example recording of ECG waveform.

line. These changes in potential, summed over many cells, can be measured by electrodes placed on the surface of the body, leading to the ECG time-series.

The ECG is a quasi-periodic signal (i.e., corresponding to the quasi-periodic nature of cardiac activity). Three major segments can be identified in a normal ECG. The P wave is associated with depolarization of cardiac cells in the upper two chambers of the heart (i.e., the atria). The QRS complex (comprising the Q, R and S waves) is associated with depolarization of cardiac cells in the lower two chambers of the heart (i.e., the ventricles). The T wave is associated with repolarization of the cardiac cells in the ventricles. The QRS complex is larger than the P wave because the ventricles are much larger than the atria. The QRS complex also coincides with the repolarization of the atria, which is therefore usually not seen on the ECG. The T wave has a larger width and smaller amplitude than the QRS complex because repolarization takes longer than depolarization[13]. Figure 1(a) presents a schematic representation of the normal ECG, while Figure 1(b) shows an example tracing of a continuous ECG time-series over a few seconds.

## 2.2 Pathophysiology

In a healthy myocardium, the pathways of depolarization through excitable cells are usually similar for consecutive heart beats. However, in the presence of abnormalities, the conducting system has multiple irregular islands of severely depressed and unexcitable myocardium [7] (Figure 2). These lead to discontinuous electrical characteristics of the heart muscle, and create a situation analogous to the presence of race conditions within the heart. The presence of several possible adjacent pathways (e.g., the pathways shown in red and blue in Figure 2) that can potentially invade the non-functioning area leads to variations in the spatial direction of the invading vector [2]. Measured electrical activity in this setting is best described in a highly variable form, stemming from subtle unstable conduction bifurcations. Furthermore, the propagation of a beat may be dependent on the route of propagation of the previous beat. The overall effect of such minor conduction inhomogeneities is not well understood, but it is believed that they correlate with myocardial electrical instability and have potentially predictive value for ventricular arrhythmias [2] or other adverse events.

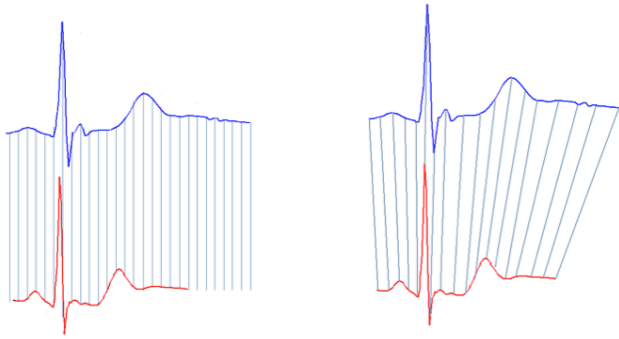


**Figure 2:** Illustration of process underlying noise-like pathological variability in ECG time-series. In healthy myocardium, electrical impulses are conducted smoothly through the heart muscle. In unhealthy myocardium (e.g., due to the presence of ischemia), unstable conduction bifurcations result in the path of impulse propagation, and consequently the ECG morphology, changing from beat to beat.

## 3. MORPHOLOGIC VARIABILITY (MV)

In this section, we describe our basic approach to measure MV in ECG time-series. In developing this methodology, we note that key challenges to measuring pathological variability that is often overlooked as noise in ECG signals include dealing with the presence of true noise and time-skew in these data, and addressing the need to discover potentially low amplitude disease signatures in the presence of high amplitude baseline activity. Section 4 describes how this basic approach can be scaled to very large volumes of ECG time-series through the use of adaptive downsampling.

For every pair of consecutively occurring beats in an ECG time-series, we quantify how the shapes of the beats dif-



**Figure 3: Comparison of time-warped shape deformations in ECG beats using DTW. In contrast to comparing activity that is time-aligned but not physiologically aligned (left), we use DTW in our study to relate similar parts of the ECG waveforms across beats in the presence of time skew.**

fer using a variant of DTW (Figure 3). Given time-series  $Q = q_1, \dots, q_n$  and  $C = c_1, \dots, c_m$ , DTW first constructs an  $n$ -by- $m$  distance matrix where each entry  $(i, j)$  represents the distance  $d(q_i, c_j)$ . The  $l_2$  norm is typically used to measure  $d(q_i, c_j)$ . DTW then finds the minimum cost path  $W = w_1, \dots, w_k, \dots, w_K$  through this distance matrix where  $w_k = (i_k, j_k)$  relates the  $i_k$ -th sample of  $Q$  to the  $j_k$ -th sample of  $C$ . The minimum cost path has the cost:

$$\sum_{k=1}^K d(q_{i_k}, c_{j_k})$$

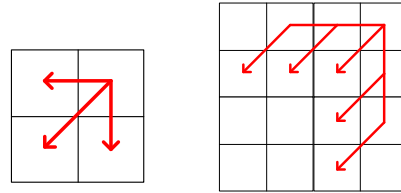
and is subject to several constraints, including boundary conditions, continuity, and monotonicity[3]. This optimal path can be found efficiently using dynamic programming with the following recurrence:

$$\gamma(i, j) = d(q_i, c_j) + \min \begin{cases} \gamma(i-1, j-1) \\ \gamma(i-1, j) \\ \gamma(i, j-1) \end{cases}$$

where  $\gamma(i, j)$  is the cumulative distance of the path from the start to cells  $(i, j)$ . From simple observation,  $DTW(Q, C) = \gamma(n, m)$  and the time and space complexity of this method is  $O(nm)$ .

