UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Doctoral Thesis

Clinical Decision Support System for the Multiparametric Stratification of Atrial Fibrillation Patients in Critical Care

Author:
Alexander Lacki

Supervisors:
D. Antonio Martinez-Millana
D. Octavio Rivera-Romero

May 2024

Declaration of Authorship

I, Alexander Lacki, declare that this thesis titled, "CLINICAL DECISION SUPPORT SYSTEM FOR THE MULTIPARAMETRIC STRATIFICATION OF ATRIAL FIBRILLATION PATIENTS IN CRITICAL CARE" and the work presented in it are my own. I confirm that:

- This work was done wholly while in candidature for a research doctoral degree at this University.
- Where I have consulted the published work of others, this is always clearly attributed and cited.
- Where I have quoted from the work of others, the source is always given as footnote or in the Bibliography.
- I have acknowledged all main sources of help and financial support.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Date _____ Signature _____

"If it were not for the great variability among individuals, medicine might as well be a science, not an art."

William Osler

Universitat Politècnica de València

Abstract

Doctor of Philosophy on Technologies for Health and Wellbeing

Clinical Decision Support System for the Multiparametric Stratification of Atrial Fibrillation Patients in Critical Care

by Alexander Lacki

Atrial fibrillation (AF) is the most commonly encountered cardiac arrhythmia, affecting over 33 million patients in the world. It is often encountered in intensive care units, where it is associated with prolonged hospitalisation, increased healthcare costs, elevated risk of thromboembolism, and higher mortality.

AF has diverse causes and mechanisms, and is considered to be a heterogeneous disease. It may be caused by cardiac and non-cardiac comorbidities, such as endocrine, pulmonary, and metabolic disorders, genetics, and inflammation. The abundance of pathophysiological mechanisms associated with AF has led to the realization that AF patients are considerably heterogeneous. This heterogeneity among patient populations have previously been identified as an unaddressed impediment in epidemiological studies.

Guidelines for the treatment and management of AF exist for the general population but are not directly applicable to ICU populations due to different AF mechanisms, risks, and effectiveness of treatments. Further, strong evidence for optimal treatment strategies is missing, resulting in a lack of consensus among clinical decision-makers, and different treatment approaches across clinical institutions.

This doctoral thesis reports the process of developing a stratification method for AF patients in the critical care setting. Novel semi-supervised clustering algorithms are developed, benchmarked, and employed to identify AF phenotypes. Treatment effects of common antiarrhythmic drugs are compared among phenotypes, and a usability assessment is performed to identify the clinical applicability of the developed methods.

Universitat Politècnica de València

Resumen

Doctorado en Tecnologías para la Salud y el Bienestar

Sistema de apoyo a la toma de decisiones clínicas para la estratificación multiparamétrica de pacientes con fibrilación auricular en cuidados intensivos

por Alexander Lacki

La fibrilación auricular (FA) es la arritmia cardíaca más común y afecta a más de 33 millones de pacientes en el mundo. A menudo se encuentra en unidades de cuidados intensivos, donde se asocia con hospitalizaciones prolongadas, mayores costos de atención médica, riesgo elevado de tromboembolismo y mayor mortalidad.

La FA tiene diversas causas y mecanismos y se considera una enfermedad heterogénea. Puede ser causada por comorbilidades cardíacas y no cardíacas, como trastornos endocrinos, pulmonares y metabólicos, genética e inflamación. La abundancia de mecanismos fisiopatológicos asociados con la FA ha llevado a la comprensión de que los pacientes con FA son considerablemente heterogéneos. Esta heterogeneidad entre las poblaciones de pacientes se ha identificado previamente como un impedimento no abordado en los estudios epidemiológicos.

Existen pautas para el tratamiento y manejo de la FA para la población general, pero no son directamente aplicables a las poblaciones de la UCI debido a los diferentes mecanismos, riesgos y efectividad de los tratamientos de la FA. Además, falta evidencia sólida sobre estrategias de tratamiento óptimas, lo que resulta en una falta de consenso entre los tomadores de decisiones clínicas y diferentes enfoques de tratamiento en las instituciones clínicas.

Esta tesis doctoral informa el proceso de desarrollo de un método de estratificación para pacientes con FA en el entorno de cuidados críticos. Se desarrollan, comparan y emplean nuevos algoritmos de agrupamiento semisupervisados para identificar fenotipos de FA. Se comparan los efectos del tratamiento de fármacos antiarrítmicos comunes entre fenotipos y se realiza una evaluación de usabilidad para identificar la aplicabilidad clínica de los métodos desarrollados.

Universitat Politècnica de València

Resum

Doctorat en Tecnologies per a la Salut i el Benestar

Sistema de suport a la presa de decisions clíniques per a l'estratificació multiparamètrica de pacients amb fibril·lació auricular en cures intensives

por Alexander Lacki

La fibril·lació auricular (FA) és l'arítmia cardíaca més comú i afecta més de 33 milions de pacients al món. Sovint es troba en unitats de cures intensives, on s'associa amb hospitalitzacions prolongades, majors costos d'atenció mèdica, risc elevat de tromboembolisme i més mortalitat.

La FA té diverses causes i mecanismes i es considera una malaltia heterogènia. Pot ser causada per comorbiditats cardíaques i no cardíaques, com ara trastorns endocrins, pulmonars i metabòlics, genètica i inflamació. L'abundància de mecanismes fisiopatològics associats a la FA ha portat a la comprensió que els pacients amb FA són considerablement heterogenis. Aquesta heterogeneïtat entre les poblacions de pacients s'ha identificat prèviament com un impediment no abordat als estudis epidemiològics.

Hi ha pautes per al tractament i maneig de la FA per a la població general, però no són directament aplicables a les poblacions de la UCI a causa dels diferents mecanismes, riscos i efectivitat dels tractaments de la FA. A més, manca evidència sòlida sobre estratègies de tractament òptimes, la qual cosa resulta en una manca de consens entre els prenedors de decisions clíniques i diferents enfocaments de tractament a les institucions clíniques.

Aquesta tesi doctoral informa el procés de desenvolupament d'un mètode d'estratificació per a pacients amb FA a l'entorn de cures crítiques. Es desenvolupen, comparen i fan servir nous algorismes d'agrupament semisupervisats per identificar fenotips de FA. Es comparen els efectes del tractament de fàrmacs antiarítmics comuns entre fenotips i es fa una avaluació d'usabilitat per identificar l'aplicabilitat clínica dels mètodes desenvolupats.

Acknowledgements

This work has only been possible with the support, mentorship, and guidance of Dr. Antonio Martínez Millana and Dr. Octavio Rivera-Romero, who I want to thank first.

Secondly, I would like to thank Dr. Ursula Ravens and Dr. Dirk Westermann for their support in improving my understanding of AF mechanisms, and providing a clinical perspective.

Further, I would like to acknowledge colleagues and co-supervisors who have helped and supported me during these last years and helped me shape this work for the better: Dr. Marilù Casini, Dr. Albert Dasí I Martinez, and Dr. Diego Boscá-Tomás together with the VeraTech team. Your support in its many different ways has been tremendous.

Finally, I would like to thank my parents and my grilfriend, Monika, for their love and support.

Contents

List of Figures

List of Tables

List of Abbreviations

AF	Atrial fibrillation
AP	Action potential
APD	Action potential duration
BB	Beta blocker
BMI	Body mass index
BP	Blood pressure
BPM	Beats per minute
BSA	Body surface area
CCB	Calcium channel blocker
CDSS	Clinical decision support system
COPD	Chronic obstructive pulmonary disease
DAD	Delayed afterdepolarization
EAD	Early afterdepolarization
ECG	Electrocardiogram
ECM	Extracellular matrix
eGFR	Estimated glomerular filtration rate
EHR	Electronic health Record
ERP	Effective refractory period
GDF-15	Growth differentiation factor 15
HAC	Hierarchical agglomerative clustering
HFpEF	Heart failure with preserved ejection fraction
HFrEF	Heart failure with reduced ejection fraction
LA	Left atrium
LAA	Left atrial appendage

MgS	Magnesium sulphate
MI	Myocardial infarction
OSA	Obstructive sleep apnoe
PCB	Potassium channel blocker
PVI	Pulmonary vein isolation
RCT	Randomized controlled trial
S-HAC	Survival hierarchical agglomerative clustering
SR	Sinus rhythm
TAM	Technology acceptance model
TTF	Task technology fit

Chapter 1

Introduction

1.1 Motivation and Rationale

Atrial fibrillation (AF) is the most prevalent cardiac arrhythmia affecting more than 33 million patients globally [175]. While AF currently consumes in excess of 13.5 billion euros in European healthcare systems [27], its prevalence is expected to further increase due to population aging, especially in rapidly developing countries such as Brazil, China, and India [175]. It is commonly encountered in critically ill patients, with incidences between 4.5% and 15% in intensive care units (ICUs) [33]. It is associated with increased healthcare costs, prolonged hospitalization, and increased risk of stroke and mortality [34, 195].

AF is a heterogeneous disease with a variety of risk factors and mechanisms. Even though significant progress has been made in the elucidation of these pathological processes in recent years, the primary realization that has arrived is that the AF population is highly heterogeneous with varying pathological mechanisms being present in different patients [226]. Such heterogeneity within patient populations has been recognized as a limitation of randomized controlled trials, since population sub-groups may respond differently to administered treatments [97].

While AF may be managed using various treatment strategies, therapies carry risks, and have undesirable side effects such as an increased risk of bleeding events or increased risk of death [64]. The treatment and management of AF therefore requires careful balancing of several objectives to maximize treatment utility while minimizing the associated risks. Nonetheless, guidelines for decision-making in the treatment of AF in critical care environments are lacking, and strong evidence for treatment selection is missing [53]. This results in a lack of consensus among critical care physicians, reliance on decision-making tools developed for community cohorts, and vastly different treatment approaches in different institutions [225].

This doctoral thesis targets these shortcomings and presents the development of a clinical decision support system for the identification of the best treatment option in AF patients in critical care environments. Biomarkers involved in the pathophysiological processes underlying AF are identified and captured from electronic health records (EHRs). Methods are developed to capture subphenotypes of the heterogeneous patient population and their characteristics and treatment responses are described. Finally, a decision-making algorithm is developed to predict optimal treatment selection, and a technology acceptance study is performed to evaluate the utility of the developed models in clinical practice.

1.2 Cardiac Anatomy and Electrophysiology

This section provides an overview of cardiac anatomy and electrophysiology. It is split into three sub-sections: a description of the heart, an overview of cardiac tissue properties and its constituents, and an introduction to cardiomyocyte function.

1.2.1 The Heart

Anatomy

The heart is a muscular organ which enables the circulation of blood through all blood vessels. Oxygenated blood is pumped throughout the body, while unoxygenated blood is transported to the lungs [76]. The pathway taken by the pumped blood is visualized in figure 1.1. Venous, unoxygenated blood enters the right atrium through the superior and inferior venae cavae. Upon atrial contraction, the blood passes through the tricuspid valve entering the right ventricle from where it is pumped through the pulmonary valve into the pulmonary artery, proceeding to the lungs. Oxygenated blood returns from the lungs through the pulmonary veins, entering the left atrium, from where it passes through the mitral valve to enter the left ventricle following atrial contraction. Ventricular contraction further pumps it through the aortic valve and into the aorta, from which the blood disperses throughout the body [93].

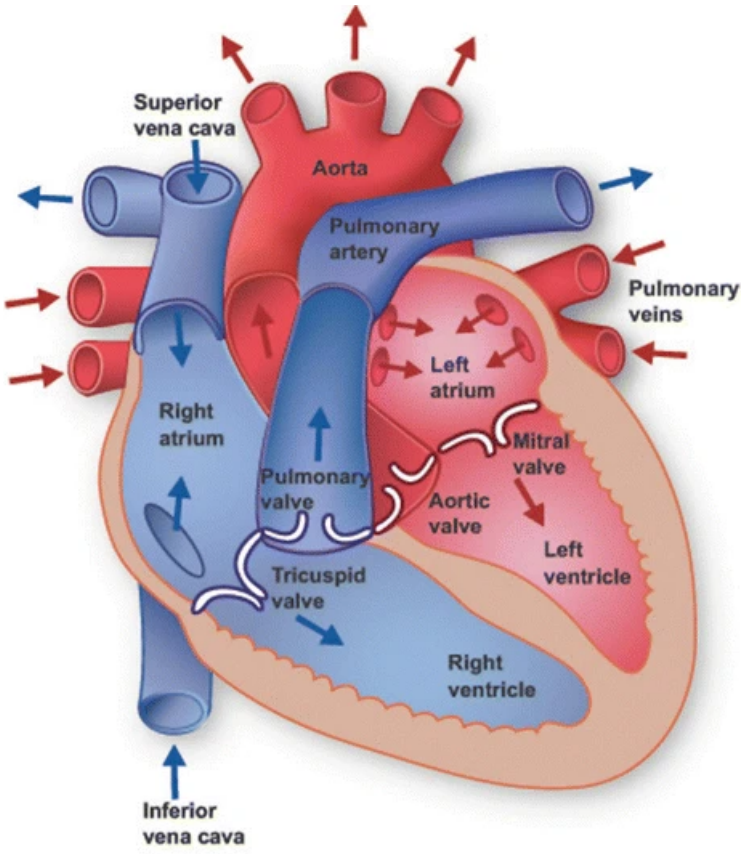

Figure 1.1: Anatomy of the heart [69]

While subtle anatomical differences between the left and right halves of the heart exist, the general operating principles remain fundamentally similar. Each of the four chambers, the atrial and ventricular lumen, are surrounded by muscles that contract to enable barometric pressure differences. These pressure differences cause passive opening and closing of cardiac valves and result in unidirectional blood flow. The efficient operation of the heart is provided by the synchronized contraction of the heart muscle, a sufficient contractile force, as well as a normal opening and closing of the heart valves [93].

Electrophysiology

The synchronous contraction of the heart is governed by the cardiac conduction system, which is portrayed in figure 1.2. The initial signal commences in the sinus node, the primary pacemaker of the cardiac conduction system, which is located in the roof of the right atrium. The signal propagates through the Bachmann's bundle as well as the left atrial tissue leading to excitation and synchronous contraction of both atria. Following complete atrial excitation, the atrioventricular node introduces a delay in propagation, which provides sufficient time for the contracting atria to displace blood into the ventricles. Following this delay, the atrioventricular node induces an impulse into the bundle of His, leading to a propagation of signal into the bundle branches, the Purkinje fibers, and the entire ventricular myocardium leading to a contraction of the ventricles [76].

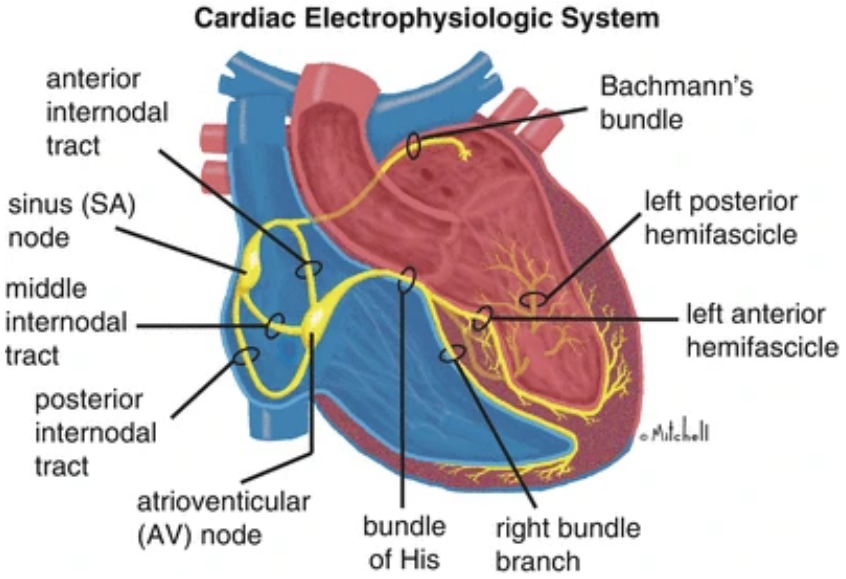

Figure 1.2: Cardiac conduction system [147]

1.2.2 Cardiac Tissue

Cardiac tissue has a complex composition with a multi-cellular structure. It consists of its contractile component, the cardiomyocyte, endothelial cells, fibroblasts, and the extracellular matrix. A schematic of cardiac tissue is presented in figure 1.3.

