Sex-Specific Computationally Generated Biomarkers for Cardiovascular Risk Stratification post Acute Coronary Syndrome

by

Alicia Chong Rodriguez

B.S Electronic & Computer Engineering
Instituto Tecnológico y de Estudios Superiores de Monterrey, 2009

Submitted to the MIT Integrated Design & Management and the Department of Electrical Engineering & Computer Science in Partial Fulfillment of the Requirements for the Degrees of Master of Science in Engineering and Management and Master of Science in Electrical Engineering and Computer Science

at the

Massachusetts Institute of Technology

June 2018

© 2018 Alicia Chong Rodríguez. All rights reserved.
The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now or known hereafter created.

Signature redacted

MIT Integrated Design and Management
Department of Electrical Engineering and Computer Science

May 25, 2018

Signature redacted

Collin M. Stultz, MD, PhD
Professor of Electrical Engineering and Computer Science
Institute for Medical Engineering and Science
Thesis Supervisor

Signature redacted

Leslie Kolodziejski, PhD
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Signature redacted

Matthew S. Kressy
Executive Director
Integrated Design and Management Program
Sex-Specific Computationally Generated Biomarkers for Cardiovascular Risk Stratification post Acute Coronary Syndrome

by

Alicia Chong Rodriguez

Submitted to the MIT Integrated Design and Management Program and the Department of Electrical Engineering and Computer Science on May 25, 2018, in partial fulfillment of the requirements for the degrees of Master of Science in Engineering and Management and Master of Science in Electrical Engineering and Computer Science

Abstract

Women who present with symptoms consistent with an Acute Coronary Syndrome (ACS) are often under-diagnosed and under-represented in clinical trials. Moreover, there are data to suggest that women with cardiovascular disease have worse outcomes, poorer prognoses, and higher mortality rates than men. Determining the risk of future adverse cardiovascular events for women who have previously suffered an ACS is therefore a problem of paramount importance in the field of cardiovascular medicine. The identification of high-risk patient subgroups typically begins with an evaluation of the patient’s history, physical exam, and the surface electrocardiogram (ECG). Indeed, the ECG plays a central role in the assessment and management of patients post ACS.

In this study, we develop and test a technique for automatically assessing the risk of death in women who presented with an ACS. The method combines both patient history and an automated analysis of the surface ECG to accurately quantify that patient’s future risk. The clustering of patients into subgroups, each having a different level of risk, is used to develop an algorithm to quantify the risk of new patients who present with an ACS. In this work, a comprehensive comparison between clustering female only data and traditional, female and male data is demonstrated as risk stratification methodologies for learning the significance or impact of our test and its inputs. The model trained on the entire population always performs worse for female population and the model trained only on female patients always provides a better performance for these patients. Comparing to existing risk scores, the female-specific model performs better.

Thesis Supervisor: Collin M. Stultz MD, PhD
Title: Professor of Electrical Engineering and Computer Science
Institute for Medical Engineering and Science
Acknowledgements

Being part of MIT has been a childhood dream come true. A goal I sometimes thought would be unattainable for so many diverse reasons. Things happened at the right moment for me and led me to the right people, whom without their support I would not have achieved what I consider a successful outcome and very decisive and exciting next steps in my life journey.

First, I would like to thank Professor Collin Stultz, my thesis supervisor, his understanding and support to my work is unparalleled and inspires me to work harder. For the last year and a half he has guided, mentored, taught me immensely. Being part of the Computational Biophysics Group has been a great experience, especially learning from Paul Myers, whom I am grateful for having received his support and has significantly contributed to this work.

I would also like to thank my IDM Director, Matthew Kressy, he created an opportunity that will mark me for the rest of my life. Thanks to Melissa Parrillo, Andy MacInnis and all my IDM friends.

My studies at MIT were supported by the Costa Rican Ministry of Science and Technology (MICITT) through a scholarship from the PINN program and the MIT Legatum Fellowship, I am very thankful for them enabling this opportunity.

To all team behind the Legatum Center, Georgina Campbell, Megan Mitchell, Julia Turnbull, Ali Diallo, Reinaldo Normand, and my class, I was honored when I became part of the Legatum Fellows and having your support has been wonderful, empowering and an experience that has strengthened me for my global (ad)venture.

I want to thank all the people whom we’ve learned from and that have supported our work in Women’s Heart Health and our venture Bloomer Health Tech, Dr. Nandita Scott, Catherine Ricciardi, Tatiana Urman, Pratik Shah, Ramesh Raskar, Ian Waitz, Anantha Chandakrassan, Roger Mark, Thomas Heldt, Ben Linville-Engler, Trish Cotter, Bill Aulet, Alicia Carelli, Nick Meyer, Elaine Chen, and everyone in the Trust Center, MIT Sandbox, MIT MISTI.

Monica, Aceil and Antonio, thank you for being by my side making Bloomer Health Tech a reality, I look forward to our growth and impactful journey.

Adrian, Meli, Mcow, Sol, Andrea, Majo, Rossy, Chabi, Lety and Joy B Thank you for love, support, friendship and for always pushing me forward, enabling me to reach my highest potential.

To my global sisters and brothers, your support, love and messages are really important to me. I admire each of you and you motivate me everyday.
To my Mom and Dad, thank you for the strong values you established in me around equality and human rights, it has led me to triggers, passion in life and never giving up to what I can accomplish. We have lived together through societal changes and you are my main role models, I look up to you and feel strong, happy and loved every day.
# Table of Contents

1 **Introduction** .................................................................................................................. 12  
1.1 Motivation .................................................................................................................. 12  
1.2 Objectives of Thesis Work ....................................................................................... 13  
1.3 Thesis Organization .................................................................................................. 14  

2 **Sex-Related Differences in the Cardiovascular System** ............................................ 15  
2.1 Cardiovascular Physiology ......................................................................................... 15  
2.2 Acute Coronary Syndromes ....................................................................................... 16  
2.2.1 ACS Risk Factors ................................................................................................. 17  
2.3 The role of ECG in diagnosis of ACS ........................................................................ 18  
2.3.1 ECG sex-specific differences ............................................................................... 20  
2.4 Patient Risk Stratification ......................................................................................... 21  