We restrict the local range of the alignment path in the vicinity of a point to prevent biologically implausible alignments of large parts of one beat with small parts of another. For example, for an entry  $(i, j)$  in the distance matrix, we only allow valid paths passing through  $(i-1, j-1)$ ,  $(i-1, j-2)$ ,  $(i-2, j-1)$ ,  $(i-1, j-3)$  and  $(i-3, j-1)$ . This is an adaptation of the Type III and Type IV local continuity constraints proposed by Myers et al. [17] and ensures that there are no long horizontal or vertical edges along the optimal path through the distance matrix, corresponding to a large number of different samples in one beat being aligned with a single sample in the other. This leads to the following recurrence relation (also shown graphically in Figure 4):

$$\gamma(i, j) = d(q_i, c_j) + \min \begin{cases} \gamma(i-1, j-1) \\ d(q_{i-1}, c_j) + \gamma(i-2, j-1) \\ d(q_{i-1}, c_j) + d(q_{i-2}, c_j) + \gamma(i-3, j-1) \\ d(q_i, c_{j-1}) + \gamma(i-1, j-2) \\ d(q_i, c_{j-1}) + d(q_i, c_{j-2}) + \gamma(i-1, j-3) \end{cases}$$



(a) Original DTW (b) Constrained DTW

**Figure 4: Illustration of possible path alignments.**

The process described here transforms the original ECG time-series into a sequence of time-warped morphology differences between consecutive beats. To characterize pathological structure within this sequence, we study its spectral characteristics. Since the activity of the heart is quasi-periodic (i.e., since the heart does not beat at an exact rate), the time gap between the samples of the sequence constructed through DTW is not uniform. We address this issue by estimating the power spectral density of the morphology differences time-series using the Lomb-Scargle periodogram [15]. For a time series where the value  $m[n]$  is sampled at time  $t[n]$ , the Lomb-Scargle periodogram estimates the energy at frequency  $\omega$  as:

$$P(\omega) = \frac{1}{2\sigma^2} \left( \frac{\sum_n [(m[n] - \mu) \cos \omega(t[n] - \tau)]^2}{\sum_n \cos^2 \omega(t[n] - \tau)} + \frac{\sum_n [(m[n] - \mu) \sin \omega(t[n] - \tau)]^2}{\sum_n \sin^2 \omega(t[n] - \tau)} \right)$$

where  $\mu$  and  $\sigma$  are the mean and variance of the  $m[n]$ , and  $\tau$  is defined as :

$$\tan(2\omega\tau) = \frac{\sum_n \sin(2\omega t[n])}{\sum_n \cos(2\omega t[n])}$$

We define our computationally generated biomarker, MV, as energy between 0.30 and 0.55 Hz (as estimated from the Lomb-Scargle periodogram) in the time-series of aggregate morphology changes constructed using DTW. The range of 0.30 to 0.55 Hz is based on theoretical and empirical observations suggesting that the discriminative ability of MV for predicting death following heart attacks is maximized over this range [22]. A flow chart of the whole process for generating MV is shown in Figure 5.

## 4. SCALING TO LARGE DATASETS

The basic methodology described in Section 3 parallels the approach investigated in our earlier clinical experiments [21]. Here, we focus on the question of how this methodology can be scaled for use with very large volumes of ECG data.

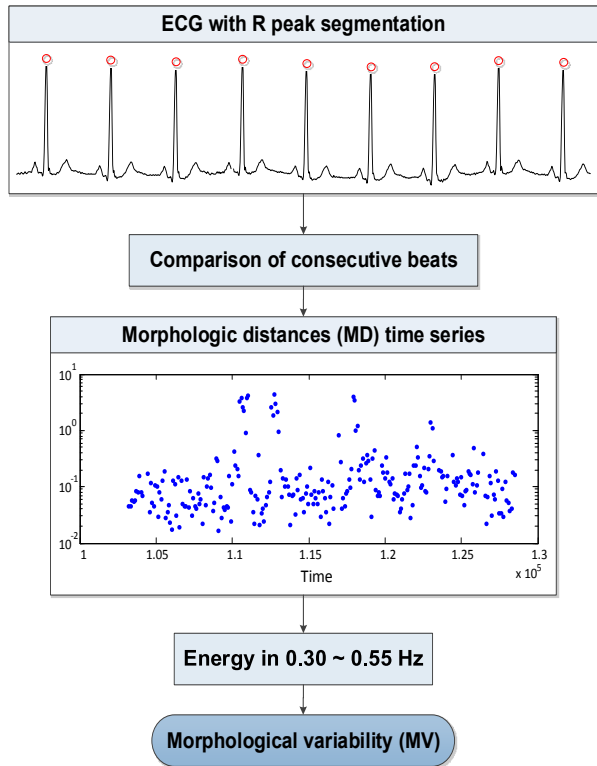


Figure 5: Flow chart of the process for generating Morphological Variability (MV).