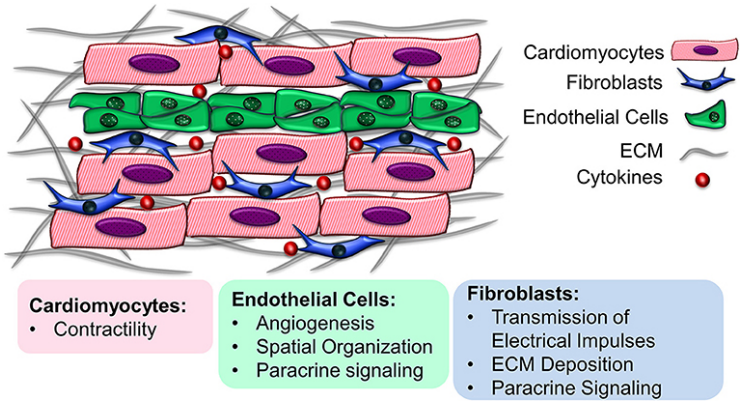

Figure 1.3: Composition of myocardial tissue [236]

The muscular component of the myocardium consists of cardiomyocytes, that are responsible for the rhythmic contraction and relaxation of the heart. Cardiomyocytes are connected through gap junctions which facilitate communication between neighboring cells allowing for the propagation of signals through the myocardium. The function of cardiomyocytes is covered in section 1.2.3.

Endothelial cells line blood vessels including capillaries. They regulate blood flow by controlling vasoconstriction and vasodilation following an activation by various receptors. Further, they are responsible for facilitating an exchange of substances between blood and myocardial tissue. These cells are of crucial importance in maintaining homeostasis by controlling the balance of a variety of cellular agonists and antagonists such as inflammatory factors, procoagulants and anticoagulants [57].

The various cells in the myocardium communicate via paracrine signaling through cytokines, which are small proteins produced and released in response to a variety of physiological and pathological processes. They are secreted by fibroblasts, endothelial cells, as well as cardiomyocytes. While a plethora

of cytokines exists, and significant cross-talk between different cytokines has been reported, it is well known that cytokines affect the autonomic innervation, which is a major regulator of heart function [121]. Among others, cytokines are released in response to inflammation and are involved in cytokine induced apoptosis. Cytokines have been proposed not only as prognostic markers but also as potential therapeutic targets in heart disease [54].

The constituents of the myocardium are connected and supported by the extracellular matrix (ECM). The ECM is composed of different types of connective tissue such as collagens and glycoproteins which form fibrils. These fibrils provide the myocardium with structural strength and elasticity, and are responsible for the transmission of the contractile force produced by cardiomyocytes. The ECM is an evolving entity which changes its composition in accordance with mechanical stress and hormonal factors [138].

Fibroblasts are cells that secrete collagen proteins, and thereby maintain the extracellular matrix. Due to their ability to secrete ECM proteins, excessive fibroblasts activation can lead to the development of fibrosis separating adjacent cardiomyocytes and reduce conduction velocities [62, 210]. In-vitro models have demonstrated the existence of electric coupling between fibroblasts and cardiomyocytes, its implication remains, however, poorly understood [26, 210]. Fibroblasts have been shown to react to mechanical stimuli, adapting their protein expression, and perform paracrine signaling to neighboring cells, affecting the function of cardiomyocytes [104].

1.2.3 Cardiomyocytes

On the smallest scale, heart function is controlled by electrical excitation of a cardiomyocyte, the heart's muscle cell, that triggers its contraction. Cardiomyocytes are electrically excitable because their cell membranes contain ion channels that are selective for various cations. The cardiomyocyte exhibits an inside negative resting membrane potential because the cell membrane is selectively permeable for K^+ ions but almost impermeable to Na^+ ions. Since at rest the inward rectifier K^+ channels are permanently open, the outwardly directed concentration gradient for K^+ is balanced by the inwardly directed electrostatic attraction of their positive charges to the negative inside of the cell. The Na+ pump [(Na^+,K^+)-ATPase] helps to maintain the concentration gradients by utilizing the energy of ATP to transport 3 Na^+ for 2 K^+ against their concentration gradients. During an action potential (AP) various ion channels will open and close in a voltage- and time-dependent manner and produce the characteristic cardiac shape of the AP depicted in figure 1.4.

The AP of cardiomyocytes is divided into five phases [178]:

Phase 0 - Depolarization of the cardiomyocyte through an activation of the fast inward Na^+ current resulting in a sharp increase of transmembrane voltage from the resting potential of -90 mV to 30mV.

Phase 1 - Early repolarization due to an inactivation of the inward Na^+ current, and transient outflow of K^+ ions resulting in a short drop in potential across the cardiomyocyte's cell membrane.

Phase 2 - Plateau phase of the action potential. Depolarizing Ca^{2+} flows into the cell via L-type Ca^{2+} channels, while repolarizing K^+ currents (I_{Kr} and I_{Ks}) flow out of the cell through rapidly activating Kv11.1 (hERG) and slowly activate Kv7.1 (KvLQT1) channels, respectively.

Phase 3 - Final repolarization of the cardiomyocyte occurs as when outward current prevails over inward current as L-type Ca^{2+} channels close, while K^+ channels remain open.

Phase 4 - Resting potential is reached again and maintained as explained above. The Na^+/Ca^{2+} exchanger (NCX) is the major actor which removes Ca^{2+}, that entered the cell during the AP, from the cell. The NCX utilizes the energy of the Na^+ concentration gradient for Ca extrusion from the cell. Because of the transport ratio of 3 Na^+ : 1 Ca^{2+} the NCX contributes some depolarizing current to the resting membrane potential.

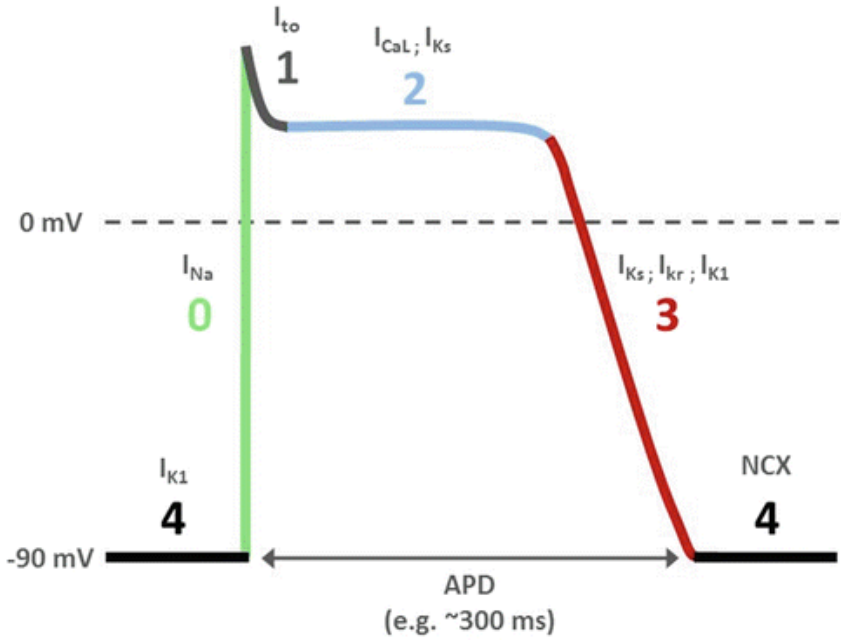

Figure 1.4: Cardiac action potential [178]

The shape and duration of the AP can vary depending on the expression of ion channels, which has been shown to be heterogeneous within the heart. Not only do atrial cardiomyocytes have shorter action potential durations (APDs) than ventricular cardiomyocytes, but intra-atrial APD variability has also been observed, with APDs in the left atrium being shorter than in the right atrium [157, 202].

The electrical stimulation is the prerequisite for the contraction of a cardiomyocyte. The sequence of events is referred to as *excitation-contraction coupling*, and the involved steps are portrayed in figure 1.5. Force development is primarily controlled by the interaction between intracellular Ca^{2+} concentration and the myofilaments. The Ca^{2+} entering the cell during the AP plateau triggers the release of further Ca^{2+} from the sarcoplasmic reticulum, the primary Ca^{2+} storage site of the cardiomyocyte [96].

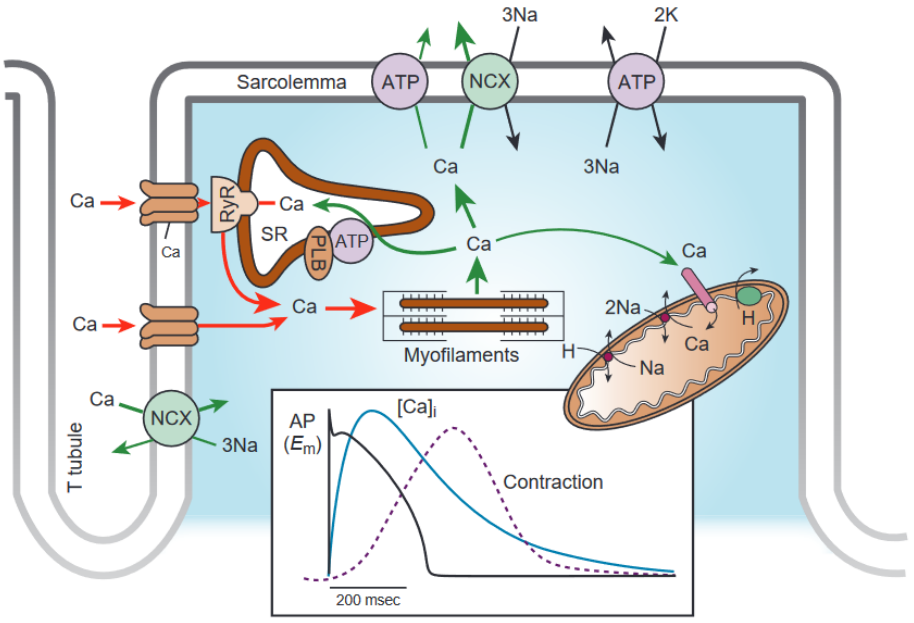

Figure 1.5: Excitation-contraction coupling [96]

The release of Ca^{2+} from the sarcoplasmic reticulum provides a sufficiently high intracellular Ca^{2+} concentration, that free Ca^{2+} binds to troponin C, resulting in a modification of the troponin-tropomyosin complex. This modification enables ATP hydrolysis supplying energy to the actin-myosin complex, and results in sliding of the myosin 'heads' over actin filaments in the force generating step. The simultaneous activation of all myofilaments within the cardiomyocyte causes the contraction of the entire cell [96].

1.3 Atrial Fibrillation

AF is a supraventricular tachyarrhythmia characterized by uncoordinated atrial activation resulting in ineffective atrial contractions. It presents with an irregular, and often rapid, ventricular response. The fibrillatory motion of the atria translates into hemodynamic incompetence, which results in blood stasis in the atria, increasing the risk of blood clot formation. AF patients, when untreated, are therefore at a five-fold increased risk of ischemic stroke [89].

AF is the most common sustained cardiac arrhythmia in adults, and poses a large burden on patients and healthcare systems [89]. It is estimated that between 2 and 4% of adults suffer from AF, and this number is expected to more than double in the upcoming years due to population aging [33]. Beyond a reduction in the quality of life, AF is associated with a variety of complications, such as myocardial infarction, heart failure, dementia, chronic kidney disease, and stroke. The associated complications of AF result in a 2-fold increased risk of death in the population of AF patients [19, 177, 33].

1.3.1 Pathophysiology

AF is not only associated with many co-occurring cardiac diseases, but other factors such as pulmonary, metabolic, endocrine diseases or genetic aberrations may also contribute [24, 106]. The mechanisms driving AF are not fully understood, but appear to be complex, of multifactorial nature, and heterogeneous across the patient populations [96]. It is generally agreed that the manifestation of AF involves two co-occurring factors: a trigger mechanism, which enables AF to commence, as well as a maintenance mechanism, which allows AF to perpetuate. Once maintained for prolonged periods, AF causes changes in atrial substrate structure and function, further promoting AF, which is referred to as remodeling [158].

Triggers of Atrial Fibrillation

Triggers that initiate AF are diverse, and include stimulation of the autonomic nervous system, atrial stretch, bradycardia, and premature atrial complexes [96]. In the majority of patients, AF is initiated by ectopic sources primarily found in the pulmonary veins [83]. These ectopic sources initiate wavefronts in addition to the already existing wavefront produced by the sinus node. The resulting interaction of these wavefronts may induce reentrant wavelets and AF. Even though the mechanisms of ectopic sources have not been fully defined, it

is assumed that such ectopic sources occur due to automaticity and triggered activity.

Normal and healthy cardiomyocytes do not show self-excitation or spontaneous activity [98], and are therefore unable to initiate ectopic triggers. Histological studies have, however, identified the presence of pacemaker cells, transitional cells, and Purkinje cells in human pulmonary veins. These cells have been shown to possess automaticity, and to perform spontaneous depolarizations [169].

Triggered activity can be categorized into delayed afterdepolarizations (DADs), or early afterdepolarizations (EADs), both being potential inducers of ectopic beats. Figure 1.6 depicts the two types of afterdepolarizations as single triggered depolarizations, as well as maintained excitation trains.

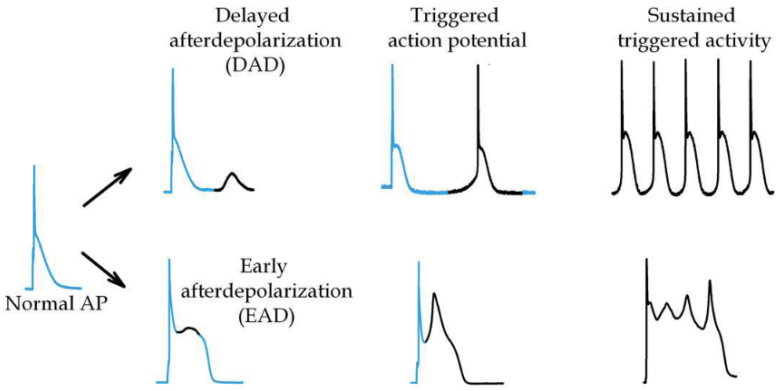

Figure 1.6: Delayed and early afterdepolarizations [193]

Abnormalities in cellular calcium handling, and, in particular, Ca^{2+} overload may lead to DADs. Increased cellular Ca^{2+} concentrations will activate the NCX, which, due to its electrogenicity, depolarizes the membrane. When reaching the threshold for a propagated AP such depolarization can trigger an ectopic excitation that may interfere with the physiological excitatory wavefront to produce AF. DADs are particularly common in patients with congestive heart failure, where cellular calcium load is increased [234].

EADs are reversals of repolarization during the second or third phase of the AP, and occur as a consequence of prolonged action potential duration, during which CA channels recover from inactivation an reopen to produce a small depolarization [158]. EADs have been observed in ventricular cardiomyocytes as a result of mechanical stretch, which prolongs action potential durations [45, 105].