3 **Sex-Specific Computational Biomarkers Based on ST-Segment Morphology** for Patient Risk Stratification ...................................................................................... 23  
3.1 Introduction .................................................................................................................. 23  
3.2 Datasets ......................................................................................................................... 24  
3.3 Feature Extraction ....................................................................................................... 26  
3.3.1 ST-segment morphology ....................................................................................... 26  
3.3.2 Normalization ......................................................................................................... 28  
3.4 Modeling ........................................................................................................................ 28  
3.4.1 Patient Feature Sets ............................................................................................... 28  
3.4.2 Determining number of Clusters K ......................................................................... 30  
3.4.3 K-means Clustering Method .................................................................................. 30  
3.5 Results .......................................................................................................................... 32  
3.5.1 Results on calculating K ....................................................................................... 32  
3.5.2 Models ................................................................................................................... 32  
3.5.3 Female-Specific Results and Analysis ................................................................... 33  
3.5.4 Independent Holdout Test ....................................................................................... 35  

4 **Conclusions and Future Work** ................................................................................... 38  
4.1 Future Work ................................................................................................................ 38  

Appendices .......................................................................................................................... 41
List of Figures

Figure 1. Traditional and Emerging Risk Factors .......................................................... 18
Figure 2. A vector view of the standard 12 Lead ECG................................................... 19
Figure 3 Diagram of an idealized ECG wave, with each feature labeled .......................... 20
Figure 4 ST-segment feature extraction process ........................................................... 27
Figure 5 High-Risk Clusters on each dataset.................................................................. 33
Figure 6 Model 2: $KM_{2\alpha} + 7Hx + 2ST$ AUC comparisons........................................ 44
Figure 7 Example of AUC results to choose $\kappa$ on TM, RxTM and FO ............................ 45
List of Tables

Table 1. Population in the Traditional Mixed Dataset (TM) ............................................. 25
Table 2. Population in the Cohort-2 Dataset ................................................................... 26
Table 3. K value for each Model and Dataset ................................................................. 32
Table 4. AUCs for each model and dataset ..................................................................... 34
Table 5. Paired sampled t-test results 1 .......................................................................... 34
Table 6. Paired sampled t-test results 2 .......................................................................... 35
Table 7. Cohort-2 AUCs ................................................................................................. 36
Table 8. Cohort-2 Sensitivity Results ............................................................................. 36
Table 9. Cohort-2 Specificity Result .............................................................................. 37
Table 10. Smart ECG Bra ............................................................................................... 39
Table 11. Physiologic changes during pregnancy, delivery and post-partum ................. 41
Table 12. Population in the Reduced-Traditional Mixed Dataset (Reduced-TM) .......... 43
Table 13: Performance of Reduced-Traditional Model compared to TM and FO ....... 43
Chapter 1

1 Introduction

1.1 Motivation

Historically, women have been underrepresented in clinical trials, and as a result, data derived from a predominantly male population are commonly and perhaps inappropriately extrapolated for clinical use in women [1]. Therefore, 24 years ago, in 1993, the National Institutes of Health (NIH) of the United States mandated that women and minorities be included in any government-funded health research [2][3], a decision made due to the alarming lack of knowledge around women’s health and the rising number of deaths from cardiovascular diseases (CVDs), the leading cause of death worldwide [3]. Even though it has led to new discoveries, inequities still exist. More recently, studies have demonstrated inadequate compliance with the NIH guidelines [2].

There has been little improvement in either the enrollment of women in trials or the use of sex-specific analyses [1]. Women constitute fewer than 30% of clinical trials for cardiovascular diseases and only half of scientific publications include analyses of the results according to the participants sex [6]. Therefore, data used for clinical decision-making, drug development and medical device design has been primarily made from clinical trials of cohorts composed mainly of men.

According to the American Heart Association, 44 million women in the U.S have some form of cardiovascular disease, about 6.6 million have coronary heart disease (CHD), 2.7 million women are living today with a history of myocardial infarction (MI), and women under the age of 50 who have had MI have double the risk of dying compared to their male counterparts [2]. Accumulated knowledge of heart health in men is not necessarily applicable to women, making it more difficult to diagnose and treat women’s heart problems effectively.
Recognizing that men and women differ in the diagnosis and outcomes of cardiovascular diseases has led to research interest and the recognition that there are crucial gaps impacting negatively women with heart disease [4]. This clear need amongst the affected population for more tools, data, analyses and conception of devices designed for accelerating data collection on women and increasing recruitment, adherence, information and knowledge to mitigate these incremental and alarming gaps constitute the motivation behind this work. Aiming to contribute to women’s health research by generating sex-specific tools in the growing field of machine learning in healthcare

1.2 Objectives of Thesis Work

The research presented in this thesis aims to demonstrate that by developing sex-specific methods the ability to identify women at high-risk of adverse outcomes after an acute coronary syndrome (ACS) can be improved. A successful risk stratification metric would help ensure that women with symptoms consistent with an ACS are assigned therapies that are appropriate for their level of risk.

In summary, this thesis has three objectives:

- Develop and test a technique for automatically assessing the risk of death in female patient populations who presented with an ACS by using electrocardiogram and other relevant features.

- Investigate if significant diagnostic information contained in large quantities of female-specific data, such as electrocardiogram recordings, is lost when these data are combined with larger quantities of male data.

- Explore the unconscious biases that can arise when building models
1.3 Thesis Organization

- Chapter 2: introduces sex-specific differences in the cardiovascular system and prior research evidence in this field. Ranging from facts, statistics to recent emerging research. Basic physiological concepts are explained and contextualized for sex-specific physiology and pathophysiology such as acute coronary syndromes. Manifestation and consequences of differences are described. Tools to collect information from these physiological parameters and patient risk stratification methodologies are introduced.

- Chapter 3: describes the datasets and features considered that will be given to the model used to create the risk metric that quantifies the patient’s risk of death within a 1-year period. Introduces unsupervised learning methods and presents the computational biomarkers based on ST-segment morphology for patient risk stratification. It makes the comparison within the different datasets used and female-specific analysis.

- Finally, Chapter 4 concludes the work with a summary of contributions and items for future work
Chapter 2

2 Sex-Related Differences in the Cardiovascular System

2.1 Cardiovascular Physiology

Understanding the foundations of sex-based differences in the cardiovascular system is important to developing new approaches for prevention, diagnosis, and treatment. The human body is made up of trillions of cells that all work together for the maintenance of the entire organism [8]. Very relevant differences between males and females exist and go beyond the obvious, reproductive system [9]. Cellular and molecular differences influence the prevalence and severity of a broad range of diseases, disorders and conditions [9].

When comparing the structure and function of the female and male heart it is commonly the same state [11]. Physiological differences include sex hormones, are commonly referred as influential to cardiovascular diseases [9].

Cyclical variations in sex hormones influence during the menstrual cycle, pregnancy, post-partum period and menopause. These changes have traditionally been considered a challenge when experiments are designed due to how it affects making observations with such variability when in contrast, men are perceived as a “cleaner” more constant model for study of the cardiovascular function [12].