The runtime of measuring MV is dominated by the time taken to quantify time-warped morphology differences between consecutive beats. For a total of  $p$  beats in an ECG time-series of length less than  $n$ , the computational complexity of this step is  $O(pn^2)$ . While reducing the number of consecutive pairs of beats to be examined (i.e., reducing  $p$ ) offers one approach to reduce the overall runtime of MV, this approach is made challenging by multiple factors (e.g., poorer estimation of spectral energy, less data available to distinguish between persistent pathological variations and true noise, increased latency for real-time decision-making etc.). As a result, our efforts largely center on addressing the quadratic runtime of DTW (i.e., reducing the  $n^2$  term above). We achieve this by exploring the idea of adaptively downsampling the time-series inputs to DTW, and describe how the dynamic programming problem inherent in DTW can be evolved to handle adaptively downsampled signals.

#### 4.1 Adaptive Down-sampling (ADAP)

Our work builds upon the use of downsampling to improve the efficiency of the basic DTW algorithm. The general idea underlying this approach is to reduce the  $O(n^2)$  runtime of DTW by reducing the sizes of its inputs. Popular existing approaches such as aggregate approximation (PAA) [9] and FastDTW [19] achieve this by downsampling signals uniformly, i.e., by a constant factor across time. We note that while the use of downsampling improves the runtime and space efficiency of DTW, the decision to carry out this downsampling by a constant factor over time causes both

rapidly and slowly changing parts of a signal to be treated similarly. Downsampling in this case may be associated with the loss of important information.

We believe that this process can be improved by exploiting slowly changing parts of a signal by downsampling them at a higher rate than rapidly changing regions. In contrast to PAA and FastDTW, we therefore propose the idea of adaptive downsampling where the rate of reduction of time series varies according to the rate of changes taking place locally. This allows for the reduction of time series, while retaining sharp changes that would otherwise be smeared if downsampling were applied uniformly to the entire signal.

We achieve adaptive downsampling using trace segmentation [11]. While we describe this approach in more detail subsequently, the basic idea underlying trace segmentation is to divide the signal into regions with equal cumulative derivative activity. This places a higher number of boundaries for downsampling in regions that are rapidly changing (i.e., have higher cumulative derivative activity).

More formally, given a signal  $Q = q_1, \dots, q_n$  and a number of frames  $\theta$  to downsample this signal to, we first calculate the cumulative difference  $D_Q[k]$  for  $k = 2, \dots, n$  between each neighboring pair of samples:

$$D_Q[k] = \sum_{i=2}^k |q_i - q_{i-1}| \quad (1)$$

with  $D_Q[1] = 0$ . The sum of the total differences in  $Q$  is given by  $D_Q[n]$ . The cumulative difference in each adaptive downsampling bin is then set to  $d_Q = \frac{D_Q[n]}{\theta}$ . Using this, downsampling proceeds by finding the sample numbers  $t_i$  for  $i = 0, \dots, \theta$  such that for all values of  $i$  we have:

$$t_i = \min\{k \mid D_Q[k] \geq d_Q \cdot i\} \quad (2)$$

The corresponding amplitudes of  $Q$  at samples  $t_i$  are given by  $x_i = q_{t_i}$ . We can then use interpolation to approximate the fractional sample numbers  $\hat{t}_i$  where we would expect  $D_Q[\hat{t}_i] = d_Q \cdot i$ . For  $i = 0, \dots, \theta$  using the notation:

$$\beta_i = \frac{D_Q[t_i] - d_Q \cdot i}{D_Q[t_i] - D_Q[t_i - 1]} \quad (3)$$

we have:

$$\begin{aligned} \hat{t}_i &= t_i - \beta_i \\ \hat{x}_i &= q_{t_i} - \beta_i(q_{t_i} - q_{t_i-1}) \end{aligned} \quad (4)$$

The resulting adaptively downsampled representation of the original signal  $Q$  is given by two series corresponding to time and amplitude:

$$\begin{aligned} T^Q &= \hat{t}_0, \hat{t}_1, \dots, \hat{t}_i, \dots, \hat{t}_\theta \\ X^Q &= \hat{x}_0, \hat{x}_1, \dots, \hat{x}_i, \dots, \hat{x}_\theta \end{aligned} \quad (5)$$

This process can be carried out in time that is linear in the size of the input. Figure 6 presents the trace segmentation approach for downsampling graphically.