Maintenance of Atrial Fibrillation

Episodes of AF may be of short duration, however a multitude of factors can contribute to its perpetuation. Therefore, beyond a trigger, AF requires a vulnerable atrial substrate to be maintained. Substrate vulnerability is mediated through electrical and structural changes, which foster heterogeneous conduction and re-entrant propagation patterns. While the mechanism of re-entrant propagation patterns remains controversial [96], three dominant hypotheses exist: re-entrant rotors [140], endo-epicardial dissociation [4] and the multi-wavelet theory [148]. All three potential mechanisms benefit from similar substrate modifications such as fibrosis, oxidative stress, pathologies of the autonomic nervous system or renin-angiotensin system, and genetic factors.

A commonly encountered indicator of substrate vulnerability is the degree of atrial fibrosis. The increased collagen content in the ECM modifies the conductive properties of the myocardium and produces heterogeneous conduction velocities, dispersion of refractoriness, conduction block, and electrical dissociation of neighboring tissue. Each of these phenomena may facilitate re-entries and wavelet formation and promote sustained AF. Fibrosis can be quantified using late gadolinium enhanced magnetic resonance imaging, and its quantity is a predictor of AF onset, as well as recurrence [2].

The autonomic nervous system is another major modifier of atrial substrate that may predispose to AF. The sympathetic nervous system has a major impact on Ca^{2+} handling of cardiomyocytes, leading to a potential shortening of APDs, which in turn creates a vulnerable substrate. The term “vagal AF“

has been used to describe AF in patients without structural heart disease, who tend to experience AF episodes during night time or post-exercise, when the parasympathetic tone dominates. Parasympathetic stimulation has been shown to prolong AF episodes in animal models, with increased heterogeneity in AP [137]. Further studies have indicated that complex relationships between sympathetic and parasympathetic activity are responsible for substrate vulnerability [134].

Oxidative stress and inflammation are central modifiers of atrial substrate that cause substrate vulnerability through a variety of processes. Inflammation commonly occurs as a result of cardiac surgery after which AF is frequently observed ("postoperative" AF). Further evidence points to inflammation decreasing the homogeneity of conduction velocities of atrial substrate promoting AF maintenance [44]. The mechanism of inflammation in AF is currently poorly defined. Inflammatory biomarkers were shown to be significantly increased in patients with AF demonstrating a bi-directional relationship between AF and inflammation [96].

Further modifiers of atrial substrate have been identified as promoting AF, such as changes in the renin-angiotensin system, ischemia, inflammatory diseases, hypertension, valvulopathies, as well as genetic factors [3].

Atrial Remodeling

Atrial remodeling is an adaptive response of the atrial tissue to mechanical or hormonal factors that alter the morphology or physiology of the atrium. Different pathological processes drive remodeling of atrial tissue constituents. Beyond AF risk factors such as ischemia, barometric overload, and aging, AF itself may cause atrial remodeling, due to its progressive nature. While initially occurring primarily in paroxysmal form due to the presence of triggers, repeated and prolonged episodes of AF lead to more persistent forms of longer duration. As such, the conversion to sinus rhythm (SR) becomes increasingly difficult the longer AF persists. This phenomenon, is often referred to as "AF begets AF", and describes the process of atrial remodeling, defined as “sustained functional or structural changes in atrial substrate“, which often promote the occurrence and maintenance of AF [158].

Prolonged maintained AF has been observed to modify the expression of cardiomyocytes' ion channels, leading to a shortening of the effective refractory period (ERP) [190]. Down-regulation of L-type calcium channels, modification of the sarcoplasmic reticulum ATPase as well as potassium channel changes

[73] have been observed in patients with AF. The mechanism responsible for the shortening of the ERP remains controversial, but it is generally agreed that a shortened ERP contributes to the substrate vulnerability by facilitating the formation of wavelets in accordance with the multiple wavelet theory [96]. Along with the ERP, the atrial contractile function decreases as a result of modifications in the intracellular Ca^{2+} handling. Both processes are, however reversible, when SR is restored and maintained for a sufficiently long period [190, 227].

Analyses of protein expression have revealed changes in connexin 40, indicating increased heterogeneity of gap junctions, which correlate with AF stability [215]. Spatial alterations in the expression of gap junctions may be responsible for heterogeneous conduction velocities, and increase the substrate vulnerability to re-entrant propagation patterns.

Structural changes like atrial dilation may occur as a result of valvulopathies [113] and other comorbidities, but also as the result of prolonged AF [150]. Atrial dilation is one of the most common markers to assess the vulnerability of the atrial substrate and has been shown to correlate with the degree of atrial fibrosis [209]. Both factors are known to increase substrate vulnerability and to promote AF [85]. Similar to contractile force and ERP, atrial dilation slowly decreases when SR is restored [78].

In summary, AF is a complex arrhythmia with rapid and disorganized electrical activity, of which the underlying pathophysiological processes are not completely understood. The general consensus is that AF requires a trigger and a maintenance mechanism, which originate from intra-atrial, but also extra-atrial pathophysiological processes. AF is associated with age, a variety of comorbidities and systemic diseases, as well as electrical and structural remodeling of atrial tissue. AF patients present significant heterogeneity in the combination of interacting pathophysiological processes, which are continuously being further elucidated, but are not completely understood [119, 226].

1.3.2 Treatment and Management of Atrial Fibrillation

The treatment and management of AF follow a holistic approach with a combination of therapies being employed to mitigate undesirable effects of AF. The major concern in AF patients is the risk of stroke, which is reduced by using anticoagulation therapy. AF may also be reverted to SR using electrical or pharmacological cardioversion. Alternatively, AF may be allowed to exist and a therapy managing AF symptoms may be preferred. This section outlines different options of treating and managing AF.

Anticoagulation and Stroke Prevention

Due to a five-fold increased risk of stroke in the AF population, anticoagulation is a crucial component of AF treatment. Stroke risk varies by the specific presentation of different risk factors that need to be assessed to determine the necessity of anticoagulation treatment. A major complication of anticoagulation therapy is the increased risk of lethal bleeding events. Along with the risk of stroke, the risk of bleeding must be evaluated to identify modifiable risk factors [89].

Stroke risk can be decreased with different pharmacological agents or through the use of an atrial appendage occlusion device. Pharmacological agents are classified into several categories:

Vitamin K Antagonists such as Warfarin have commonly been used for anticoagulation, but are gradually being replaced with novel oral anticoagulants (NOACs) due to easier dosing.

Heparins and Heparinoids accelerate the neutralization of several coagulation factors.

Direct Thrombin Inhibitors such as Dabigatran (NOAC) prevent thrombin from forming fibrinogen which prevents blood from clotting.

Direct Factor Xa inhibitors like Edoxaban (NOAC) prevent the transformation of prothrombin into thrombin inhibiting the formation of blood clots.

While NOACs are generally superior to vitamin K antagonists, valvulopathies and presence of artificial heart valves, warrant the use of vitamin k antagonists [89]. Among NOACs the choice of anticoagulation agent is largely unclear with similar outcomes observed across different types, but meta analyses indicate

increased mortality in diabetic patients treated with NOACs vs. vitamin K antagonists [166].

Blood clotting in AF patients is a result of blood stasis due to a lack of atrial contraction. The highest degree of blood stasis has been observed in the left atrial appendage, which led to the development of non-pharmacological stroke prevention therapies. One such therapy is the implantation of an occlusion device into the LAA, which has shown to considerably reduce hemorrhagic stroke risk. Alternatively, surgical exclusion of the atrial appendage may be performed. The non-inferiority of occlusion or exclusion of the LAA to pharmaceutical anticoagulation is primarily driven by the decrease in risk of hemorrhagic stroke associated with pharmacological anticoagulation. The decrease in hemorrhagic stroke risk is, however, traded for an increased risk in ischemic stroke. This makes LAA occlusion or exclusion potentially superior to pharmaceutical anticoagulation in patients with an increased risk of bleeding events. Randomized controlled trials to support this assumption are however missing [89].

Rhythm Control

Rhythm control strategies aim to restore and maintain SR. Possible rhythm control strategies are electrical cardioversion, pharmacological cardioversion, lifestyle changes, or combinations thereof. Rhythm control is primarily aimed at reducing symptoms and an improving in patients' quality of life. It is indicated as the first line treatment, and should be attempted in case of uncertainty regarding its outcome [89].

Electrical Cardioversion

Electrical cardioversion can be performed using a synchronized direct current shock applied to the patient's chest via two electrodes. The exact mechanism by which electric cardioversion terminates arrhythmia is not fully understood. The two primary hypotheses are that the applied electric shock depolarizes the cardiomyocytes' transmembrane potential past the activation threshold for APs [52], and that depolarization of a critical mass of tissue prevents reentrant patterns from continuing [240]. The other hypothesis states that the electric shock produces a prolongation of APDs that is sufficient to terminate chaotic conduction patterns [208].

Before electric cardioversion is performed, transesophageal echocardiography is commonly performed to identify potential thrombi in the atrium. The failure

to detect a thrombus could, upon restoration of SR, lead to an ischemic stroke as the thrombus enters the circulatory system [89].

A combination of pharmacological cardioversion with electrical cardioversion was shown to be superior to only electrical cardioversion.

Pharmacological Cardioversion Pharmacological cardioversion can be performed using antiarrhythmic drugs (AADs), which are categorized into five classes depending on their mechanism of action [228]:

Class 1 AADs are Na^+ channel blockers. They act as stabilizers of the membrane by limiting cardiomyocyte excitability.

Class 2 consist of beta blockers that reduce sympathetic activity of the autonomous nervous system in the heart.

Class 3 agents prolong the APD by blocking K^+ channels leading to a prolongation of the effective refractory period, which prevents reentrant conduction patterns.

Class 4 are L-type Ca^{2+} channel blockers, thereby slowing the sinoatrial and atrioventricular node and increasing the effective refractory period leading to a decrease is re-entrant propagation patterns.

Class 5 contains AADs which do not belong to any of the above categories. This category contains AADs such as digoxin, adenosine, and atropine.

- Digoxin increases vagal activity thereby reducing conduction in the AV node.
- Adenosine decreases conduction velocity through the atrioventricular node while decreasing sympathetic activity and increasing parasympathetic tone.
- Atropine decreases parasympathetic tone and increases the conduction through the atrioventricular node.

Rate control

Rate control is a treatment approach in which AF is allowed to persist, but the ventricular rate is controlled to minimize symptoms. The best type and intensity of rate control treatment remains to be identified with randomized controlled trials showing no significant differences in composite outcomes between different treatment arms [89].

Pharmacological therapy with the use of beta blockers, diltiazem, verapamil, or a combination of drugs is performed to achieve a stable and low heart rate. The choice of drug is primarily driven by patient symptoms, comorbidities, and side-effects [89].

Conflicting evidence exists on the efficacy of rate control compared to pharmacological rhythm control. While some studies suggest that rate control is superior to pharmacological rhythm control with lower rates of adverse events being observed in a number of trials [194], other studies report the opposite observation [224]. Consensus exists, however, that patients treated with rate control strategies have a higher symptom burden than patients treated with rhythm control strategies [194].

When pharmacological rate control is not successful, an atrioventricular node ablation and pacemaker implantation may be performed to control the ventricular rate. This procedure has been shown to have low complication rates and low mortality risk [89].

Chapter 2

State of the Art

This chapter describes the state of the art in prediction of AF outcomes. Section 2.1 describes risk-scores, some of which are used in clinical practice to guide AF treatment, and the predictive biomarkers that these scores employ. Section 2.2 introduces the state of the art in phenotype identification from a machine learning perspective. Section 2.3 describes advances and limitations of the data-driven identification of AF sub-phenotypes. Finally, section 2.4 introduces CDSS and their design and evaluation frameworks.

2.1 Clinical Stratification for Atrial Fibrillation

In community cohorts, clinical stratification methods are based on risk scores that predict the risk of an unwanted outcome. Different scores may be applied to predict the risk of (1) stroke, (2) major bleeding, (3) AF onset/incidence, or (4) AF recurrence following catheter ablation. This section presents the most common risk scores according to the *European Society of Cardiology Guidelines on Atrial Fibrillation* [89].

2.1.1 Risk Scores for Stroke Prevention

AF increases the risk of stroke five-fold. The patient specific risk, however, varies with the present risk factors. Great efforts are made to personalize anticoagulation treatment aimed at stroke prevention. Table 2.1 shows common risk scores employed in the estimation of stroke risk in AF patients.

Risk Score	Development Population	Development Model	External Validation C-Index
CHA_2DS_2–VASc [136]	35 countries Age: 66.0±14.0 FU: 1 year	Logistic Regression	0.61 (NA)
ABC-stroke [87]	North America Latin America Europe Asian Pacific Age: Median 70 FU: Median 1.9 years	Cox Regression	0.66 (0.58, 0.74)
ATRIA [200]	California Age: 71.7±11.6 FU: 6.50-7.25 years	Cox Regression	0.70 (0.67, 0.72)
Intermountain [79]	NS Age: 71±12 (m) 68±12 (w) FU: 5.8±4.1 years	Cox Regression	0.71 (w) 0.72 (m)
GARFIELD-AF [66]	Caucasian Hispanic-latino Afro-Caribbean Asian Mixed/other Age: Mean 71 FU: Up to 1 year	Stepwise Regression	0.75 (0.73-0.77)

Table 2.1: Risk Scores for Stroke Stratification. NS = Not stated. FU = Follow-up.

Out of the presented risk scores, the CHA_2DS_2–VASc score is the recommended choice to guide anticoagulation treatment [89]. It makes use of the following variables:

Variable	Score
Congestive heart failure/ Left ventricular dysfunction	1
Hypertension	1
Age > 75	2
Diabetes mellitus	1
Previous stroke	2
Vascular disease	1
Age 65-74	1
Female gender	1
Maximum	9

Table 2.2: Biomarkers used by the CHA_2DS_2–VASc score

The CHA_2DS_2–VASc estimates stroke risk on a level from 0 to 9 by adding scores for each variable present in a specific patient. In patients scoring 2 points or more, anticoagulation therapy is recommended.

2.1.2 Risk Scores for Major Bleeding

Risk of major bleeding events are of major concern in AF patients. This is especially true for patients being treated with anticoagulants, which increase the risk of bleeding events. A personalized assessment of bleeding risk is therefore of importance. Several bleeding risk scores have been proposed to stratify patients according to bleeding risk. Table 2.3 presents the most common risk scores.

Risk Score	Development Population	Development Model	External Validation C-Index
HAS-BLED [172]	Euro Heart Cohort Age: 66±13 (nb) 73±10 (b) FU: 1 year	Logistic Regression	0.72 (0.65-0.79)
ABC-bleeding [88]	ARISTOTLE Age: 70 (19–97) FU: Median 1.7 years	Cox Regression	0.71 (0.68-0.73)
HEMORR$_2$HAGES [71]	US Medicare beneficiaries Mean age: 80.2 years FU: Mean 0.82 years	Logistic Regression	0.60 (0.51–0.69)

Table 2.3: Risk Scores for Major Bleeding Events. FU = Follow-up

While a consensus on the utility of the HAS-BLED score has not been reached, the weight of evidence is in favor of its efficacy. As such, guidelines recommend its consideration for the identification of modifiable bleeding risk factors, and to identify patients at increased risk of bleeding. Such patients should receive a more frequent follow-up, while a high bleeding risk is not necessarily a contraindication for anticoagulation [89].