Some of the physiological changes that occur in women during this cyclical variations include: increase in blood volume and cardiac output, an increase in red blood cell mass, a reduction in systemic vascular resistance and shifts in blood flow between organ systems [12].
Less evident physiological sex differences in the cardiovascular system include: the female heart being on average about twenty-five per-cent smaller than the male heart, including smaller diameter of major blood vessels such as the coronary artery and less collateral arteries branching off the major epicardial coronary arteries [11]. When comparing a female to a male of the same age and ethnicity, the smaller heart translates into a lower Stroke Volume (SV), but it doesn’t translate to less Cardiac Output (CO) [10].

2.2 Acute Coronary Syndromes

Acute Coronary Syndrome refers to a spectrum of conditions that are primarily attributable to inadequate blood flow to the myocardium following acute plaque rupture or erosion leading to thrombus formation [15]. This spectrum of conditions includes: unstable angina (UA, partial/intermittent occlusion, no myocardial damage), non-ST-segment elevation myocardial infarction (NSTEMI, partial/intermittent occlusion, myocardial damage) and ST-segment elevation myocardial infarction (STEMI, complete occlusion, myocardial damage) [15].

Women, on average, have worse outcomes, more complications, higher mortality rates, and poorer recovery rates [16]. Internationally, women also have worse outcomes, in similar proportions to the US despite various differences in health care and culture in non-US countries [16]. Women present with ACS on average ~10 years later than men in all regions of the world [16][17]. Regardless of age, within 1 year of a first acute myocardial infarction (AMI), more women than men will die, and the median survival time after a first myocardial infarction (MI) is significantly lower for women [16].

The trend of declining cardiovascular events in the overall population has not been seen in younger women over the past decade [16][18]. From 2001 to 2010, women demonstrated either no change (in women 30–39 years of age) or a slight absolute increase (in women 40–49 years) in hospitalization rates for acute myocardial infarction
Furthermore, young women <55 years of age have longer lengths of stay, higher hospital readmission rates, and higher in-hospital mortality, as compared to young men [18].

Some groundbreaking discoveries about ACS in women may explain these sex-differences, such as those accomplished through The Women’s Ischemic Syndrome Evaluation, known as WISE study, the Women’s Health Initiative, known as WHI and PROSPECT trial (Providing Regional Observations to Study Predictors of Events in the Coronary Tree. For example, differences in plaque morphology. Young women in particular (pre-menopausal) have shown higher rates of plaque erosion and thrombus formation, which means plaques, composed of smooth muscle and proteoglycan rich matrix without a necrotic lipid laden [13]. This shows erosion leading to thrombus formation with the core of the plaque intact and thrombi form after coming in contact with smooth muscle. In contrast, men have higher rates of plaque explosion or rupture, which is the morphology without a necrotic core. Women had lower rates of plaque rupture, smaller necrotic cores and less calcium, but similar overall plaque burden [13].

These types of differences raises concerns for the female population that presented ACS and demonstrate the need of understanding risk assessment, comorbidities, symptoms, treatment, and outcomes from a sex-specific perspective to enable further progress in this field.

2.2.1 ACS Risk Factors

Evolving knowledge of sex-specific presentations, improved recognition of novel risk factors, and expanded understanding of the sex-specific pathophysiology of ACS have resulted in improved clinical outcomes in women [10]. Identified risk factors in women presenting with ACS differ from men [1][3][10][16][24] and include multiple comorbidities, age and race/ethnicity. Obesity, smoking, hypertension have been found
more frequently in women than men with ACS [16]. Comorbidities were more frequent in women < 55 years old and these differences were less pronounced as age increase [16].

Traditional Risk Factors in both men and women include genetic predisposiion, age (women generally will be older than men to have ACS), hypertension, smoking (higher risk when women are below 45 years), diabetes and obesity.

Additionally, there are several non-traditional or emerging cardiac risks that are unique to or predominant in women, including early menopause or menarche, gestational diabetes mellitus (DM), hypertension, preeclampsia an eclampsia during pregnancy and systematic inflammatory disorders [25] (See Fig 1).

![Image of risk factors]

**Figure 1. Traditional and Emerging Risk Factors [25]**

### 2.3 The role of ECG in diagnosis of ACS

As soon as ACS is suspected, patients should seek medical care. However, women on average, delay the period between the onset of symptoms and their entry into the healthcare system [3]. Failure to recognize symptoms of ACS can cause critical delays in treatment [3]. Once ACS is suspected, standard hospital tests are performed, including
the recording of the ECG. The electrocardiogram (ECG) has an important role in the assessment of cardiac rhythm and in the evaluation of patients with suspected myocardial ischemia. The ECG is a non-invasive measurement of the cardiac electrical activity. The electrical activity across the heart on each heartbeat can be observed with this measurement in a waveform known as PQRST. The electrical potential from the depolarization and repolarization of the myocardium (i.e. heart muscle) [26]. The electrical conduction signal of each heart beat originates from the sinoatrial node and travels through the atria to the atrial-ventricular node, then travels down to the ventricular septum and up walls of the ventricles.[26] This electrical activity enables the heart to maintain its pumping blood function through the electrical impulses that spread from one cell to the next, thus creating a cardiac cycle.

A standard clinical ECG consists of 12 “leads”, which are schematically shown below. Lead locations can be observed in Fig 2.

![Figure 2. A vector view of the standard 12 Lead ECG.](image)

Each vector will provide a view of the PQRST waveform, a detailed diagram on the features that can be extracted from this waveform is provided in Fig 3.
In particular the ST-segment is a key indicator of ACS and episodes through ST-segment elevation or depression, these features will be used to build the models in Chapter 2. They play an important role as a predictor for morbidity and mortality.

2.3.1 ECG sex-specific differences

Two parameters of early repolarization (i.e., the amplitude of the junction between the termination of the QRS complex and the beginning of the ST segment, known as J point and the angle between the ST-segment and the baseline) were the best discriminators between the ECGs of males and females [29][30]. Women, on average, have a longer corrected QT interval and are more susceptible to the effects of QT prolonging medications. Differences in the heart rate and QT interval on a surface ECG are among the most commonly reported findings between male and female ECGs [30]. In summary combined results of recent and earlier studies have established that a typical male ECG can be differentiated from a typical female ECG by the following five characteristics [29]:

1. Higher takeoff of the ST segment i.e., greater amplitude of the ST junction
2. Shorter period between the J-point and the onset of T wave
3. Steeper slope of the ST segment
4. Steeper ascent of the T wave
5. Higher T wave amplitude.
2.4 Patient Risk Stratification

Risk stratification is the process of separating patient populations into high-risk and low-risk groups. There are existing tools for estimating the risk of death in patients after they experience ACS that are commonly based on echocardiography and clinical risk scores (for example, the TIMI risk score) [21]. These tools identify patients at high-risk of death post acute coronary syndrome. The tool is beneficial for the adequate management of the disease, since patients at high risk of cardiovascular death can benefit from invasive procedures and therapies, such as cardiac catherization, percutaneous coronary intervention, or surgery – all of which have been shown to reduce risk in the appropriate high risk patient population. By contrast, patients at low risk may not benefit, as the risk of the procedure itself may be higher than their own risk of a subsequent adverse event [31].