For ECG time-series, trace segmentation can preserve important information related to sharply changing parts of the signal (e.g., the QRS complex). This is illustrated in Figure 7. In contrast to PAA, a similar number of adaptively

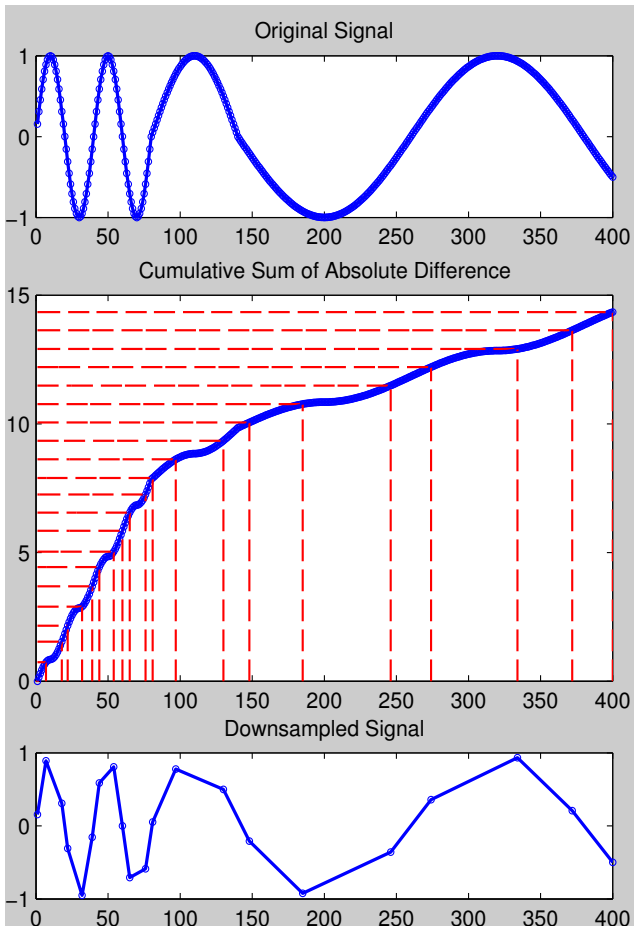


Figure 6: An illustration of the trace segmentation process.

downsampled segments provide a better characterization of notching within the R wave and also the sharpness of the S wave. While PAA achieves good results in a variety of real-world application domains, we believe the distinctions retained by adaptively downsampling are relevant to the specific goal here of measuring MV to predict death following heart attacks.

## 4.2 DTW with ADAP

DTW searches for the optimal alignment between two sequences in an efficient manner using dynamic programming. For uniformly downsampled signals, the dynamic programming process is essentially unchanged, although it is applied to reduced representations of the original signals. For adaptively downsampled signals, however, the cost of alignment cannot be calculated in a similarly simple manner from the Euclidean distance between the samples of the downsampled representations. Since the original signal is now divided into segments of variable lengths, this length information needs to be factored into consideration when deriving the distances for the DTW dynamic programming recurrence.

We represent two adaptively downsampled signals  $Q$  and  $C$  as comprised of segments  $s_q(1), \dots, s_q(\theta)$  and  $s_c(1), \dots, s_c(\theta)$  respectively, with  $\theta$  corresponding to the number of downsampled segments. The amplitude of each segment  $s_q(i)$  is

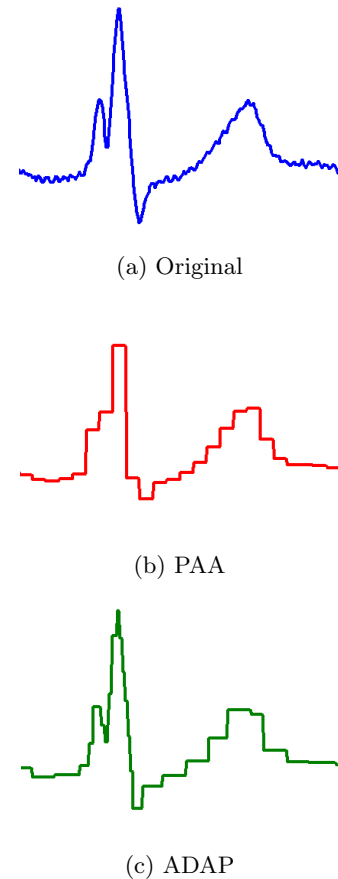
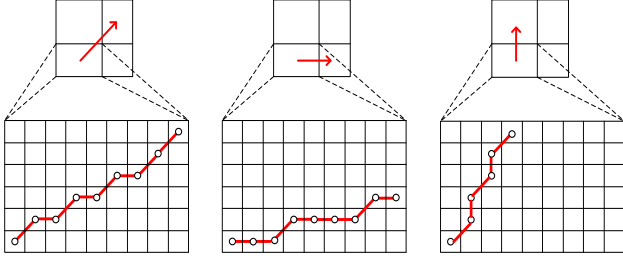


Figure 7: Adaptive downsampling of ECG signals.

represented by  $x_q(i)$  and the duration by  $l_q(i)$  (similar notation is used for the amplitude and duration of each segment  $s_c(i)$ ). Using this notation, we describe the process through which the dynamic programming of DTW can be modified to handle adaptively downsampled segments.