2.1.3 Risk Scores for AF Onset

Risk scores aimed at identifying patients with an elevated risk of developing AF have been developed. Such high-risk individuals could potentially benefit from preventative treatment such as prophylactic stroke prevention [135]. Even though a variety of scores have shown promising results, their use in clinical practice remains low, and guidelines do not make any recommendations for their use [89]. Table 2.4 presents risks scores developed for the assessment of AF onset risk.

Risk Score	Development Population	Development Model	External Validation C-Index
FHS [189]	Caucasian Age: 61 (45-95) FU: 10 years	Cox Regression	0.78 (0.76-0.80)
CHARGE-AF [5]	81% white 19% black Age: 65 (46-93) FU: Up to 7 years	Cox Regression	0.78 (0.75-0.78)
WHS [58]	Caucasian Female Age: Median 54 years (IQR 49-59) FU: Median 14.5 y	Cox Regression	0.72 (0.68–0.75)
MHS [7]	Israel 54% Female Age: Mean 63 years FU: 10 years	Cox Regression	0.72 (0.68–0.75)
JMC [84]	Japanese 35% Female Age: Mean 52 years FU: 5.5±1.6 years	Cox Regression	0.77 (SE 0.02)
C_2HEST [131]	Asian 43% Female Age: Mean 47 years FU: Median 2.6 years	Cox Regression	0.75 (0.73-0.77)
Shandong [47]	Chinese 33% Female Age: 57 (45-85) FU: Median 2.6 years	Cox Regression	0.77 (NA)

Table 2.4: Risk Scores for AF Onset. FU = Follow-up

2.1.4 Risk Scores for AF Recurrence

Risk scores have also been developed for the assessment of AF recurrence risk following catheter ablation. Table 2.5 shows common risk scores to assess AF recurrence.

Risk Score	Development Population	Development Model	External Validation C-Index
Apple [117]	Germany Age: 60±10 FU: 1 year	Logistic Regression	0.63 (0.60–0.66)
DR-FLASH [118]	Germany Age: 61±10 FU: 2 years	Logistic Regression	0.80 (0.74–0.87)
MB-LATER [152]	Serbia Age: 56.9±11.8 years FU: 29.1±10.1 months	Cox Regression	0.62 (0.54-0.69)
ATLAS [143]	Portugal Age: 59±11 years FU: 4.20±2.70 years	Cox Regression	0.75 (NA)
CAAP-AF [229]	California Age: 62.3 ± 10.3 years FU: 2.5 ± 1.7 years	Cox Regression	0.65 (NA)
BASE-AF2 [28]	California Age: 4.6 ± 10.5 years FU: 6 months	Cox Regression	0.65 (0.58-0.71)
ALARMEc [231]	NA 29% Female Age: 58 (50-65) years FU: Median 15.6 months	Cox Regression	0.49 (0.42-0.56)

Table 2.5: Risk Scores for AF Recurrence. FU = Follow-up

According to clinical guidelines for community cohorts, factors associated with an increased risk of AF recurrence are LA size, AF duration, patient age, renal

dysfunction, and fibrosis in the myocardium, which can be visualized by means of MRI. Guidelines recommend that these risk factors be taken into account for treatment selection [89].

2.2 Machine Learning

Machine learning is a branch of artificial intelligence, which is concerned with the study of algorithms that provide learning capability to computers without being explicitly programmed [187]. Machine learning algorithms may be categorized into four categories: Supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning [151]. The following sections will discuss the first three categories. For an overview of reinforcement learning, the reader is referred to [103].

2.2.1 Supervised learning

Supervised learning is a category of machine learning algorithms which use an input to predict an output. The goal of such algorithms is to to statistically model the input-output relationship in order to make predictions on unseen samples. Examples of such tasks are regression tasks, where a continuous value (i.e. blood pressure, age) is to be predicted, classification tasks, where binary (i.e. male/female) or multi-class label (i.e. nationality) is required, or survival analysis tasks in which the time to an event is of interest is predicted [29]. Figure 2.1 visualizes the difference between classification and regression.

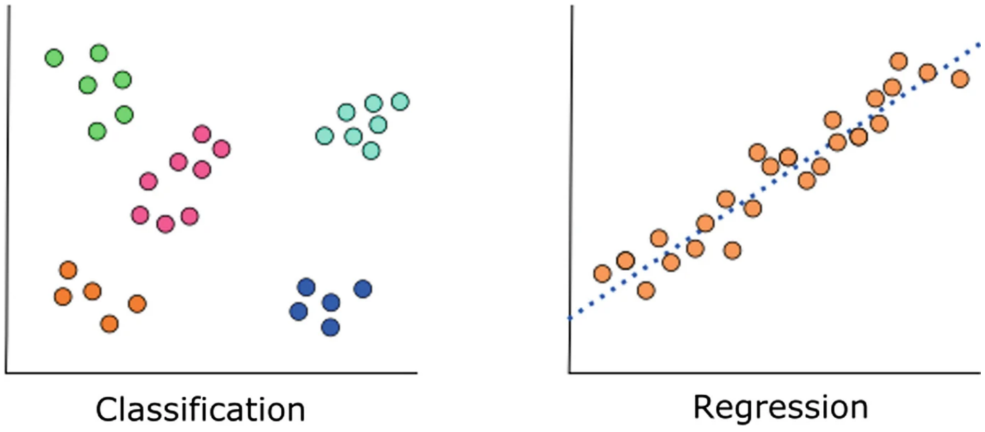

Figure 2.1: Classification vs. Regression [18].

An abundance of methods have been developed to perform supervised learning tasks, ranging from linear and logistic regression models such as used for the risk scores introduced in section 2.1, to the use of decision tree ensembles and artificial neural networks. For an overview of supervised machine learning methods, the reader is referred to [198].

2.2.2 Unsupervised learning

Unsupervised learning uses solely input data and makes no use of labels. The goal of such algorithms may be dimensionality reduction, outlier detection, or clustering. Clustering algorithms learn a dataset's underlying structure in order to draw conclusions about the existence of distinct groups within the data, and are commonly used to identify patient phenotypes in heterogeneous cohorts. They are primarily driven by high-density regions in the feature space, and output group assignments for each observation in the training dataset [127].

Clustering algorithms use a set of features to distinguish and separate observations into groups. These algorithms can generally be classified into the following categories [6]:

Partitioning techniques divide a dataset into a fixed number of groups, each represented by a centroid in the feature space. An illustration of this is the K-Means clustering algorithm [100]. In K-Means clustering, centroids are initially chosen randomly in the feature space, and data points are assigned to their closest centroid based on a distance metric such as Euclidean distance. Subsequently, centroids are adjusted iteratively by computing the mean of points associated with each centroid until convergence criteria are met. Although K-Means converges rapidly, it is sensitive to noise and outliers, disregards variations in cluster shapes, sizes, and densities, and requires the user to specify the number of clusters, potentially facing convergence issues due to local minima [201].

Hierarchical methods conduct a hierarchical clustering of data points, yielding a dendrogram illustrating the clustering sequence from a single inclusive cluster at the apex to individual data points as singleton clusters at the base. An example of a dendrogram is shown in figure 2.2. Hierarchical clustering techniques merge or split clusters in accordance with distance metrics in iterative steps [201]. Dendrograms aid users in identifying suitable distance thresholds to define final clusters. Although hierarchical approaches furnish useful insights due to their ability of representing hi-

erarchies, they are associated with high time complexity, rendering them computationally intensive [11].

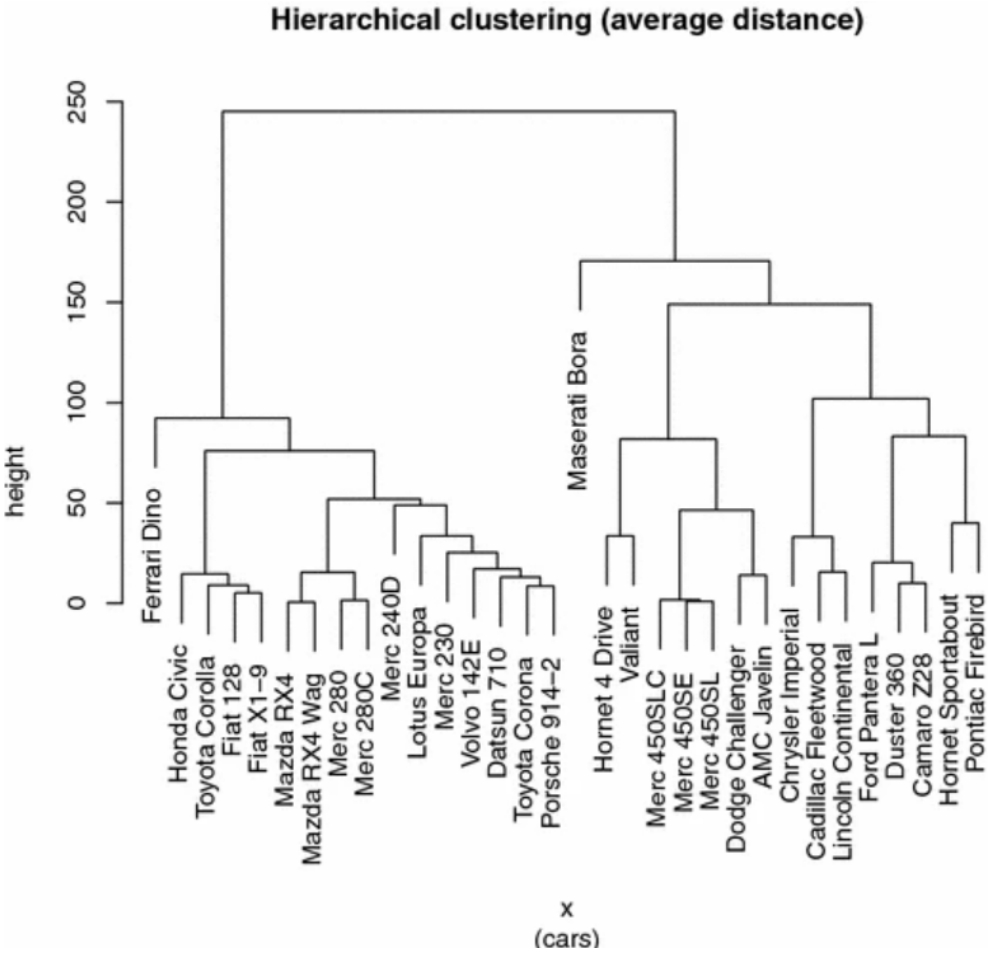

Figure 2.2: Example of a Dendrogram from [159].

Density-based methods utilize local density assessments to similar observations into the same clusters. Rather than relying solely on pairwise proximity, these algorithms identify clusters as regions of high density separated by areas of lower density. Density-based approaches are resilient to noise and capable of identifying clusters of various shapes but struggle with

reliably describing cluster shapes and are unsuitable for high-dimensional datasets [201, 11].

2.2.3 Semi-supervised learning

Semi-supervised learning algorithms address the problem of partially missing labels by using unlabeled data together with labeled data. Semi-supervised algorithms may be used to in different applications such as regression, classification, anomaly detection, and clustering tasks [29].

The underlying proposition of semi-supervised learning is that the distribution of feature p(x) contains information about the posterior distribution of interest p(y|x). If this proposition is met, then the inclusion of unlabeled observations may improve the accuracy of predictions [238].

For this premise to hold, several assumptions are made about the analyzed dataset [56]:

1. The **smoothness assumption** states that for two observations, which are similar in the feature space, the labels should be the same.
2. The **low-density assumption** states that the decision boundary of a classifier should pass through low-density regions in the input space, as defined by the feature space p(x).
3. The **manifold assumption** states that observations in a high-dimensional feature space are usually concentrated on lower-dimensional structures called manifolds. Observations in such structures are said to usually carry the same labels. If such manifolds can be identified, and the labels for a subset of observations in each manifold is known, the labels for the remaining observations can be inferred.

Given the use of clustering algorithms for phenotype identification in patient cohorts, semi-supervised clustering are of particular interest to this work. While semi-supervised classification is a relatively well understood problem, semi-supervised clustering is a relatively small research area [56].

In semi-supervised clustering, observations may be partially labeled, or have must-link or cannot-link constraints between observations [126]. Other works have suggested the use of survival data into the clustering process, to form clusters which are predictive of an event of interest [13]. Such works will be discussed further in section 4.2.1.

2.3 Atrial Fibrillation Phenotypes

AF patients are commonly classified according to their clinical sub-phenotype (paroxysmal, persistent, long-standing persistent), which describes the AF burden and AF duration. This definition captures exclusively the temporal pattern of AF episodes that patients experience, and does not represent the underlying pathophysiological mechanisms that are causing AF. An alternative, knowledge driven, classification into sub-phenotypes have been proposed, classifying AF into being secondary to structural heart disease, focal, polygenic, and other classes corresponding to the understanding of the AF mechanism [115]. The validation of this classification's clinical use is, however, lacking [89]. In contrast to the knowledge-driven classification, recent works employ data-driven methods to uncover sub-phenotypes in the AF population.

Unsupervised clustering analyses of large datasets have, in the past, identified patient sub-phenotypes that correspond to different pathological processes and show different treatment responses [176]. Within the AF population, several studies have investigated the composition of the population by means of the hierarchical agglomerative clustering (HAC) algorithm. Most notably, Inohara et. al. [95] identified patient sub-phenotypes based on 60 clinical biomarkers in a first application of unsupervised machine learning algorithms to an AF cohort. The authors used hierarchical agglomerative clustering [221] to identify four sub-phenotypes, which the authors interpret as:

1. An **atherosclerotic-comorbid** cluster, with the majority of patients having a diagnosis of coronary artery disease and the highest proportion of prior myocardial infarction. Given the highest rates of heart failure, hypertension, hyperlipidemia, diabetes, chronic kidney disease, and anemia, patients in this cluster also exhibited to highest CHA_2DS_2–VASc scores.

2. A **tachy-brady/device implantation** cluster, with patients characterized by their resemblance to patients with tachycardia-bradycardia. Patients in this cluster had the highest prevalence of device implantation due to sinus node dysfunction or atrioventricular node ablation. They are further characterized by the highest European Heart Rhythm Association symptom scores.

3. A **low comorbidity** cluster, which was the largest cluster identified. Patients in this cluster were the second-youngest, had the lowest rates of diabetes, OSA, heart failure, and COPD. In accordance with relatively

good health, patients in this cluster exhibited the highest LVEF and lowest CHA_2DS_2–VASc score.

4. A **youger behavioral disorder** cluster characterized by the highest rate of liver disease, alcohol abuse, drug abuse, and prevalence of smoking. Patients in this cluster were the youngest and most likely to be male. Further, this cluster had the highest BMIs and the highest prevalence of COPD and OSA, but, remarkably, the best renal function measured by eGFR.

In addition to the identification of existing sub-phenotypes, the authors demonstrated differences in the treatment strategies employed in the different sub-phenotypes. Significant differences exist in the prevalence and selection of anti-arrhythmic drugs, prevalence of rate control, as well as the use of anticoagulation therapy. Further, the authors quantified the rates of clinical outcomes within different clusters. Significant differences were found between the obtained sub-phenotypes, with differences in death, stroke rates, bleeding events, and hospitalization. The authors conclude that the results of their study confirm the heterogeneity of the AF population, while demonstrating that the traditional clinical classification of paroxysmal, persistent and permanent AF did not drive cluster formation, and are therefore not defining features of the population.