Risk stratification may be helpful for the planning of early treatment strategy with percutaneous intervention or drugs [27]. Commonly used risk scores for risk stratification include other patient data as well and serve as a tool for identifying patients at a high risk of death. The Global Registry of Acute Coronary Events (GRACE) risk score [27] is the largest multinational hospital discharge risk score considered to accurately predict long-term mortality post acute coronary syndrome and the Thrombolysis in Myocardial Infarction (TIMI) risk score [27] is a simple prognostication scheme that categorizes a patient's risk of death and ischemic events and provides a basis for therapeutic decision making.

Prior comparisons on the value of current risk scores for risk stratification and prognosis such as GRACE and TIMI have been studied specifically in female patients with non-ST segment elevation acute coronary syndrome (NSTE-ACS) [27]. The results of these studies, after patients where followed for over a year show differences in risk stratification and prognoses, while demonstrating the importance of other variables such as age of women in the study.
Differences in risk stratification could be one of the reasons a large international registry found women with ACS were generally treated less aggressively, were less likely than men to undergo coronary angiography, to have revascularization, to receive antiplatelet, including glycoprotein IIb/IIIa inhibitors, statin and angiotensin-converting enzyme inhibitors therapies[27] Women who have a higher incidence of non-obstructive coronary disease are less likely to be treated with guideline-based therapies, and have more complications with reperfusion interventions compared to men [2].

Establishing the right risk level for women can guide treatment and reduce the rate of under-treatment. Additionally, due to advances in technology, there has been a rise in collecting data and computing abilities in the last decades, with this progress, the potential of automating a lot of processes that physicians have used for a long time is high. This includes ECG long-term recordings and processing methods of these large quantities of data. Combining clinical data obtained from large-scale clinical trials and other studies with statistical processing methods, including machine-learning techniques, enable automated and further exploration of large amounts of data. This could potentially enable predicting prognoses adequately in female patients with ACS through risk stratification.
Chapter 3

3. Sex-Specific Computational Biomarkers Based on ST-Segment Morphology for Patient Risk Stratification

3.1 Introduction

An unsupervised learning approach that utilizes k-means clustering was used to identify ACS patient subgroups of varying risks and evaluates the utility of those subgroups in patient risk stratification. Clustering algorithms seek to learn, from the properties of the data, an optimal division or discrete labeling of groups of points [19]. In other words, the goal for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data. Through K-means clustering, the expectation is to discover the inherent groupings in the sex-specific data.

The main idea consists of defining k centroids, one for each cluster. Initially the centroids are randomly placed. Then each point belonging to the dataset will be assigned to the nearest centroid. The distance is calculated using the following objective function, in this case a squared error function [19]:

The objective function is

\[ x' = \sum_{j=1}^{k} \sum_{i=1}^{n} || x_i^{(j)} - c_j ||^2 \quad \text{(eq. 3)} \]

where \( || x_i^{(j)} - c_j ||^2 \) is a chosen distance measure between a data point \( x_i^{(j)} \) and the cluster center \( c_j \) [19].

When all points have been assigned, the first step is complete and an initial cluster assignment is made to all patients in the dataset [19].
The new centroids are then calculated from the k clusters arising from the previous step [19]. After getting these k new centroids, a new assignment has to be done between the dataset points and the nearest new centroid. This process is repeated and as a result the k centroids change their location step by step until the centroid locations converge. In other words, the centroids do not move any more.

Selecting the number of cluster centers is an important part of this process, which will be explained further in this Chapter. Through clustering, the relationships between the patient features and risk levels are examined. If the patients naturally separate into clusters based on their level of risk, insight into the relationship between the patient features and the risk of dying can be obtained. A comparison of these clustering results on datasets with female only data and traditional mixed datasets will be made, expecting a better cluster assignment for female patients from the female-only data. More on the data that will be used is below.

3.2 Datasets

Cohort-1 has been used in previous works [20] [21] [22] [23] for evaluating several computational biomarkers. It consists of records for 6,354 patients from the MERLIN-TIMI 36 (Metabolic Efficiency with Ranolazine for Less Ischemia in Non-ST-Elevation Acute Coronary Syndrome—Thrombolysis in Myocardial Infarction 36) clinical trial [10]. Only 2291 of those are female (36%), within the guidelines ratio of female and males in clinical trials [2]. There are features that include demographics, health history, laboratory tests and medications that were collected for each patient. Continuous ECG monitoring, sampled at 128 Hz, was recorded in two-thirds of these patients for up to seven days. Outcomes such as cardiovascular death and the date of death within a year after enrollment were obtained through a patient follow-up.

The criteria used to derive these two datasets from Cohort-1 consists of information in patients medical records that patients can promptly input in a risk metric and are
Comorbidities/risk factors of acute coronary syndromes. The following 7 binary features: gender, history of diabetes, history of hypertension, current smoker, prior MI, prior angiography, and at least one day of continuous ECG recordings. Patients with fewer than 50 usable beats in that first day of ECG data or who were missing any of the seven patient features were not considered.

The patients who fit the criteria from Cohort-1 dataset were used to build the two data subsets used in this work for training and validation of the models.

In order to compare how well the models identify high-risk female patients and analyze if we are losing significant diagnostic information when using traditional methods that combine female and male data, two datasets were derived from the Cohort-1 dataset:

- A traditional mixed dataset (the entire Cohort-1 dataset, also referred to as TM)
- A female-only dataset (all female patients only, also referred to as FO)

<table>
<thead>
<tr>
<th>Table 1. Population in the Traditional Mixed Dataset (TM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
</tr>
<tr>
<td>No. (%)</td>
</tr>
<tr>
<td>Total No. (%)</td>
</tr>
<tr>
<td>Clinical</td>
</tr>
<tr>
<td>Age, avg ±SD (IQR), y</td>
</tr>
<tr>
<td>Diabetes (%)</td>
</tr>
<tr>
<td>Hypertension (%)</td>
</tr>
<tr>
<td>Current Smoker (%)</td>
</tr>
<tr>
<td>Prior MI (%)</td>
</tr>
<tr>
<td>Prior Angiography (%)</td>
</tr>
<tr>
<td>Outcomes</td>
</tr>
<tr>
<td>Death (%)</td>
</tr>
</tbody>
</table>
The Cohort-2 dataset consists of records of 990 patients from the TIMI-Disperse2 clinical trial [check]. The common features between Cohort-1 and Cohort-2 include the same 7 binary features: gender, history of diabetes, history of hypertension, current smoker, prior MI, prior angiography, and at least one day of continuous ECG recordings. Patients with fewer than 50 usable beats in that first day of ECG data or who where missing any of the seven patient features were not considered.