Figure 8 shows, from left to right, three separate possibilities when aligning adaptively downsampled segments. In each case, the alignments of the adaptively downsampled segments are illustrated at the top, and the alignments of the original signals are illustrated below. The leftmost subfigure shows the situation where the adaptively downsampled segments are diagonally aligned, i.e. segment  $s_q(i)$  is aligned with segment  $s_c(j)$ , while  $s_q(i+1)$  is aligned with segment  $s_c(j+1)$ . Intuitively, we expect the warping path between the samples comprising the segments  $s_q(i)$  and  $s_c(j)$  in the original signals to be close to the diagonal. Without solving for the optimal path of alignment between these original samples, we approximate the cost of alignment between the segments  $s_q(i)$  and  $s_c(j)$  as the product of  $d(s_q(i), s_c(j))$  (i.e., the Euclidean distance of  $x_q(i)$  and  $x_c(j)$ ) and  $\max(l_q(i), l_c(j))$  (i.e., an estimate for the length of a diagonal path). We adopt a similar approach for the subfigure shown in the middle of Figure 8. In this case, the adaptively downsampled segment  $s_q(i)$  is aligned with both  $s_c(j)$  and  $s_c(j+1)$ . Again, without solving for the optimal path of alignment between the original samples for these segments, we expect the path of alignment for the



**Figure 8: Illustration of derivation for adaptive recurrence equation. Left subfigure shows the diagonal case, the middle shows the horizontal case, while the right subfigure shows the vertical case.**

samples comprising  $s_q(i)$  and  $s_c(j)$  to be roughly horizontal. We therefore approximate the length of this path to be the product of  $d(s_q(i), s_c(j))$  and  $l_q(i)$ . The situation shown in the rightmost subfigure (i.e., a roughly vertical path of alignment for the samples comprising  $s_q(i)$  and  $s_c(j)$ ) is treated analogously.

We note that our approach of modifying the dynamic programming of DTW for use with adaptive downsampling approximates the path length in each case, and this approximation may not be optimal. However, this approach provides a simple way to augment the dynamic programming of DTW. In particular, in this setting, the recurrence relation for the cumulative path distance  $\gamma(i, j)$  till the adaptively downsampled segments  $i$  and  $j$  can be represented as:

$$\gamma(i, j) = \min \begin{cases} \gamma(i, j, \mathbf{d}) \\ \gamma(i, j, \mathbf{h}) \\ \gamma(i, j, \mathbf{v}) \end{cases}$$

where the cumulative path distance  $\gamma(i, j)$  depends on the direction in which the path proceeds next (i.e., diagonal  $\mathbf{d}$ , horizontal  $\mathbf{h}$ , or vertical  $\mathbf{v}$ ) and:

$$\gamma(i, j, \mathbf{d}) = d(s_q(i), s_c(j)) \max(l_q(i), l_c(j)) + \min \begin{cases} \gamma(i-1, j-1, \mathbf{d}) \\ \gamma(i-1, j, \mathbf{h}) \\ \gamma(i, j-1, \mathbf{v}) \end{cases}$$

$$\gamma(i, j, \mathbf{h}) = d(s_q(i), s_c(j)) l_q(i) + \min \begin{cases} \gamma(i-1, j-1, \mathbf{d}) \\ \gamma(i-1, j, \mathbf{h}) \\ \gamma(i, j-1, \mathbf{v}) \end{cases}$$

$$\gamma(i, j, \mathbf{v}) = d(s_q(i), s_c(j)) l_c(j) + \min \begin{cases} \gamma(i-1, j-1, \mathbf{d}) \\ \gamma(i-1, j, \mathbf{h}) \\ \gamma(i, j-1, \mathbf{v}) \end{cases}$$

### 4.3 Path Constraints

In Section 3 we described a modification to the basic DTW recurrence relation to find more biologically plausible alignments (i.e., the situation shown in Figure 4). In this case, we made use of the recurrence relation:

$$\gamma(i, j) = d(q_i, c_j) + \min \begin{cases} \gamma(i-1, j-1) \\ d(q_{i-1}, c_j) + \gamma(i-2, j-1) \\ d(q_{i-1}, c_j) + d(q_{i-2}, c_j) + \gamma(i-3, j-1) \\ d(q_i, c_{j-1}) + \gamma(i-1, j-2) \\ d(q_i, c_{j-1}) + d(q_i, c_{j-2}) + \gamma(i-1, j-3) \end{cases}$$

We adopt an analogous approach to constrain DTW with adaptive downsampling for more meaningful alignments. Since the original signal is divided into unequally sized segments, we note that the above recurrence would not be directly applicable. Instead of restricting valid paths to pass through no more than 3 consecutive horizontal or vertical steps, we therefore restrict the path to traverse through at most  $k$  steps such that no such implausible alignment would occur. In other words, a segment  $s_q(i)$  is only allowed to align with segments of  $s_c$  such that the total length of those  $k$  segments is no greater than three times the length of  $s_q(i)$ , which can be expressed as  $3 \cdot l_q(i) \leq \sum_{n=1}^k l_c(j-n)$ .