Similar studies [94, 206, 217, 223] have used HAC to identify sub-phenotypes of AF patients. Generally a low number of clusters (3-5) were identified, and differences in outcomes and treatments were described. Of note, all studies rely solely on geometrical measures of the covariate space to cluster patients. Such unsupervised machine learning approaches have previously been shown to produce sub-phenotypes that are not consistent with patient outcomes making the clinical value questionable [14, 72]. In fact, the AF sub-phenotypes identified in the above studies lack the resolution to allow for a treatment selection similar to the one currently performed in clinical practice. Contraindications such as thyroid dysfunctions an heart failure can generally not be identified in the over-agglomerated sub-phenotypes, making a treatment selection based on the sub-phenotype impractical and imprecise.

While different AF sub-phenotypes may be identified using an unsupervised clustering algorithm, the utility of the sub-phenotype is limited, if the treatment response of different sub-phenotypes is the same, or worse, single sub-phenotypes contain groups with different treatment responses. The latter of which, may be assumed to occur in the previously mentioned studies due to the low number of identified clusters. While this shortcoming of unsupervised

clustering has been recognized by literature, and semi-supervised clustering (SSC) methods are increasingly being employed to tackle this shortcoming [14], semi-supervised methods have yet to be employed within the analysis of AF populations.

The identification of patient sub-phenotypes *within the context of outcome prediction* is of critical importance in the development of interpretable models for precision medicine [36, 50]. SSC algorithms that combine patient covariates and survival outcomes may therefore be preferred, as they produce clusters that are not only biologically meaningful, but also clinically relevant.

2.4 Clinical Decision Support Systems

CDSS are computer-based systems that provide its users with patient-specific recommendations or alerts to support clinical decision-making. The development and implementation of CDSS is, however, complex and success and adoption rates vary. Prior works have recognized the complexity of CDSS developement, and proposed formal frameworks for the development and evaluation of CDSS [80].

2.4.1 Five Rights of CDSS

Fundamentally, a CDSS should follow five basic principles: provide the right information, to the right person, in the right format, through the right channel, at the right time [162]. Olakotan et. al. [161] have previously systematically evaluated a large number of alert-based CDSS and identified common issues in terms of the five basic principles and the socio-technical aspects suggested by [222] (table 2.6).

Within the socio-technical dimension the right channel choice, several studies provided an excess of information with poorly designed alerts such as choice of font or counter-intuitively arranged medication lists. A desensitization from “alert fatigue“ was observed where an overwhelming number of alerts were generated, which may lead to increased mental workload, and, ultimately, alerts being ignored. To alleviate such information overload, an expert panel has suggested that CDSS should only provide essential information at the time when decisions are made [167], while others suggest the replacement of alert-based CDSS with "autopilots" to support decision-making before a decision is made by the provider [233]. Such “autopilots“ could evaluate EHRs and pro-

Table 2.6: Five rights principles with the according technological dimension and common issues according to [161]

Socio-technical dimension	Five rights principle	Issue
Technology	Right channel	Excess information Desensitization from alert fatigue
	Right information	Non-specific patient information
Human	Right person	Inability to fine-tune alerts
Organization	Right policies	Institutional/ government rules impeding alert use
Process	Right format	Interruptive/ non-interruptive alerts
	Right time	Poor alert timing

vide decision-makers with suggestions that maximize treatment benefits and minimize risks [192].

Several studies reported that alerts provided insufficient information which was difficult to interpret or unclear, slowing down clinicians' workflow. Similarly, information provided may not match the end-users needs, especially if end-users of different specialities are expected to use the CDSS. The information presented by a CDSS should therefore be tailored to the end-user in question, and, if necessary, be presented differently for different end-users. The presentation of the right information may go beyond the information required for decision-making, but provide information that further supports clinical workflows. A clinician faced with a treatment choice may therefore be presented the suggested treatment, with additional information such as patient weight, height, and dosage information [185, 207].

Information should be provided to the right person, in accordance with the persons needs. Alerts should be customizable, and not interfere with existing workflows. Several studies have reported alerts being raised during data-entry, which were usually bypassed [35, 165]. Depending on the type of alert, different providers should be alerted, and providers being able to decline alerts if the information is known to them. Such customizability requires appropriate training, which poses a challenge to the adoption of CDSS. Inadequate training of the end-user may increase workload [165].

Organizational rules may limit the usability of CDSS. Such impediments include institutions not allowing for the customization of certain software com-

ponents such as alert fonts or colors [237]. Further, the recommendations from CDSS may not be in accordance with internal rules or policies. Similarly, the necessity to provide justification for overriding CDSS recommendations may increase workload and reduce CDSS effectiveness [156].

Information should be presented in the right format. Within the scope of alerts, this may refer to alerts being interruptive or non-interruptive. Which alarms should be raised in a fashion that interrupts the current activity, and which should be raised passively, depends on the severity of the alert, but also depends on user preference [164].

Information should be presented at the appropriate time. Several studies reported that alerts were only generated following the prescription or dispensation of medication, increasing clinicians' workload. Unnecessary delays were identified in a variety of studies, which were caused by malfunctions, long computation times, and internet connectivity problems [31, 213].

2.4.2 CDSS Design Principles

Beyond outlining common pitfalls of previous studies, frameworks for the design of CDSS have been proposed. An example thereof is the framework proposed by Zikos [239], who defined seven design principles for a successful CDSS:

Principle 1: CDSS should mimic the cognitive process of clinical decision makers

Clinical reasoning is often characterized by a repeating loop of clinical assessment and data acquisition. A clinician may assess a patient, and request a laboratory test to fill "reasoning gaps". This fundamental element of clinical decision-making should be considered during the design of CDSS, and such processes should be replicated where appropriate [10].

Principle 2: CDSS should provide recommendations with longitudinal insight

Responses to treatments are often evaluated using specific physiological values, as responses to therapies may change physiological measurements. A measurement should therefore be evaluated in the context of prior measurements if a treatment response is to be evaluated. The temporal trends of physiological measurements should therefore be considered in the design of CDSS.

Principle 3: CDSS should 'know' the time when decisions will be made

The timing at which a clinical decision is made must be considered during CDSS development. In particular, data availability may be an issue when a patient is admitted to a hospital, because patient information not having been collected yet. Predictive models trained on data that was collected throughout a patients entire hospital stay may therefore show optimistic performance, which is not attainable in practice.

Principle 4: CDSS should provide predictions in a dynamic manner

In clinical practice, information is obtained and updated continuously, as laboratory measurements, vital signs, and diagnostic tests are gathered. Predictions should therefore be made dynamically as data arrives, and the reliability of the prediction should be evaluated internally by the CDSS. Only when a prediction can be made with reasonable confidence, should it be presented to the decision-maker. Such approaches are, however, limited by their computational demands, which limit system response time.

Principle 5: The 'Historical Decision' bias: CDS models should be outcomes-based

Historical decision bias occurs when models are trained on historical data that contains decision mistakes, such as misdiagnoses. Such models may show satisfactory performance when validated, but perform poorly in practice, because they reproduce past mistakes. To alleviate this issue, it is recommended to develop systems that make outcomes-based predictions instead of predicting diagnoses or treatment selections.

Principle 6: CDSS should model a-priori known interactions between clinical attributes

Different clinical variables are often combined to reach a conclusion. Existing knowledge on the interactions of different clinical variables and their implications should therefore be modeled explicitly. An example thereof is the diagnosis of a disease with the use of a laboratory test and a symptom. The modeling of such interactions, however, requires extensive knowledge of the clinical domain in question.

Principle 7: Dimensionality reduction helps train models on-the-fly, but should be done with caution

Medical data is often highly dimensional. The International Classification of Diseases (ICD-10-CM), for example, contains more than 69,000 individual disease codes. Such high dimensional coding makes data mining infeasible

and requires dimensionality reduction, which may be accomplished using dimensionality reduction algorithms or the creation of variable groups based on domain knowledge.

2.4.3 CDSS for the Management of Atrial Fibrillation

Several CDSS have been proposed to guide support the treatment of AF patients. The scope of these systems has been education, improving communication between patients and healthcare providers, improving patient involvement in the decision-making process, and providing treatment recommendations.

A popular choice in designing CDSS for AF management is the use of point scores to predict the risk of ischemic stroke and major bleeding to support the decision of initiating anticoagulation therapy. Numerous studies have used such an approach with varying results. The DARTS II CDSS [211], for example, provides a computerized decision aid the prescription of anticoagulation therapy in outpatient clinics. The system uses the Framingham equation [232] to predict patients' stroke risk, and estimates the effect of warfarin therapy on stroke risk and bleeding risk based on information from systematic literature review. The predicted risks are communicated to patients, using the risk communication screen shown in figure 2.3. In addition to a numerical presentation of risks in terms of yearly rates, patients are presented with 100 smiley faces to visualize the computed risk for stroke and stomach bleed.

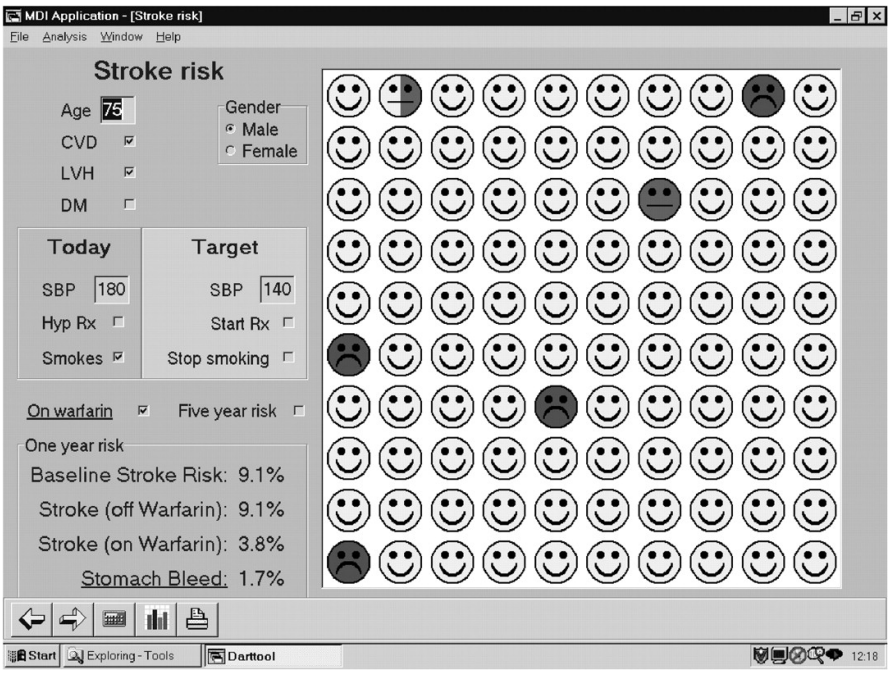

Figure 2.3: DARTS II Clinical Decision Support System [212]

In an RCT evaluating the impact of DARTS II [212], patients were randomized to either usual care according to guidelines, or the supplementation of usual care with the CDSS. In the CDSS arm, patients interacted with the CDSS, and a shared decision between patient and physician is was made whether to initiate anticoagulation therapy. The DARTS II CDSS led to a reduction in decision conflict between patients and clinicians, but also a reduction of warfarin use, which, the authors conclude, may ultimately lead to an increased risk of stroke in patients treated with the CDSS.

Similarly, Fraenkel et. al. [67] proposed a decision support tool to support the education of AF patients and support shared decision-making between patients and care providers. The tool uses the CHA_2DS score to determine the risk of stroke and the $HEMORR_2HAGES$ score to predict major bleeding risk, both adjusted for three treatment options: No treatment, aspirin, warfarin. The expected incidence rates under the treatment options are communicated to the patient, to educate patients on the available choices and reduce physician-patient conflict. Similarly to Thomson et. al., Fraenkel chose to present risks under different treatments using 100 smiley faces. In addition

to communicating the risks under different treatments, the CDSS included an educational component, which explained the pathogenesis of stroke in AF patients. In an RCT the tool was shown to improve patient understanding, and patient-clinician communication, but did not significantly impact treatment choice [68].

CDS-AF is a CDSS which evaluates the CHA_2DS_2–VASc score to determine if a given patient should be anticoagulated or not [108]. Upon the logging of a patient into the hospital's electronic journal, the CDSS is activated, and the clinician is presented with an overview of the patient's diagnoses, obtained from the EHR system, and given the possibility to adjust these diagnoses. The CHA_2DS_2–VASc is computed, and the yearly stroke risk is estimated. The clinician can thereafter decide to initiate anticoagulation therapy if indicated by current guidelines, or postpone the decision. In a RCT [107] the authors demonstrate a higher adherence to guidelines and lower rates of bleeding events in the CDSS group than the control group. No improvements were, however, observed in the rates of ischemic stroke. A similar tool was proposed by van Doorn et. al. [51]. Its evaluation in a RCT has shown no improvements in stroke incidence, major bleeding risk, or anticoagulation over- or underuse. Similarly, Silbernagel et. al. [199] have evaluated a similar system in a prospective study, and observed a modest improvement in guideline adherence.

Similarly, Sheibani et. al. [197] have proposed a mobile CDSS in the form of a smartphone application that integrates the CHA_2DS_2–VASc score to predict stroke risk and the HAS-BLED score to determine the risk of major bleeding events. The application allows clinicians to manually select risk factors, and presents the corresponding stroke and bleeding risk. Treatment recommendations were based on guidelines from the American Heart Association [15], and propose the use of anticoagulants or aspirin. The authors reported an increased rate of guideline adherence in a prospective study, but the impact on stroke or bleeding was not evaluated.

LaHaye et. al. [124] have implemented a CDSS using the CHA_2DS_2–VASc score for stroke prediction, and the HAS-BLED score to asses the risk of major bleeding events. The impact of different treatments on stroke and bleeding risks are computed based on relative risks reported by meta analyses and RCTs. In contrast to other studies, the authors explicitly implemented an algorithm for recommending specific treatments based on their expected stroke risk, bleeding risk, and treatment cost in the form of a utility function. Depending on the computed risks for stroke and bleeding, the CDSS may recommend no treatment, the use of aspirin, apixaban, or dabigatran. An evaluation of the proposed system remains undone.

Ru et. al. [182] have proposed and evaluated another CDSS for recommending anticoagulation treatment in patients treated by general practitioners. The proposed CDSS is integrated into the hospital information system, and is opened by the GP, when a patient is diagnosed with AF for the first time. After completing patient related information, the CDSS computed the CHA_2DS score and the CHA_2DS_2–VASc score to predict stroke risk, as well as the HAS-BLED, ORBIT, and ATRIA score to predict the risk of major bleeding events. A recommendation of initiating anticoagulation is based on Chinese guidelines [91]. In a prospective study the authors showed an increase in adherence to guidelines, and a lower incidence rates of adverse effects in the CDSS group compared to the control group.

Michalowski et. al. [144] have further refined this approach by introducing *AFGuide*, a mobile application to support anticoagulation treatment in the outpatient setting. The authors implement rules from the Canadian AF guidelines and drug-drug interactions as an executable first order logic. The former are used to identify all possible anticoagulation therapies for a given patient, while the latter is used to identify potentially occurring adverse events, which may occur when treating a patient with multiple comorbidities. Further, *AFGuide* predicts patients' adherence to therapies and preferences using differential factors that characterize patients and therapies. These differential factors are obtained from previously published review papers and patient websites. These factors are used to develop scoring functions predicting individual adherence-to-therapy. Stroke and bleeding risk are assessed using unspecified point scores, which are synthesized with the results of clinical trials to provide a final risk assessment. The first order logic rules are applied to identify feasible anticoagulation therapies for a patient in question, and are annotated with the associated stroke and bleeding risks, and weighted by the scoring functions arising from the adherence-to-therapy models. Physicians are then provided with a ranked list. A therapy explanation module further provides justifications and explanations for the generated recommendations, which is supplemented by with clinical evidence such as systematic reviews. The module acts passively, and will only provide an explanation if requested. An evaluation of *AFGuide* remains to be performed.