Table 2. Population in the Cohort-2 Dataset

<table>
<thead>
<tr>
<th></th>
<th>Disperse Data (n = 861)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
</tr>
<tr>
<td>Total No. (%)</td>
<td>861</td>
</tr>
<tr>
<td>Clinical</td>
<td></td>
</tr>
<tr>
<td>Age, avg ±SD (IQR), y</td>
<td>63 ±11 (54 to 72)</td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>479 (55.63)</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>598 (69.45)</td>
</tr>
<tr>
<td>Current Smoker(%)</td>
<td>205 (23.80)</td>
</tr>
<tr>
<td>Prior MI(%)</td>
<td>76 (8.82)</td>
</tr>
<tr>
<td>Prior Angiography(%)</td>
<td>548 (63.64)</td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
</tr>
<tr>
<td>Death (%)</td>
<td>14 (1.63)</td>
</tr>
</tbody>
</table>

3.3 Feature Extraction

3.3.1 ST-segment morphology

ECG recordings have information relevant to ACS, especially contained in the ST-segment. In order to represent and quantitatively model the changes in morphology of the ST-segment a mathematical method that was proposed recently to distinguish between ischemic and non-ischemic ST segment events in the ECG tracing is used [23].
The method consists of using orthogonal basis functions to represent the morphology of the ST-segment.

The ECG recording for each patient is first preprocessed to remove noise and identify usable beats. Then the ECG was first segmented to delineate the PQRST wave of each beat and then the ST-segment of each beat is identified and transformed into a set of Legendre polynomial coefficients, which provide relevant features of the morphology of the ST-segment. The first coefficient is a constant and represents the level with respect to the isoelectric line of the ST-segment, the second coefficient corresponds to a linear function and represents the slope of the ST-segment and higher order coefficients correspond to non-linear functions and represent higher order curvature of the ST-segment. Up to 16 coefficients were used. Figure 4 shows this process.

![Diagram of ST-segment feature extraction process]

Figure 4 ST-segment feature extraction process [22].
All patients in both datasets had at least one day of continuous ECG data, but the number of usable beats varied across patients, therefore the average across 24 hours was calculated for each coefficient C0-C15. Also, a fixed number of the first 50 usable beats were used.

3.3.2 Normalization

Normalization of these coefficients was done in the same way as prior work [23] by calculating the standard deviation for each of the 16 coefficients, C0-C15, across all beats for all patients within each of the three datasets and then dividing each coefficient by its corresponding standard deviation.

Another feature that was normalized was the age. The age was scaled as follows:

\[ x' = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  
(eq. 4)

where, \( x_{\text{min}} \) is the minimum age, \( x_{\text{max}} \) is the maximum age across all patients within each dataset.

The Female-Only Dataset had different age maximum and minimums as can be observed in Table 3 and these were used for the normalization of this dataset.

3.4 Modeling

3.4.1 Patient Feature Sets

Four K-means clustering models that incorporate the following sets of patient features were constructed with the Traditional mixed dataset and Female only dataset for comparison:
\( KM_{2\mu + 7Hx} \) is the first model it uses 9 features that include the information from the average of coefficients C0 and C1 (level and slope) within the first 24 hours of ECG monitoring and now seven features from the patient’s medical history and demographic data. That is, the seven features derived are, age, gender, diabetes, hypertension, current smoker, prior MI and prior angiography. Age was continuous-valued all other features were binary. This model will be referred as 1-TM for the traditional-mixed dataset and 1-FO for the female-only dataset.

\( KM_{2\mu + 7Hx + 2ST} \) is the second model uses 109 features that include the information from the average of coefficients C0 and C1 (level and slope) within the first 24 hours of ECG monitoring, the seven features from the patient’s medical history and demographic data, and C0 and C1 (level and slope) corresponding to the first 50 usable beats from the continuous ECG monitoring as 50 beats was shown in a previous study [22] to be enough to build a good predictive model. This model will be referred as 2-TM for the traditional-mixed dataset and 2-FO for the female-only dataset.

\( KM_{16\mu + 7Hx} \) is the third model uses 23 features that include the information from the average of coefficients C0-C15 (level, slope, curvatures) within the first 24 hours of ECG monitoring and now seven features from the patient’s medical history and demographic data. That is, the seven features derived are, age, gender, diabetes, hypertension, current smoker, prior MI and prior angiography. Age was continuous-valued and all other features were binary. This model will be referred as 3-TM for the traditional-mixed dataset and 3-FO for the female-only dataset.

\( KM_{16\mu + 7Hx + 16ST} \) is the fourth model and uses 823 features that include the information from the average of coefficients C0-C15 (level, slope, curvatures) within the first 24 hours, the seven features from the patient’s medical history and demographic data, and C0-C15 (level, slope, curvatures) corresponding to the first 50 usable beats from the
continuous ECG monitoring. This model will be referred as 4-TM for the traditional-mixed dataset and 4-FO for the female-only dataset.

All four models were applied to the two datasets, Traditional mixed dataset and Female only dataset using a bootstrapping procedure in which each dataset was randomly partitioned into a training set that consists of 80% of the data and a testing set that consists of 20% of the data. In splitting the data, ensuring that the training and testing sets had the same proportion of females and the same proportion of deaths within each gender was taken into account. This random partition was repeated 100 times in the 3 different datasets. Afterwards, for each partition of each of the three datasets, the six models were trained on each of their corresponding training sets (80% of the data) and testing sets (20% of the data). In other words, the Traditional mixed dataset that has a population size of \( n = 4395 \) was partitioned into a training set of size \( n = 3517 \) and a testing set of size \( n = 878 \), and this was done 100 times creating 100 different pairs of training and testing sets. The same was done for the Female only dataset that has a patient size of \( n = 1521 \) and was partitioned into training set of size \( n = 1217 \) and testing sets of size \( n = 304 \), also 100 times.