## 5. EVALUATION

We evaluated our research on ECG data from patients in the DISPERSE2-TIMI33 trial [5]. Patients in the DISPERSE2-TIMI33 trial were enrolled if they experienced ischemic symptoms at rest for a duration exceeding 10 minutes with either biochemical marker evidence of myocardial infarction (defined as Troponin-T, -I, or creatinine kinase-MB elevation greater than the local myocardial infarction decision limit) or ECG evidence of ischemia (defined as the presence of new or presumably new ST-segment depression  $\geq 0.05$  mV, transient ST-segment elevation  $\geq 0.1$  mV, or T-wave inversion  $\geq 0.1$  mV in 2 or more contiguous leads). As part of this study, continuous electrocardiographic (ECG) data were recorded for a median duration of 4 days. Three-lead Life-Card CF Holter monitors were placed within 48 hours of the initial event, and the data were sampled at 128 Hz. Patients were followed up for a period of 90 days for cardiovascular death. In our study, we used data from the first 24 hours of ECG recording during hospitalization to predict the risk of death following heart attacks. There were a total of 765 patients in the DISPERSE2-TIMI33 trial with available ECG signals used in this analyses, with 14 deaths during follow-up. On average, each 24 hour recording contained 103,180 instantaneous heart rate measurements.

We evaluated MV, MV measured with downsampling using PAA, and MV measured with adaptive downsampling for cardiovascular risk stratification in multiple ways. First, we measured the areas under the receiver operating characteristic curves (AUROCs), which reflects the ability of the different MV approaches to discriminate between patients who died during follow up and those that remained event free. The AUROC is widely used in medicine, and is generally considered the standard for evaluating risk stratification methods [1]. As part of this evaluation, we compared the AUROC values for the downsampled MV approaches to the basic MV algorithm without downsampling using the method proposed by DeLong et al. [6] to assess whether the changes are statistically significant. Second, we also assessed the MV models with downsampling relative to the basic MV algorithm without downsampling by measuring the integrated discrimination improvement (IDI) proposed

**Table 1: Univariate association of MV and other clinical variables with death following heart attacks.**

Parameter	Hazard Ratio	P-value
Age>65	3.72	0.024
Women	2.76	0.054
Smoker	0.53	0.225
Hypertension	6.66	0.067
Diabetes	2.77	0.049
Hyperlipidemia	0.66	0.422
Previous Heart Attack	1.94	0.210
Previous Angina	2.86	0.103
ST depression>0.5mm	2.69	0.091
MV	5.16	0.002

by Pencina et al. [18]. This was done by translating the MV values obtained through each approach into regression-based probabilistic risk estimates, and then measuring the difference between the mean predicted probabilities of events and non-events.

In addition to evaluating the predictive accuracy of MV measured through each approach, we also evaluated the runtime of the algorithms as the average time taken across ten runs to compute MV for all patients. These experiments were performed on a machine with quad-core Intel Xeon X3450 processors (2.67 GHz, 8MB Cache) and 8 GB RAM. The distance metrics were uniformly implemented in C++ on the Red Hat Enterprise Linux Server release 5.6 (Tikanga).

Finally, we also assessed how the relative ranking of patients between the different MV approaches changed with adaptive and non-adaptive downsampling. This metric was used to study how downsampling moves patients relative to each other while measuring MV. To measure this information, we computed the average absolute difference in the ranking of each patient by MV across different approaches.

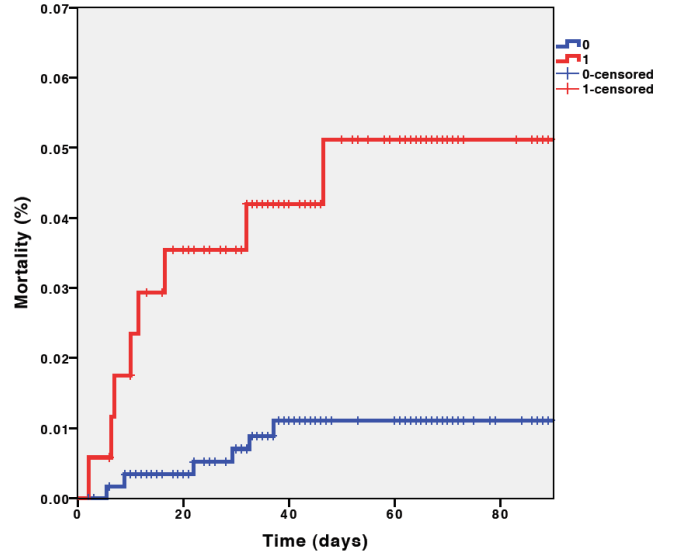
## 6. RESULTS

### 6.1 Association of MV and death

The basic MV algorithm achieved an AUROC of 0.75 for discriminating between high and low risk patients following heart attacks. When the MV predictions were dichotomized at a simple threshold ( $MV > 50$  vs.  $MV \leq 50$ ), patients with high MV were found to be at a significantly increased risk of death following heart attacks (Figure 9). For comparison, we show the relative increase in risk between patients with high and low MV, as well as the relative increases in risk for a variety of existing clinical variables in Table 1. In the DISPERSE2-TIMI33 dataset, MV identified a group of patients at a higher relative risk than any of these other metrics. These results are consistent with our earlier findings reported in the clinical literature [20].

### 6.2 Scaling up MV to Large Datasets

Table 2 compares the AUROC for the basic MV algorithm with the AUROCs obtained for MV measured with downsampling using PAA and MV measured with adaptive downsampling. For both the downsampling approaches, we experimented with downsampling the original heart beat signals to down to 30, 50 and 70 samples.