A recent meta-analysis [186] has shown that CDSS that recommend antithrombotic therapy based on guideline recommendations significantly improve guideline adherence (RR: 1.03 [1.01-1.04]) in outpatient settings. Even though the analysis revealed that the risk of thromoembolic events does not significantly differ between CDSS and control groups, the risk of major bleeding tends to be lower. A further scoping review has found that the use of CDSS decreases deci-

sion conflict between patients and physicians, and increases patient knowledge about the risks of AF and AF treatments [196].

QRhythm [203] is a novel CDSS that supports treatment selection in outpatient populations. It uses 9 predictive variables to predict the probability of clinicians prescribing an antiarrhythmic rhythm control medication, a rate control medication, performing a electric cardioversion, or performing an ablation. Physicians are used to input data for a patient in question, and a linear regression model predicts the probabilities of the patient in question receiving each treatment. The system uses a two-step process to train its linear regression model: In the first stage, a retrospective cohort of 100 patients is used for model optimization, predicting the assignment of each treatment. During clinical use, a reinforcement learning algorithm is used to optimize the model such that the occurrence of stroke, hospitalization, and symptomatic AF recurrence are minimized. The CDSS' user interface is shown in figure 2.4.

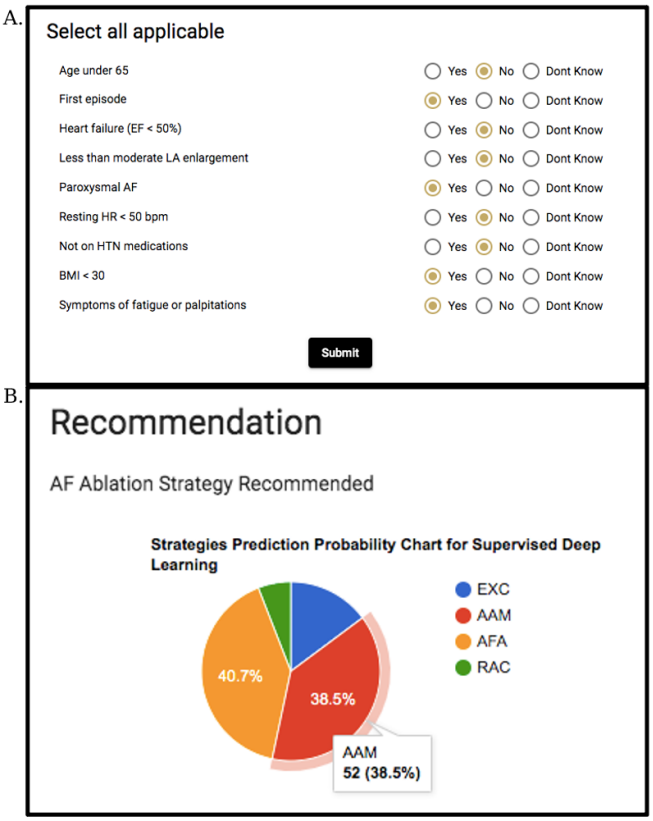

Figure 2.4: QRhythm CDSS [203]
Distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

Recently, CDSS are becoming increasingly complex with recent works spanning an abundance of functions and aiming for the facilitation of a holistic treatment approach. In particular, they aim to integrate the ABC pathway proposed by the ESC guidelines [89]: **A**voiding stroke, **B**etter symptom control, **C**ardiovascular and comorbidity risk optimization. For example, the CATCH-ME Healthcare Professional App [120] is a mobile application developed for AF patients and their treating physicians. It supports patient education in the form of booklets on AF pathology, symptoms, prognosis, associated comorbidities, treatment strategies, and tips on self-management. It allows patients to submit their health issues, and provides a personal health record and a symptom diary. Physicians are provided with a different version of the app, which allows them to view treatment guidelines, consult patient data, and access interactive treatment algorithms.

The mAFA platform [81] is a smartphone app that was developed to support patients in the management of AF in an outpatient setting. It contains classical decision support tools such as the CHA2DS2-VASc and HAS-BLED scores, guideline-based treatment recommendations, and educational materials. It provides dynamic assessments of bleeding risk, and flags modifiable risk factors to patients and physicians. Figure 2.5 shows the CDSS assessing a patient's bleeding risk, and the quality of anticoagulaiton based on the time in therpeutic range. Symptoms are quantified using the European Heart Rhythm Association symptom assessment scale, and heart rhythm and heart rate are captured using a photoplethysmography system. A prospective RCT has shown that mAFA reduces the use of OACs, and decreases the risk of bleeding, stroke, AF recurrence, heart failure, and hospitalization [82].

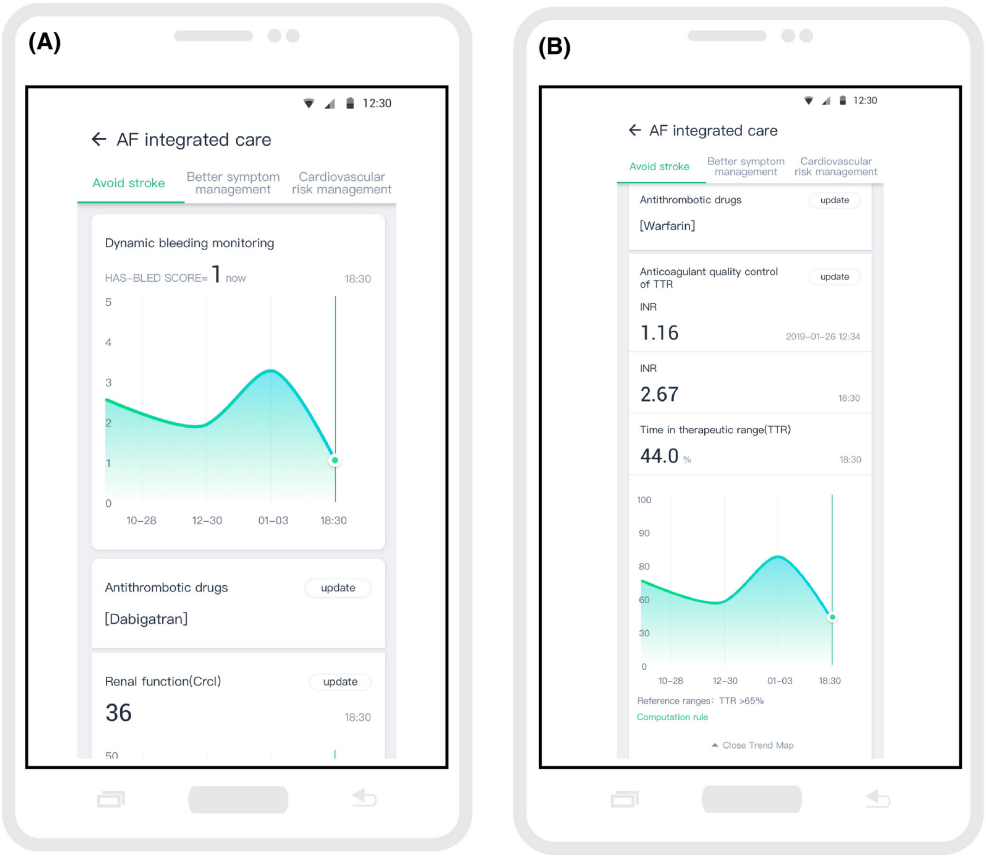

Figure 2.5: mAFA Clinical Decision Support System [81]. A) Dynamic bleeding monitoring; B) Anticoagulation quality.

A similarly CDSS has been proposed by Cox et. al. [38], which makes use of electronic health records to support clinical decision-makers and patients in the management of AF in primary care. It allows patients to enter clinical data such as symptoms, heart rate and blood pressure, which are relayed to their physician via the CDSS. Physicians are supplied with a web-based platform, which displays patient-reported data, as well as laboratory results and prior diagnoses. The platform automatically reports CHA_2DS_2–VASc score and HAS-BLED scores, and provides alerts to recommend changes in treatment based on clinical guidelines.

A summary of selected CDSS is provided in table 2.7. To date, the development of CDSS for the management of AF is limited to outpatient settings. Primarily, CDSS are designed to support patients in education and self-management, and aid in adhering to clinical guidelines, with the integration of point scores for treatment decisions being a popular method. The application of predictive algorithms is primarily used for binary decisions such as initiating anticoagulation treatment or not. CDSS for the selection of specific drugs remain rare, and are often limited to guideline recommendations. Recent advances show a trend towards mobile applications, which provide an abundance of functionalities, and are aimed at enabling holistic treatment, patient monitoring, and physician-patient communication. While RCTs show that these systems often improve guideline adherence, and sometimes clinical outcomes, an application of clustering-based methods which have been used to address patient heterogeneity have not been employed by any CDSS to date. Similarly, CDSS are limited to the outpatient setting, with critical care not being considered by a single CDSS.

Table 2.7: Summary of CDSS for the management of AF

End-User	Patient	[67], [212]
	Physician	[51], [182], [108], [124], [144], [197], [199], [203]
	Patient and Physician	[38], [81], [120]
Setting	Outpatient	[38], [51], [67], [81], [108], [120], [124], [144], [182], [197], [199], [203], [212]
	Inpatient	-
Time of Action	Pre-treatment	[51], [67], [108], [124], [144] [182], [197], [199], [203], [212]
	Continuous	[38], [81], [120]
Input Data	Demographics	[38], [67], [81], [108], [120], [124], [144], [182], [197], [199], [203], [212]
	Medical History	[38], [51], [67], [81], [108], [120], [124], [144], [182], [197], [199], [203], [212]
	Vital Signs	[38], [67], [81], [120], [144], [182], [203], [212]
	Treatments	[38], [67], [81], [120], [124] [144], [182], [197], [199], [203], [212]
Data Processing Method	Linear Regression	[124], [203]
	Logistic Regression	[38], [51], [67], [81], [108], [120], [124], [144], [182], [197], [199], [212]
	Cluster Analysis	-
Purpose	Patient Education	[67], [81], [120]
	Self-management	[81], [120]
	Monitoring	[38], [81], [120]
	Patient-Physician Communication	[81], [120]
	Binary Treatment Indication	[38], [51], [67], [81], [108], [120], [124], [182], [197], [199], [203], [212]
	Drug Selection	[38], [67], [81], [124], [144], [203]

Chapter 3

Hypotheses and Objectives

Current literature has proposed diverse CDSS, as well as analyzed the heterogeneity of AF populations and their outcomes. While state of the art approaches identify data-driven AF phenotypes using unsupervised clustering algorithms, an application of semi-supervised methods remains undone. Similarly, to what extent a CDSS based on phenotype classification may find acceptance among clinical decision makers remains unstudied. This work aims to close these gaps by evaluating the following hypothesis with the help of several objectives outlined in this chapter.

3.1 Hypotheses

The underlying assumption of this work is that semi-supervised clustering algorithms are capable of identifying AF phenotypes which are biologically meaningful and clinically useful. Further, it is assumed that the use of decision-making algorithms can provide an useful method of selecting AF treatment. Based on these assumptions this work evaluates the following hypothesis:

Semi-supervised clustering methods can identify sub-phenotypes of AF patients with varying treatment effects and guide treatment selection.

3.2 Objectives

The hypothesis is tested within the scope of the following primary objectives.

1. Development of a semi-supervised clustering algorithm.
2. Identification and description of sub-phenotypes and their treatment responses using said semi-supervised clustering algorithm.
3. Evaluation of the proposed framework in terms of usability by clinical decision-makers.

The primary objectives are supported by the following secondary objectives.

1. Benchmark study of the semi-supervised clustering algorithm.
2. Standardization of clinical records into a common data model.
3. Development of a domain ontology allowing for the capture of relevant concepts from clinical records.

A conceptual view of the hypothesis and supporting objectives is shown in figure 3.1.

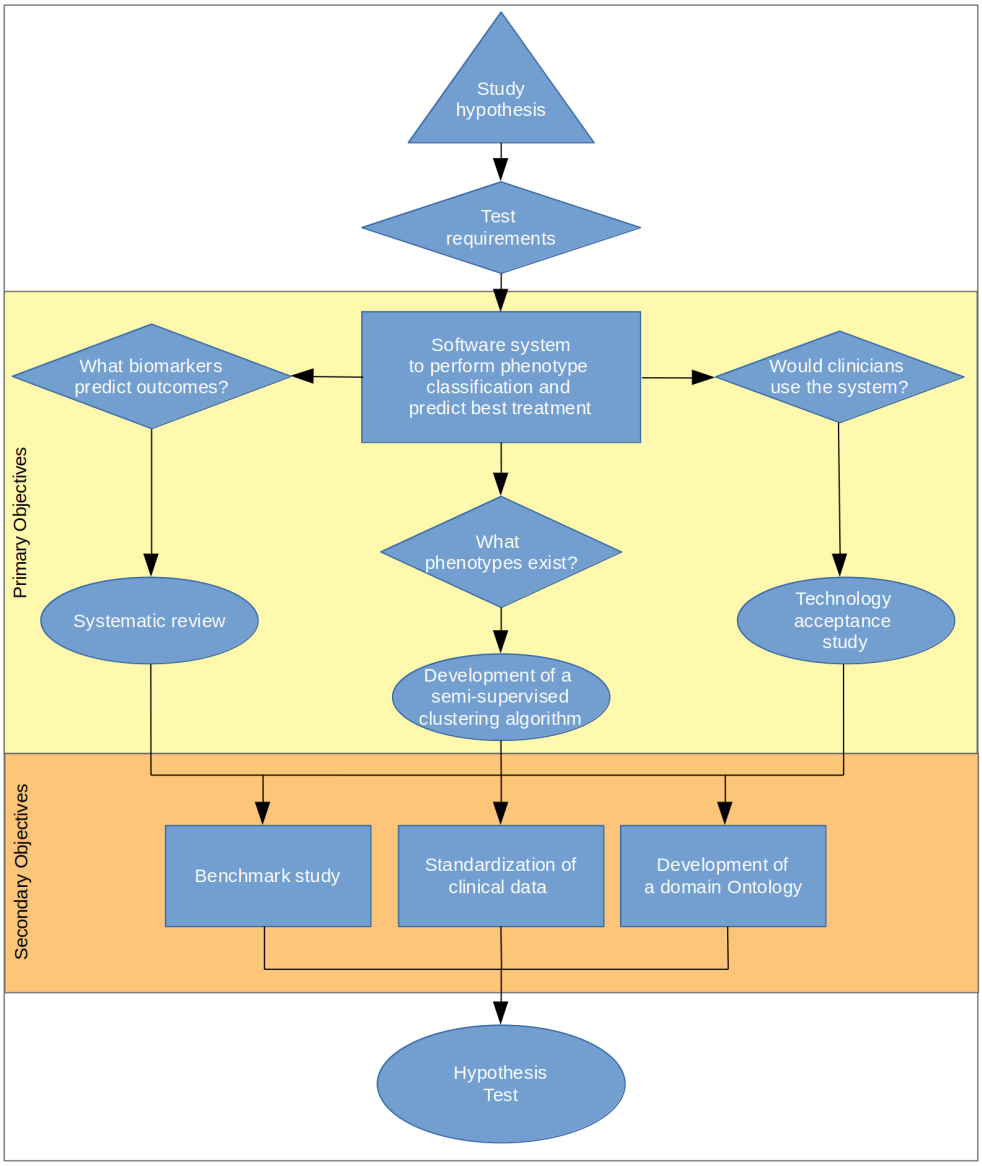

Figure 3.1: Conceptual view of the study hypothesis and objectives

Chapter 4

Materials and Methods

This chapter describes the materials and methods used within the scope of this thesis. Section 4.1 describes a systematic literature review to identify biomarkers predictive of outcomes in AF patients. Section 4.2 introduces survival analysis and describes the implementation and evaluation of a novel semi-supervised clustering algorithm. Section 4.3 presents the materials and methods used for the identification of AF phenotypes, including the used database, its processing, and the algorithms developed for the analysis. Finally, section 4.4 outlines the the methods for the usability study.