3.4.2 Determining number of Clusters K

To choose the number of clusters for each of the four models on each of the two datasets, many different approaches where tried. Finally the approach selected involves using only one of the bootstrap partitions created above and treating the number of clusters, \( K \), as an optimization parameter. That is, the models were trained with a fixed value of \( K \) and the AUC was evaluated on the testing set. \( K \) was swept from 1 up to 100, and the \( K \) with the highest AUC was selected.

3.4.3 K-means Clustering Method

After choosing the number of clusters with the method described above, K-means clustering was performed once for each of the training sets of the 100 bootstrap
partitions using random centroid initialization, and squared Euclidean distance as a distance metric.

The risk of each cluster was evaluated after training by calculating the percentage of patients in each cluster who died. In the Traditional mixed datasets and female only models the death percentages of 3.39% and 3.81% were used, respectively, to define high risk and low risk clusters. These percentages correspond to the death rates in each dataset. A cluster with a risk equal to or higher than its dataset’s death percentage was considered high risk, and clusters with risk lower than the death percentage were considered low risk.

Afterwards, to evaluate the utility of the identified clusters at patient risk stratification on the testing set, for each patient in the testing set, the cluster closest to them was calculated by using the squared Euclidean distance. Each patient was given a predicted risk score equal to 1 if they belonged to a high-risk cluster or 0 if they did not belong to a low risk cluster.

Finally, the AUC of the test set was calculated using the actual patient labels and their predicted label. AUC gives the probability that a high-risk patient will be assigned a higher predicted risk score by our model than a low-risk patient. This evaluation metric takes into account both the true positive rate and the false positive rate of each model. On the traditional-mixed dataset the AUC was calculated for all the patients in the test set and also the AUC_FEM was calculated for only the female patients in the test set for comparison purposes.

The ability to generalize to unseen data was tested through independent holdout datasets consisting of patients from Cohort-2. The entire two datasets derived from Cohort-1 were used to train the models. The AUC was calculated to measure performance, other metrics observed for all models are the specificity, sensitivity and
3.5 Results

8 models were developed on two datasets with four feature sets. Throughout this section a comparison of these models and their results are provided.

3.5.1 Results on calculating K

The value of K that is chosen for each model is the one that gives the highest AUC. Table 8 has the selected K's for each of the 8 models. You can find an example of the calculation process in the Appendix.

Table 3. K value for each Model and Dataset

<table>
<thead>
<tr>
<th>#</th>
<th>Features</th>
<th>Number of Clusters (K)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>TM</td>
<td>Female-Only</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>KM2μ + 7Hx</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>KM2μ + 7Hx + 2ST</td>
<td>11</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>KM16μ + 7Hx</td>
<td>7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>KM16μ + 7Hx + 16ST</td>
<td>61</td>
<td>58</td>
<td></td>
</tr>
</tbody>
</table>

K is different in each model, and there is no general pattern that can be observed. In the case of Model 1 (KM2μ + 7Hx) for example, both datasets have K = 4 whereas in the rest of the models the TM dataset has a slightly higher value of K compared to the female-only models.

3.5.2 Models

After performing K-means clustering to the entire datasets, TM and Female Only, the clusters with the highest risks percentage of deaths are marked as high-risk clusters. The number of high-risk cluster within each model also varies. Fig 6 shows examples, model
model 2 \((KM_{2\mu} + 7Hx \ 2ST)\) used to show the risk assignments for each cluster in for both TM and FO datasets and how many patients belong to each cluster.

![Figure 5](image_url)

**Figure 5**: Models 1(above) and Model 2(below) High-Risk Clusters on each dataset. Shows the risk of the identified clusters for all the models. Each bar represents a cluster, and the height of the bar represents the risk (percent of patients who have died) in each cluster. The red clusters are the high risk cluster (3.39% for TM and 3.81% for FO). Numbers represent the number of patients that belong to each cluster.

3.5.3 **Female-Specific Results and Analysis**

The results of the AUCs are reported as averages over the 100 testing sets for the 8 models. The AUC of the test set was calculated using the actual patient labels and their predicted label, which is determined by the risk of the cluster they belong to. If they belong to a high-risk cluster, their predicted label is 1, whereas if they belong to a low
risk cluster, their predicted label is zero. Also, AUC_FEM was calculated separately for the female patients in each test set for each dataset.

Table 4. AUCs for each model and dataset

<table>
<thead>
<tr>
<th>#</th>
<th>Features</th>
<th>TM AUC</th>
<th>TM AUC_FEM</th>
<th>Female-Only AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>KM2μ + 7Hx</td>
<td>0.6720</td>
<td>0.6009</td>
<td>0.6334</td>
</tr>
<tr>
<td>2</td>
<td>KM2μ + 7Hx + 2ST</td>
<td>0.6818</td>
<td>0.6361</td>
<td>0.6336</td>
</tr>
<tr>
<td>3</td>
<td>KM16μ + 7Hx</td>
<td>0.6455</td>
<td>0.6009</td>
<td>0.6219</td>
</tr>
<tr>
<td>4</td>
<td>KM16μ + 7Hx + 16ST</td>
<td>0.6099</td>
<td>0.5958</td>
<td>0.5985</td>
</tr>
</tbody>
</table>

For each of the four feature sets a comparison was made between the overall AUC for the TM dataset and the AUC_FEM of the female populations within that dataset to determine if the female population performs differently from the overall population. A second comparison was done between the AUC of the female only dataset and the AUC_FEM of the female populations in the TM dataset to determine if building a female-specific model with the same feature set improves the predictive performance for the female population.

By comparing the overall AUCs in the TM with the AUCs of their corresponding female populations (AUC_FEM) it can be seen that the overall AUC is always higher than its corresponding AUC_FEM. P-values were calculated using a paired sample t-test and p<0.01 for all comparisons, indicating that the differences in the AUCs are statistically significant.

Table 5 Paired sampled t-test results between performance of female test sets and all-population test sets in TM model

<table>
<thead>
<tr>
<th>#</th>
<th>Features</th>
<th>p-value difference in TM AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>KM2μ + 7Hx</td>
<td>(p&lt;0.001)</td>
</tr>
<tr>
<td>2</td>
<td>KM2μ + 7Hx + 2ST</td>
<td>(p&lt;0.001)</td>
</tr>
<tr>
<td>3</td>
<td>KM16μ + 7Hx</td>
<td>(p&lt;0.001)</td>
</tr>
<tr>
<td>4</td>
<td>KM16μ + 7Hx + 16ST</td>
<td>p = 0.0099</td>
</tr>
</tbody>
</table>
Next, by comparing the AUC_FEM in TM to the Female Only AUC in each feature set, the Female Only AUC is always higher except in Model 2 ($KM_{2\mu} + 7Hx + 2ST$), but the differences in the AUCs are only statistically significant with $p<0.05$ in Models 1 ($KM_{2\mu} + 7Hx$) and Model 3 ($KM_{16\mu} + 7Hx$).