**Figure 9: Kaplan-Meier mortality curve for patients in high MV ( $MV > 50$ ; shown in red) and low MV ( $MV \leq 50$ ; shown in blue) groups. Patients with high MV were at a consistently elevated risk of death over the 90 day period following a heart attack.**

**Table 2: Comparison of AUROCs between DTW, PAA-DTW, and ADAP-DTW**

Methods	AUROC	P-value
DTW	0.748	Referent
PAA <sub>30</sub>	0.658	0.331
PAA <sub>50</sub>	0.669	0.345
PAA <sub>70</sub>	0.693	0.384
ADAP <sub>30</sub>	0.718	0.345
ADAP <sub>50</sub>	0.737	0.729
ADAP <sub>70</sub>	0.736	0.721

In general, downsampling the original signal led to a reduction in the discriminative ability of MV (although this difference was not significant at the 5% level given the sample size). In all of our experiments, however, MV with adaptive downsampling achieved a higher AUROC than downsampling with PAA for a similar factor of reduction. These results suggest that our use of adaptive downsampling retained more information that was relevant to the task of distinguishing between high and low risk patients than the use of PAA for this application.

We also assessed changes in the clinical utility of MV with downsampling by computing the IDI for each downsampled approaches relative to the basic MV algorithm. These results are presented in Table 3. In this case, the data from our experiments show that (consistent with the AUROC case) the use of downsampling with PAA led to a small decrease in performance. Conversely, the use of adaptive downsampling actually resulted in an increase in discriminative performance relative to the basic DTW algorithm as measured by the IDI. The differences for both downsampling with PAA and with adaptive downsampling relative to the basic DTW algorithm was not significant at the 5% level given the sample size.



**Table 3: IDI comparing DTW with PAA-DTW and ADAP-DTW**

Methods	IDI	P-value
PAA <sub>30</sub>	-0.009	0.164
PAA <sub>50</sub>	-0.009	0.178
PAA <sub>70</sub>	-0.007	0.209
ADAP <sub>30</sub>	0.017	0.160
ADAP <sub>50</sub>	0.014	0.204
ADAP <sub>70</sub>	0.020	0.160

**Table 4: Average change in patient ranks relative to the basic DTW algorithm (shown as percentages of the DISPERSE2-TIMI33 population).**

Methods	Change
PAA <sub>30</sub>	20.3%
PAA <sub>50</sub>	21.2%
PAA <sub>70</sub>	20.9%
ADAP <sub>30</sub>	18.3%
ADAP <sub>50</sub>	18.8%
ADAP <sub>70</sub>	18.4%

**Table 5: Timing of the different MV algorithms.**

Methods	Time (sec)
DTW	146,940
PAA <sub>30</sub>	5,663
PAA <sub>50</sub>	7,311
PAA <sub>70</sub>	9,207
ADAP <sub>30</sub>	7,831
ADAP <sub>50</sub>	13,292
ADAP <sub>70</sub>	20,782

The relative changes in ranks of patients between the basic DTW algorithm and the DTW approaches with downsampling are shown in Table 4. Consistent with the AUROC and IDI results, DTW with adaptive downsampling resulted in a smaller relative change in rank within the DISPERSE2-TIMI33 population relative to the basic DTW algorithm.

Finally, the timing results for the different methods are shown in Table 5. While downsampling reduced the runtime of the basic DTW algorithm substantially in each case, this reduction was greater for PAA than with the use of adaptive downsampling. This result can be attributed to the additional work that needs to be done to solve the modified dynamic programming problem for adaptively downsampled DTW. Comparing the PAA and adaptively downsampled approaches based on time rather than downsampling factor, however, still showed a higher level of performance with adaptive downsampling than with the use of PAA (e.g., for PAA<sub>70</sub> AUROC: 0.693, IDI: -0.007, average rank change: 20.9% and time: 9,207 vs. for ADAP<sub>30</sub> AUROC: 0.718, IDI: 0.017, average rank change: 18.3% and time: 7,831)

## 7. CONCLUSION

In this paper, we studied the question of mining noise-like variations in long-term ECG time-series to identify patients at an increased risk of death following heart attacks. To achieve this, we made use of a modified DTW- and Lomb-Scargle periodogram-based approach that first transforms ECG time-series into sequences of beat-to-beat time-aligned morphology differences, and then relates properties of these sequences to patient risk. While the ideas underlying this work derive from our earlier experiments reported in the clinical literature [21], we focused here on the question of how this basic approach can be scaled to very large ECG time-series databases and for use in embedded devices. As part of this work, we investigated a novel approach to address the quadratic runtime of DTW. In particular, we proposed the idea of adaptive downsampling, i.e., downsampling slowly changing parts of a signal much more than rapidly changing parts of the same signal, to reduce the size of the inputs presented to DTW while retaining a good representation of the original time-series being compared. We also described changes to the dynamic programming underlying DTW to exploit such adaptively downsampled signals, where the downsampled segments may be of varying lengths.