4.1 Systematic Literature Review

Beyond the biomarkers currently used for patient stratification, literature provides insight into recently discovered biomarkers. These biomarkers are identified through a systematic literature review, which is presented in this section.

4.1.1 Search Strategy

Relevant records are identified through a search in two electronic databases: PubMed and Cochrane. Potentially relevant meta-analyses, systematic reviews, and random controlled trials from January 1st 2016 to 31st December 2021 are identified using the following query:

```
atrial fibrillation AND
(biomarkers OR markers OR predictors OR predicts)
```

4.1.2 Inclusion and Exclusion Criteria

Article screening is performed based on three inclusion criteria:

1. **The article focuses on atrial fibrillation.** Literature review will be restricted to a main focus on AF. Articles mentioning AF i.e. as a possible consequence of another disease, do not meet inclusion criteria.

2. **Relates a biomarkers to relevant patient outcomes.** Articles will be included if an association between a biomedical entity and a relevant patient outcome is made. Outcomes considered relevant are the onset of AF, efficacy of an anti-arrhythmic drug, recurrence of AF after an intervention, stroke, major bleeding events, or death.

3. **Study is performed in a hospital setting.** Literature search is restricted to studies that evaluate biomarkers in a hospital setting. Wearables such as holter monitors, or smartphone technologies do not meet inclusion criteria.

The final assessment is performed based on the following three exclusion criteria.

1. **The study focuses on post-treatment biomarkers.** Studies that evaluate biomarkers and their changes post-treatment are not included in the synthesis.

2. **Study uses exclusively common risk scores.** For the scope of identifying potential biomarkers, studies that exclusively use common risk scores are excluded. The biomarkers used in common risk scores will be analyzed separately in section 2.1.

3. **Biomarkers are collected invasively.** Biomarkers are often measured invasively during the performance of a procedure. Such studies are ex-

cluded, since they do not aid in the identification of the best treatment option.

4.1.3 Selection Process

Following a removal of duplicate records, unique records are screened using record titles and abstracts according to the inclusion criteria. Records that meet the inclusion criteria are further assessed for eligibility using the exclusion criteria. The remaining records are included in the quantitative analysis. The complete record text is screened to identify biomarkers used for outcome prediction. Biomarkers are grouped according to semantic categories, and the number of publications using each biomarker is identified.

4.2 Semi-Supervised Clustering with Survival Data

As previously outlined in section 2.3, the use of unsupervised clustering algorithms produce sub-phenotypes with little clinical utility. To provide a biologically meaningful and clinically relevant sub-phenotypes, previous works have proposed the use of semi-supervised clustering algorithms, which include survival data in the clustering process.

Several approaches to develop SSC have been proposed by literature [174]:

- **Constraint-based SSC** in which either a) the resulting clusters are forced to satisfy constraints, b) a penalty factor is added to the objective function penalizing cluster assignments that violate the constraint conditions, or c) constraints are given by independent class labels which are used to initialize cluster centers and require cluster centers to satisfy the given constraints. While such constraints are most commonly given by must-link and cannot-link formulations between instances, intra-cluster and inter-cluster constraints have also been proposed [42].
- **Distance-based SSC** where, during data preprocessing, the similarity measure between samples is modified to accommodate not only the distance in covariate space, but also a distance measure based on additional information. Clustering according to the resulting similarity measure yields clusters that aim to satisfy both requirements. While the additional distance measure may be a continuous function, literature also introduced the translation of constraints into distance measures [20].

- **Constraint and Distance-based SSC** which combines methods from both of the above.

Based on these proposed approaches, two semi-supervised clustering algorithms are developed. The first, S-HAC, extends the popular hierarchical agglomerative clustering algorithm by incorporating survival data in the form of distance computation and agglomeration constraints. Secondly, KMeans is extended to incorporate survival data, providing LSS-KMeans. Both algorithms are outlined in the following sections, which are succeeded by a the description of a benchmark study which compares the performance of both algorithms to previously proposed semi-supervised clustering algorithms.

4.2.1 Background

Survival Analysis

Survival analysis refers to a group of data analysis methods which measure the risk of an event over time. Common uses of survival analysis include the analysis of the failure of electrical components, prediction of customer churn, or patient outcomes in clinical studies.

Survival analysis uses several concepts that determine the application of its tools, and define how analyses are performed:

An observation is defined by a **birth**, which defines when the observation in question becomes relevant to the analysis. Depending on the analysis performed, the birth may refer to the start of an electrical stress test in a component, the time a customer signs up for a service, or the time at which a patient is diagnosed with a disease or a treatment is administered. Within the scope of this study a patient's birth is considered to be the time at which a patient is first observed to have atrial fibrillation, which is derived from periodic recording performed by the nursing staff.

The concept of **death** refers to the time at which the event of interest is observed. It may refer to the actual death of an individual, but may also constitute the failure of an electrical component, a customer's cancellation of a service, or the positive effect of a medical treatment. In the analysis performed in this work, several events are considered of interest, and are as follows:

- **Rhythm Control -** defined as the reversion of SR following a treatment administration.

- **Rate Control -** defined as the reduction of heart rate below 100 beats per minute following a treatment administration.
- **Mortality -** as captured from the Social Security Administration Death Master File.

Censorship occurs when the observation of a death is no longer possible due to another event prohibiting its observation. An example of censorship is the loss to follow-up in a medical study, or the intentional termination of a study. Within the scope of this work, censorship is defined depending on the outcome that is being observed:

Outcome	Censoring Events
Rhythm Control	Mortality Discharge
Rate Control	Rhythm Control Mortality Discharge
Mortality	Discharge

Table 4.1: Censorship events for evaluated outcomes

Covariates are variables that describe the individuals under investigation. Covariates of interest to the outcome of atrial fibrillation patients have been covered by risk scores, and identified from recent literature.

Survival functions describe the survival probability in a study population throughout time. In particular, they describe the probability that *death* occurs later than a specified time t. The survival function is defined as:

$$S(t) = Pr(T > t) \tag{4.1}$$

The **hazard function** describes the rate of *death* during a time interval [t, t+dt], with the condition of survival until time t. It is derived from the survival function, and is defined as:

$$\lambda(t) = \lim_{dt \to 0} \frac{Pr(t \leq T \leq t + dt)}{dt \times S(t)} = \frac{S'(t)}{S(t)} \tag{4.2}$$

Semi-Supervised Clustering with Survival Data

Clustering with survival data is a relatively under-explored field [13, 32, 141]. Two general approaches to semi-supervised clustering can be observed: The extension of unsupervised clustering methods, such as introduced in the previous section, to include survival data, and the modification of supervised methods to produce clusters. The former approach was primarily taken in initial advances, which used survival-based algorithms such as Cox proportional hazard models [37] for covariate selection [14], or covariate weighting [72]. The subset of covariates, or weighted covariates were subsequently used in unsupervised clustering algorithms. Recent works are primarily based on the use of high-performance supervised algorithms, such as artificial neural networks [32, 128, 141], random survival forests [181], or genetic algorithms [75] to produce cluster assignments. While such methods improve predictive accuracy, they introduce significant computational overhead, and reduce algorithmic transparency and interpretability [170].

To date, the extension of unsupervised clustering algorithms to incorporate survival data is limited to extensions of the KMeans algorithm. Kmeans is a partitioning clustering algorithm, and is one of the most popular clustering methods. It is applied to datasets where observations are defined by a set of m covariates, which constitute the m-dimensional space in which observations are represented as points. KMeans is initialized by randomly selecting k observations as initial cluster centroids, which are iteratively re-assigned until the algorithm converges and a steady state is reached. The algorithm can be summarized as follows:

1. Initialize k cluster centroids randomly.
2. Assign each observation to the nearest centroid.
3. Calculate the mean of each cluster.
4. Update the centroids to be the means of the corresponding clusters.
5. Repeat steps 2-4 until convergence is reached.

While KMeans operates exclusively on observations' covariates, Bair et. al. have proposed to leverage associated survival data to obtain more predictive clusters [14]. In their proposed approach, the authors utilized univariate Cox proportional hazards models to quantify the correlation between covariates and the outcome of interest, which allowed them to associate each covariate with a Cox score. Cross-validation was then employed to determine the optimal

threshold for selecting which covariates to include in the KMeans algorithm. Only solutions that demonstrated significant differences in survival across the resulting clusters based on a log-rank test were accepted. Once the best threshold was identified, the entire dataset was clustered using KMeans, including only the covariates with Cox scores above a determined threshold.

This method was subsequently refined by Gaynor et al. [72], who introduced the concept of *supervised sparse clustering*. Building upon the approach proposed by Bair, Gaynor et. al. computed Cox scores for the available covariates. Instead of identifying a threshold for including said covariates, Gaynor assigned weights to the covariates based on their Cox scores. The weighted covariates were then used in the unsupervised clustering algorithm, using the *sparse clustering* algorithm proposed by Witten et al. [230]. The *sparse clustering* algorithm is an unsupervised extension of KMeans, which maximizes covariate differences among clusters using a secondary weighting. In essence, Gaynor et. al. weigh covariates based on their Cox score, and, subsequently, perform KMeans clustering performing a second weighting to maximize covariate differences between clusters.

Numerous modifications of KMeans have been proposed, but the inclusion of survival data into KMeans remains limited to the use of Cox proportional hazard models for feature selection or feature weighting. While such approaches are very scalable, and may provide more predictive clusters than the unsupervised version of KMeans, they suffer from several limitations inherited from their components. For example, the use of a Cox proportional hazards model assumes, among others, that covariates have a linear relationship with the risk of an event. This relationship may, however, be more complex, and critical covariates may be excluded from the clustering process, simply because the assumptions of a Cox proportional hazards model do not hold true. Similarly, the KMeans algorithm assumes that clusters are spherical, and of similar sizes, and will often fail to produce accurate clusters when these assumptions do not hold.

4.2.2 Survival Hierarchical Agglomerative Clustering (S-HAC)

Note: Parts of this section, as well as subsequent sections have previously been published, and some passages have significant overlap with the following sources: [123].

A novel constraint and distance-based SSC algorithm is proposed, which is described with the aid of a simulation:

Considering a dataset comprised of 1,250 observations, which are described by two covariates. The covariates are randomly sampled from a bi-variate Gaussian distribution with $\mu = (0.0, 0.0)$ and $\sigma = (1.0, 1.0)$. The dataset is split into three distinct groups along the x-axis, with each group being characterized by a different risk of death and risk of censoring. The three groups and their corresponding survival curves are shown in figure 4.1A and B.

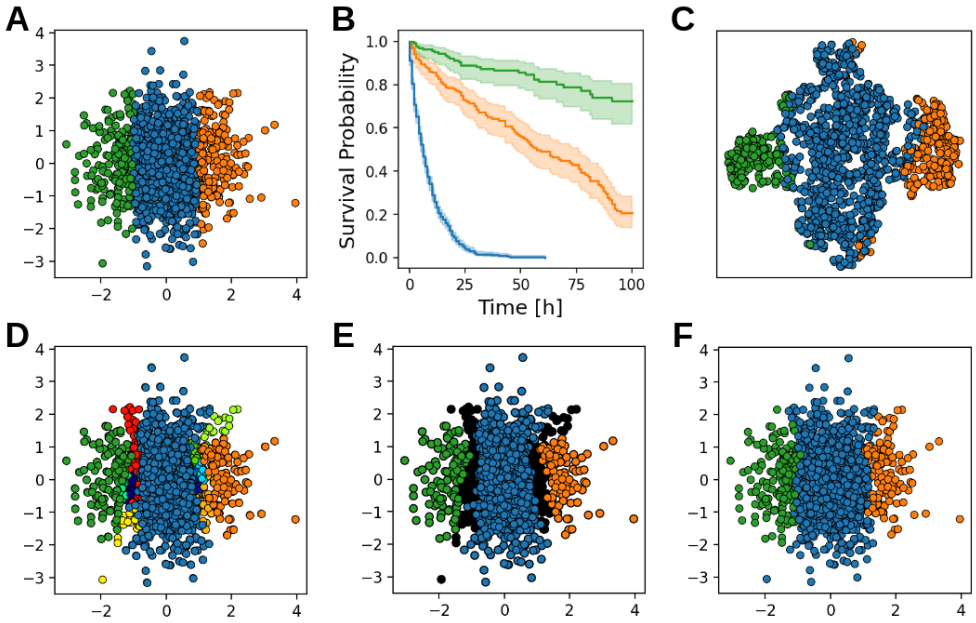

Figure 4.1: A) Simulated bi-variate Gaussian distribution of observations belonging to three different groups. B) Corresponding Kaplan-Meier curves for each phenotype. C) T-distributed stochastic neighbor embedding of the hybrid pairwise distance matrix. D) Clusters identified from the pre-clustering stage. E) Core clusters identified, with smaller clusters dissolved into singletons (black). F) Final clustering outcome [123].

The novel algorithm, called S-HAC, performs the following steps to identify coherent and predictive clusters:

Step 1 - Distance Matrix Computation: The dissimilarity between observations is described using a distance matrix, which is derived from both, the observations' covariates, as well as their survival times and event indicators. The expected survival of each observation is derived from the survival data form its m nearest neighbors. For each observation

i survival times and event indicators $\{(t_{i,1}, e_{i,1}), (t_{i,2}, e_{i,2}), ..., (t_{i,m}, e_{i,m})\}$ are aggregated. Given two such sets, two observations i and j, the survival distance between observations is computed using the log-rank test statistic $Z_{i,j}$. In order to maintain similarities in the covariate domain, a covariate distance is computed based on observations' covariates. This may be performed using Euclidean distance, Manhattan distance, or any other commonly used distance metric. The final distance matrix is computed as a weighted sum of survival distance $Z_{i,j}$, and covariate distance $D_{i,j}$, and is formally expressed as:

$$H_{i,j} = \alpha * Z_{i,j} + (1 - \alpha) * D_{i,j} \tag{4.3}$$

The impact of incorporating the survival distance into the distance matrix computation is illustrated using a t-distributed stochastic neighbor embedding (T-SNE) [139] of the resulting distance matrix (figure 4.1C). Instead of a 2-dimensional Gaussian distribution, the subjects now form three high-density clusters corresponding to the simulated groups.

Step 2 - Pre-Clustering: Hierarchical agglomeration is performed using the hybrid distance matrix H. During this process, any two clusters proposed for merging by the dendrogram are compared in terms of their survival using the log-rank test. An agglomeration of two clusters is rejected if the log-rank test indicates significant difference in their survival distributions. The intermediate result of this process is depicted in figure 4.1D.

Step 3 - Core Cluster Identification: The previous clustering solution demonstrates an under-agglomeration with many groups stemming from noise. A *minimum cluster size* is introduced to alleviate this issue. Small clusters that do not meet the minimum size requirement, are subsequently dissolved into singletons (figure 4.1E).

Step 4 - Post-Clustering: The agglomeration is continued, while preventing the agglomeration of two singleton clusters using a cannot-link constraint. Again, any two clusters indicated for merging are evaluated using the log-rank test, and are only merged if no significant difference in the survival distributions exists. The final clusters are shown in figure 4.1F.