Table 6 Paired sampled t-test results between performance of TM female test set and Female Only performance within each model

<table>
<thead>
<tr>
<th>#</th>
<th>Features</th>
<th>p-value difference in TM AUC_FEM znd FO AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>KM$2\mu + 7Hx$</td>
<td>$p = 0.038$</td>
</tr>
<tr>
<td>2</td>
<td>KM$2\mu + 7Hx + 2ST$</td>
<td>not stat significant</td>
</tr>
<tr>
<td>3</td>
<td>KM$16\mu + 7Hx$</td>
<td>$p = 0.0295$</td>
</tr>
<tr>
<td>4</td>
<td>KM$16\mu + 7Hx + 16ST$</td>
<td>not stat significant</td>
</tr>
</tbody>
</table>

Comparing Model 1 ($KM_{2\mu} + 7Hx$) and Model 3 ($KM_{16\mu} + 7Hx$) it can be seen that the feature sets with the demographics and comorbidities ($7Hx$) have AUC differences between the female only data set and the AUC Fem in the TM dataset are statistically significant. Adding these extra features increases the overall AUCs but only slightly impacts AUC FEM. This may imply that the comorbidities combined with the male population don't provide that much information on the female population. Whereas on the female only dataset has better performance than the AUC FEM.

It is interesting to observe that even though TM has around 65% more data, the Female model has better or similar performance to TM AUC FEM.

3.5.4 Independent Holdout Test

The generalizability of using cluster assignments to assess risk is measured by using an independent dataset. The female subset of Cohort-2 consists of 309 female patients. The assessment is done on the 8 models that were trained and validated using Cohort-1.
For all feature sets, the Female Only models performed better than the TM models except in 3-TM. Therefore, further analysis was performed to better understand these results on feature set 3. The results were analyzed using statistical measures such as sensitivity (also called true positive rate) and specificity (all called true negative rate) to understand how the class imbalance can affect the AUCs.

By measuring the sensitivity of the model, the proportion of patients that died and got correctly assigned to a high-risk cluster can be analyzed. From Table 8, model 3-TM assigned the majority of positive patients correctly to high-risk clusters; model 3-FO had fewer positive assignments.

### Table 8: Cohort-2 Sensitivity Results

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Dataset</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>KM16µ + 7Hx</td>
<td>TM</td>
<td>0.8750</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.6250</td>
</tr>
</tbody>
</table>

Having high sensitivity results puts models at risk of assigning too many patients to clusters of high-risk level. Therefore the specificity of the models was obtained to measure the proportion of patients that survived and were assigned to a low risk cluster.
Table 9 Cohort-2 Specificity Result

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Dataset</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>KM16μ + 7Hx</td>
<td>0.5372</td>
</tr>
<tr>
<td></td>
<td>TM</td>
<td>0.5282</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.7055</td>
</tr>
</tbody>
</table>

The specificity of the 3-TM model is 0.5282, indicating that it classifies nearly half of the negative examples incorrectly; the 3-FO model, by contrast classifies more than 70% of the negative examples correctly. Since 3-TM also has a high sensitivity, most of the patients are getting assigned to high risk clusters. In summary, even though the 3-TM model has a higher sensitivity and AUC than the 3-FO model, it incorrectly classifies nearly half of the patients as high risk. Therefore, even though the AUC of the 3-TM model is higher than the AUC of the 3-FO model, the 3-TM model is not necessarily a better predictor of female risk.

Table: Cohort-2 Accuracy

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Dataset</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>KM16μ + 7Hx</td>
<td>0.5282</td>
</tr>
<tr>
<td></td>
<td>TM</td>
<td>0.5282</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.7076</td>
</tr>
</tbody>
</table>

In the accuracy table, it is observed that 3-TM again performed poorly. Female-Only Models had better accuracy for these patients. Therefore, from what has been observed, Model 3-TM, even though has a high AUC it has low specificity and low accuracy.
4 Conclusions and Future Work

The aim of this work was to use unsupervised learning to risk stratify female ACS patients. By comparing the performance of the models built using a dataset of 100% female patients and a dataset of an entire patient population with conventional under-representation of females, only 34.6% and higher representation of males, 65.4%. Features sets included the ST-segment morphology from the electrocardiogram and comorbidities. Area Under the Curve (AUC) was used for female predictive performance and demonstrated the impact, for the female population, of creating sex-specific models as compared to traditional-mixed models.

The results obtained demonstrate 1) models trained on an entire population, combining males and females, always performs worse or very similar for their female populations and 2) Female only models perform better than models with entire population even though they contain significantly less patients. In the study population, the TIMI NSTE-ACS risk score has an AUC of 0.5832 for the female patients, therefore this can be used as a benchmark that the highest performing female-specific model from this study, with an AUC of 0.6336 performs higher for the female population than a widely used risk score.

Sex-specific analysis with emerging knowledge offers promise in determining risk of future cardiovascular death for women who have previously suffered an ACS.

4.1 Future Work

The most important future work is to gather more female-specific data. Not only are women underrepresented in clinical trials, additional data on female-specific factors
(such as number of pregnancies, menopause age, etc) and emerging risk factors (such as gestational diabetes mellitus, preeclampsia, etc) may be stronger features for better performing risk stratification of female patients with ACS, but they are not usually collected in clinical trials that have both male and female patients.

We are working on continuous, valuable physiological data acquisition novel technologies, such as the fabric-engineered Smart ECG Bra by Bloomer Tech that connects via Bluetooth to a mobile app and enable the undertaking of collecting these female clinical data which can be used to further understand individual and evolving risk of patients.

![Smart ECG Bra](image)

**Table 10:** Smart ECG Bra

Additionally, conducting research, collecting female-specific data and building sex-specific models can come with many barriers to progress, including, personal unconscious biases on methods and systematically making errors. One example of this is cleaning the data properly and ensuring that, even though it was collected for both, female and male patients, the sample is not influenced by the male-dataset, that is, to perform pre-processing and data cleanup with only the female datasets. Especially now, that machine learning and artificial intelligence are growing and have been used effectively in the healthcare field. A framework to avoid pitfalls and biases in building Female-specific models is also important to design in order to ensure quality, standards
and progress in this field.
Appendices

Appendix A

Additional Cardiovascular System Sex Differences

During pregnancy for example, stroke volume can increase by 30% during first and second trimester, heart rate increases by 15-20 beats per minute and the cardiac output increases by up to 50% whereas arterial blood pressure and systemic vascular resistance decrease. All of these changes, systemic increases and decreases, can even persist for up to 3 months post-partum. See Table 2.