We evaluated our ideas on real-world data from patients within the DISPERSE2-TIMI33 trial. Our experiments suggest that measuring MV with adaptive downsampling substantially reduces runtime while providing similar performance to the basic MV algorithm that is not optimized for large volumes of data. In addition, the use of adaptive downsampling leads to more accurate performance than downsampling through the commonly used approach of PAA. We believe that these ideas may have additional merit in clinical applications where physiological time-warping is an issue during stratification.

## 8. ACKNOWLEDGEMENT

We thank Dr. Benjamin Scirica and the TIMI Study Group at the Brigham and Women’s Hospital for the data used in our experiments. We are also grateful to Profs. John Guttag and Collin Stultz at MIT for their input on improving this

work. This research was supported, in part, by the National Science Foundation and the American Heart Association.

## 9. REFERENCES

- [1] D. Altman. *Practical statistics for medical research*. Chapman & Hall/CRC, 1991.
- [2] S. Ben-Haim, B. Becker, Y. Edoute, M. Kochanovski, and O. Azaria. Beat-to-beat electrocardiographic morphology variation in healed myocardial infarction. *The American journal of cardiology*, 68(8):725–728, 1991.
- [3] D. Berndt and J. Clifford. Using dynamic time warping to find patterns in time series. In *AAAI94 workshop on knowledge discovery in databases*, pages 359–370, 1994.
- [4] F. BJ. Defibrillators are lifesaver, but risks give pause. *New York Times*, 2008.
- [5] C. Cannon, S. Husted, R. Harrington, B. Scirica, H. Emanuelsson, G. Peters, R. Storey, 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. *Journal of the American College of Cardiology*, 50(19):1844–1851, 2007.
- [6] E. DeLong, D. DeLong, and D. Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, pages 837–845, 1988.
- [7] N. El-Sherif, B. Scherlag, R. Lazzara, and R. Hope. Re-entrant ventricular arrhythmias in the late myocardial infarction period. *Circulation*, 55(5):686–702, 1977.
- [8] J. Goldberger, M. Cain, S. Hohnloser, A. Kadish, B. Knight, M. Lauer, B. Maron, R. Page, R. Passman, D. Siscovick, 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. *Circulation*, 118(14):1497–1518, 2008.
- [9] E. Keogh, K. Chakrabarti, M. Pazzani, and S. Mehrotra. Dimensionality reduction for fast similarity search in large time series databases. *Knowledge and Information Systems*, 3(3):263–286, 2001.
- [10] H. Krumholz, P. Douglas, L. Goldman, and C. Waksmonski. Clinical utility of transthoracic two-dimensional and doppler echocardiography. *Journal of the American College of Cardiology*, 24(1):125–131, 1994.
- [11] M. Kuhn, H. Tomaschewski, and H. Ney. Fast nonlinear time alignment for isolated word recognition. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'81.*, volume 6, pages 736–740. IEEE, 1981.
- [12] L. Lilly. *Pathophysiology of heart disease*. Lippincott, Williams & Wilkins, 2003.
- [13] L. Lilly. *Pathophysiology of Heart Disease:: A Collaborative Project of Medical Students and Faculty*. Lippincott Williams & Wilkins, 2010.
- [14] D. Lloyd-Jones, R. Adams, T. Brown, M. Carnethon, S. Dai, G. De Simone, T. Ferguson, E. Ford, K. Furie, C. Gillespie, et al. Heart disease and stroke statistics - 2010 update. *Circulation*, 121(7):e46–e215, 2010.
- [15] N. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and space science*, 39(2):447–462, 1976.
- [16] R. Myerburg. Sudden cardiac death: exploring the limits of our knowledge. *Journal of Cardiovascular Electrophysiology*, 12(3):369–381, 2001.
- [17] C. Myers, L. Rabiner, and A. Rosenberg. Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 28(6):623–635, 1980.
- [18] M. Pencina, R. D’Agostino Sr, R. D’Agostino Jr, and R. Vasan. Evaluating the added predictive ability of a new marker: from area under the roc curve to reclassification and beyond. *Statistics in medicine*, 27(2):157–172, 2008.
- [19] S. Salvador and P. Chan. Fastdtw: Toward accurate dynamic time warping in linear time and space. In *KDD Workshop on Mining Temporal and Sequential Data*, pages 70–80, 2004.
- [20] Z. Syed, B. Scirica, S. Mohanavelu, P. Sung, E. Michelson, C. Cannon, P. Stone, C. Stultz, and J. Gutttag. Relation of death within 90 days of non-st-elevation acute coronary syndromes to variability in electrocardiographic morphology. *The American Journal of Cardiology*, 103(3):307–311, 2009.
- [21] Z. Syed, B. Scirica, C. Stultz, and J. Gutttag. Computationally generated cardiac biomarkers for risk stratification after acute coronary syndrome. *Science Translational Medicine*, 4(131):131ra49, 2011.
- [22] Z. Syed, P. Sung, B. Scirica, D. Morrow, C. Stultz, and J. Gutttag. Spectral energy of eeg morphologic differences to predict death. *Cardiovascular Engineering*, 9(1):18–26, 2009.
- [23] WHO. The top 10 causes of death. <http://www.who.int/mediacentre/factsheets/fs310>, 2011.