Table 11. Physiologic changes during pregnancy, delivery and post-partum

<table>
<thead>
<tr>
<th>Cardiovascular Function</th>
<th>Pregnancy</th>
<th>Labor &amp; Delivery</th>
<th>Post-partum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke Volume (SV)</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Heart Rate (HR)</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Cardiac Output (CO)</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Blood Pressure (ABP)</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Vascular Resistance (TPR)</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
</tr>
</tbody>
</table>

This means that the male heart has the benefit of pumping blood more easily as can be observed in Equation 1 [28], where in order to make the equation equivalent, females will have a higher Heart Rate (HR), on average higher resting heart rates by an average of 3-5 beats/min. Heart rate is controlled by a small patch of specialized heart cells located in the right atrium and referred to as the sino-atrial node [11], which will be discussed further in section 2.3.

\[ HR \times SV = CO \]  \hspace{1cm} (Eq. 1)

Additionally, male blood carries 10% more oxygen that females [11]. Under conditions
of cardiovascular stress (i.e. exercise), men respond by increasing their vascular resistance (TPR) which increases their mean blood pressure, women on the other hand, increase their HR which in turn increases the CO that increases the mean blood pressure as observed in Equation 2[10].

\[ P_a = CO \times TPR \]  \hspace{1cm} (Eq. 2)

There are also key differentiators in heart rate variability (HRV), a measure of the beat-to-beat variation in heart rate, often used as a substitute measure the autonomic function because it depends on a balance between parasympathetic and sympathetic activity. HRV will gradually diminish as
Appendix B

A reduced-traditional mixed dataset (the same total number of patients in the female-only dataset but with the percentage of males and females in the traditional dataset) was created and tested to see performance in same size datasets. Here we show preliminary example of the results, more iterations and reduced-mixed datasets need to be tested before stating validity of these results.

Table 12. Population in the Reduced-Traditional Mixed Dataset (Reduced-TM)

<table>
<thead>
<tr>
<th>Data</th>
<th>No. (%)</th>
<th>Overall</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n = 1521)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total No. (%)</td>
<td>1521</td>
<td>995 (65.41)</td>
<td>526 (34.58)</td>
<td></td>
</tr>
<tr>
<td>Clinical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, avg ±SD (IQR), y</td>
<td>63 ±11 (55 to 71)</td>
<td>62 ±11 (54 to 70)</td>
<td>63 ±11 (58 to 74)</td>
<td></td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>509 (33.46)</td>
<td>295 (19.39)</td>
<td>214 (14.07)</td>
<td></td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>1115 (73.31)</td>
<td>682 (44.83)</td>
<td>433 (28.47)</td>
<td></td>
</tr>
<tr>
<td>Current Smoker(%)</td>
<td>401 (26.36)</td>
<td>318 (20.90)</td>
<td>83 (5.46)</td>
<td></td>
</tr>
<tr>
<td>Prior MI(%)</td>
<td>503 (33.07)</td>
<td>363 (23.86)</td>
<td>140 (9.20)</td>
<td></td>
</tr>
<tr>
<td>Prior Angiography(%)</td>
<td>513 (33.72)</td>
<td>366 (24.06)</td>
<td>147 (9.66)</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death (%)</td>
<td>52 (3.41)</td>
<td>32 (3.22)</td>
<td>20 (3.80)</td>
<td></td>
</tr>
</tbody>
</table>

Table 13: Performance of Reduced-Traditional Model compared to TM and FO
These results show that the Reduced-traditional mixed model always performs significantly worse for the female patients than the Traditional-mixed model and the Female only model. Fig 5 shows a plot on the best performing model and how AUC FEM is compares from both TM and FO models in relation to the reduced model.

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Dataset</th>
<th>K</th>
<th>AUC Test</th>
<th>AUC Fem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1    KM2μ + 7Hx</td>
<td>TM</td>
<td>4</td>
<td>0.6720</td>
<td>0.6009</td>
</tr>
<tr>
<td></td>
<td>RxTM</td>
<td>4</td>
<td>0.6632</td>
<td>0.5656</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>4</td>
<td>0.6334</td>
<td>0.6334</td>
</tr>
<tr>
<td>2    KM2μ + 7Hx + 2ST</td>
<td>TM</td>
<td>11</td>
<td>0.6818</td>
<td>0.6361</td>
</tr>
<tr>
<td></td>
<td>RxTM</td>
<td>4</td>
<td>0.6590</td>
<td>0.5479</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>5</td>
<td>0.6336</td>
<td>0.6336</td>
</tr>
<tr>
<td>3    KM16μ + 7Hx</td>
<td>TM</td>
<td>7</td>
<td>0.6455</td>
<td>0.6009</td>
</tr>
<tr>
<td></td>
<td>RxTM</td>
<td>5</td>
<td>0.6324</td>
<td>0.5485</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>6</td>
<td>0.6219</td>
<td>0.6219</td>
</tr>
<tr>
<td>4    KM16μ + 7Hx + 16ST</td>
<td>TM</td>
<td>61</td>
<td>0.6099</td>
<td>0.5958</td>
</tr>
<tr>
<td></td>
<td>RxTM</td>
<td>67</td>
<td>0.5929</td>
<td>0.5305</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>58</td>
<td>0.5985</td>
<td>0.5985</td>
</tr>
</tbody>
</table>

Figure 6: Model 2: $KM_{2μ + 7Hx + 2ST}$ AUC comparisons
Appendix C

Choosing the right K for the model is an important process that was explored extensively. Some of the methods used are shared in this appendix.

Method 1: Sweeping from 1 to 100 clusters and selecting the cluster with the best performance and overall had the best performance of the three methods described, making this the chosen method of this work.

Figure 6 shows an example of the results on each K selected for the $KM_2 + 7Hx$ models built with the TM, RxTM and FO datasets. By coincidence K=4 in all of them.

Figure 7. Example of AUC results to choose K on TM, RxTM and FO

Method 2: Using the average variance within each cluster or elbow method. To choose two factors were considered: the average variance across all clusters, and the size of the smallest cluster. Choosing a number of clusters that would minimize the average variance across clusters, while maximizing the size of the smallest cluster with a constraint stating that no cluster can have less than 10 patients Figure below shows an example of this process.
Bibliography


[5] Noel Bairey Merz, MD: Gender Differences in Atherosclerosis and Coronary Heart Disease: A perspective from the WISE Study; ACC; March 6, 2004.


[26] https://millar.com/research/applications/telemetry-biopotential