4.2.3 Locally Smoothed Survival KMeans

Current extensions of unsupervised clustering algorithms, which make use of survival data are limited to the use of the KMeans algorithm. Its extensions make use of Cox proportional hazards models to determine covariate importance, and consequently employ the KMeans algorithm on a subset of covariates, or weigh covariates based on their importance. Similarly to S-HAC, this survival information may be used directly in the KMeans algorithm:

1. Compute q quantiles of the observation times in the dataset, creating $q+1$ intervals.
2. For each observation in the dataset, identify its n nearest neighbors.
3. Aggregate durations $(T_1, T_2, ..., T_n)$ and event indicators $(E_1, E_2, ..., E_n)$ from said neighbors.
4. Compute the cumulative hazard $\hat{H}_i$ present in each interval i, based on the time and event indicators from step 3, to produce a survival vector S.

$$\hat{H}_i = \int_{q_i}^{q_{i+1}} H(t)dt \tag{4.4}$$

$$S = (\hat{H}_0, \hat{H}_1, ...\hat{H}_{q+1}) \tag{4.5}$$

5. Given a covariate vector X, and survival vector S for each observation, scale both sets of vectors such that the average vector is of unit length.
6. Multiply the survival vectors by a factor β to control for the importance of the survival vector in the clustering.
7. Concatenate vectors X and S.
8. Run the KMeans algorithm on the set of concatenated vectors.

Several comments on this method are in order. Unlike previous approaches, which use a subset of covariates, or weighted covariates as the input of KMeans, a vector S is created, which represents the localized hazard profile for each observation, and concatenated it with the corresponding covariate vector, to use the concatenated vector for KMeans clustering. This procedure has several motivations. Firstly, by deriving the survival vector from n nearest neighbors,

noise in the time and event indicators of single observations is attenuated. Secondly, it ensures that the hazard profile for a specific observation remains representative of similar observations, which could not be guaranteed with a global approach. Lastly, the approach directly provides the KMeans algorithm with information regarding observations' survival, which, just like covariates, may drive cluster formation by representing observations with similar survival profiles as more similar in the clustering domain.

4.2.4 Simulation Study

To visualize the behaviors of S-HAC and LSS-K-Means, and compare it to alternative algorithms in a controlled fashion, a simulation study is performed using three simulated datasets, which are shown in figure 4.2:

Sim-a) A collection of 1,250 observations with two covariates distributed in a 2-dimensional Gaussian distribution $\mu = (0.0, 0.0)$ and $\sigma = (1.0, 1.0)$. Three survival distributions are created using cutoff points across the horizontal axis ($x < -1.0, -1 \leq x < 1, x \geq 1$), corresponding to three different survival distributions of constant hazard ($H_0 = 0.1, H_1 = 0.0125, H_2 = 0.004$) and a constant censoring rate of 0.01. Note that although the dataset consists of three distinct groups, identifying them is not a trivial task due to the covariates being distributed in a Gaussian distribution, which poses a challenge that unsupervised clustering algorithms would be unable to overcome.

Sim-b) A simulated distribution of two concentric circles from the Scikit-learn library [168], referred to as *noisy circles*, with a total of 1,250 observations. For the two circles, we simulate survival data with hazards of 0.02 and 0.05, and a censoring rate of 0.01. Note that this dataset poses a significant challenge to algorithms extending KMeans, because there is no linear decision-boundary that would effectively separate the two clusters.

Sim-c) A collection of 1,250 observations with two covariates distributed in a 2-dimension uniform distribution ($-3 < x < 3, -3 < y < 3$). The observations are split into two groups according to a sine wave ($y = 2sin(2x)$), with observations having hazards of 0.1 and 0.004 above and below the sine wave, respectively. A constant censoring rate is set at 0.01. This dataset poses another challenge to semi-supervised clustering algorithms, because it does not only lack structure in the covariate domain, but it requires a particularly precise, non-linear, decision boundary to accurately classify observations into the correct clusters.

All simulations are censored at $T = 100$.

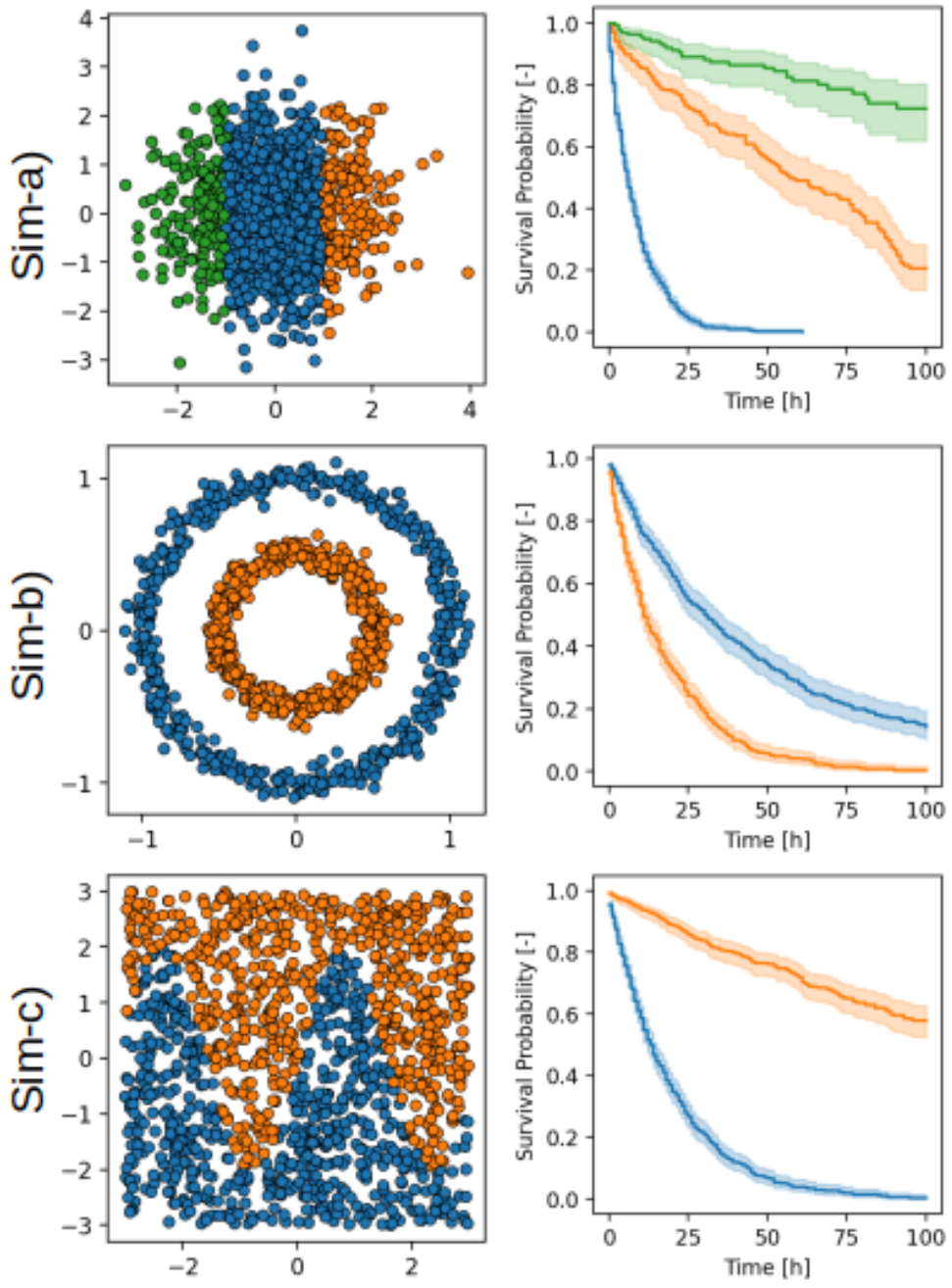

Figure 4.2: Simulated examples. Covariate distributions in the left column with the corresponding Kaplan Meier curves in the right column.

Table 4.2: Hyperparameters used in hyperparameter sweep.

Hyperparameter	Tested Values
k*	2, 3, 4, 5, 6, 7, 8
No. of neighbors	30, 40, 50, 60, 80
No. of quantiles	5, 10
β	1.0, 1.5, 2, 2.5

* Only used in the benchmark study. In the simulation study, k was set to the number of simulated clusters.

The ability of LSS-K-Means to identify the simulated clusters is compared to KMeans, as well as the previously proposed semi-supervised clustering algorithms extending KMeans: SSC-Bair and SSC-Gaynor. For KMeans, SSC-Bair and SSC-Gaynor we select k such that it matches the number of simulated clusters. For LSS-K-Means we perform a hyperparameter sweep shown in table 4.2, while setting k to the number of simulated clusters. The resulting cluster solutions are visually evaluated, and numerically assessed using the classification accuracy and the adjusted Rand index (ARI).

4.2.5 Benchmark Study

S-HAC and LSS-K-Means are quantitatively evaluated by comparing them with unsupervised clustering algorithms, and previously proposed semi-supervised clustering algorithms using survival data. Four openly accessible datasets are used: i) FLCHAIN, comprised of data from an investigation of the effect of non-clonal serum immuniglobilin free light chains on patient survival [49], b) SUPPORT, a dataset from a study evaluating a prognostic model for ICU patients [116], c) GBSG2, a dataset previously used to study the effects of hormone treatment on breast cancer recurrence [191], and d) METABRIC, a dataset created to predict the survival of breast cancer patients from clinical data, gene expression data, and mutations [39]. The datasets' characteristics are shown in table 4.3.

Table 4.3: Dataset Characteristics.

	FLCHAIN	SUPPORT	GBSG2	METABRIC
No. of Subjects	7,894	9,105	686	1,904
Events (%)	27.5	68.1	43.6	42.1
No. of Covariates	26	59	10	560
Missing (%)	2.1	12.6	0.0	$<$0.01
T_{max}	5,215	2,029	2,659	355

The benchmark study evaluates the performance of the newly developed S-HAC and LSS-K-Means algorithms to vanilla KMeans, vanilla HAC, SSC-Bair [14], and SSC-Gaynor [72]. The benchmark is performed using 10-fold cross validation, and the algorithms are evaluated based on the log-rank statistics and concordance indices. With each algorithm a clustering is performed using the training set, and survival distributions in each cluster are estimated using Kaplan-Meier estimators. The Kaplan-Meier estimators are used to estimate the median survival time, which is used as a risk score in the computation of concordance indices. The empirical distribution of observations in the validation sets are used to compute the log-rank statistic. For SSC-Bair cluster predictions in the validation set are performed using a nearest centroid method, as suggested by the authors [14]. For the remaining algorithms, cluster assignments are derived from the nearest neighbor in the training set. In terms of the tested hyperparameters, the algorithms HAC, SSC-Bair, and SSC-Gaynor are evaluated using k=$\{2, 3, 4, 5, 6, 7, 8\}$, and the metrics for best performing k are reported. Given S-HAC's multiple hyperparameters, a hyperparameter sweep is performed, which is shown in table 4.4.

Table 4.4: Hyperparameters tested for the benchmark study

Hyperparameter	Tested Values
No. of Neighbors	30, 60, 90, 120
Min. Cluster Size	30, 50, 70, 90
Significance Level	0.05, 0.10, 0.20
α	0.5

4.3 Procedure to Identify Atrial Fibrillation Phenotypes

Note: Parts of this section, as well as subsequent sections have previously been published, and some passages have significant overlap with the following sources: [122].

4.3.1 Study Data Base

AF phenotypes are identified from the The Medical Information Mart for Intensive Care (MIMIC-III) database. MIMIC-III is a single-center database containing information from critical care units in the Beth Israel Deaconess Medical Center in Boston, Massachusetts, collected in the years 2008-2019. The dataset is comprised of 49,785 distinct ICU admissions made by 38,597 individual adult patients (median age 65.8 years; 55.9% male) [102]. While a newer database, MIMIC-IV [101], exists, it has been dismissed due to the absence of critical biomarkers.

In MIMIC-III, patients are associated with medical diagnoses in the form of ICD-9 codes, as well as time-stamped laboratory measurements and observations, and demographics. Medical observations, collected from dedicated critical care information systems (CareVue [Philips Healthcare, Andover, USA], MetaVision [iMDsoft, Israel]) provide time-stamped laboratory measurements, and vital signs. Time-stamped, de-identified clinical notes, such as ECG reports, echo reports, and discharge summaries are further present to enable natural language processing to capture certain observations not present in tabular format.

Database Standardization

Observational healthcare data is created with different goals in mind. These goals may range from enabling research, to facilitating payments, or to enable healthcare professionals through electronic health records. Biomedical data therefore appears in different formats and uses different types of vocabularies to represent semantically similar concepts introducing a high degree of complexity to biomedical research. Examples thereof are administered drugs, which are coded in the form of RxNorm codes, with individual codes representing a specific medication including the dosage and producer. Such codes may be readily mapped to other codes corresponding to the present ingredients using the appropriate vocabulary. Different drug administrations, all containing the

same ingredient in different dosages and routes of administration may therefore be identified.

To enable such semantic operations, the Observational Health Data Sciences and Informatics Program [1] proposes the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), portrayed in figure 4.3. The database used in this work [123, 122] is standardized to the OMOP CDM using the methods provided by [163] and stored in a PostgreSQL database, to ensure that the methods developed in this work can be applied to a wide range of databases.

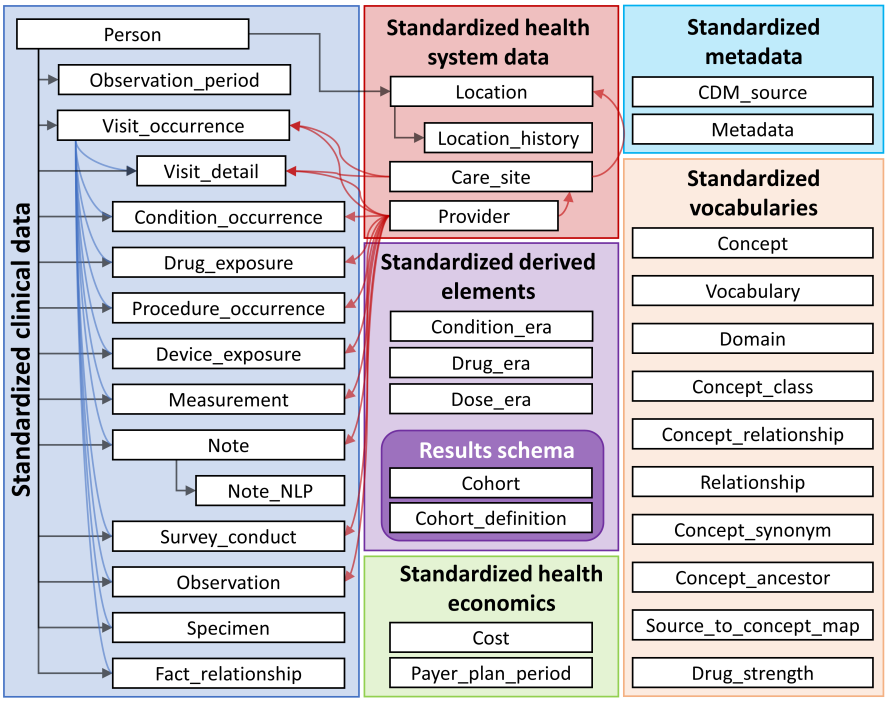

Figure 4.3: Overview of tables in the OMOP CDM [160]

[1] Observational Health Data Sciences and Informatics, https://www.ohdsi.org/

4.3.2 Knowledge Representation

Medical knowledge is often represented in the form of hierarchies. Examples thereof are medical terminologies and classification systems such as SNOMED-CT or International Classification of Diseases (ICD), which both arrange concepts in semantically meaningful subclass-superclass relationships. Such semantic relationships are intrinsic to medical concepts and allow for meaningful data processing as well as interpretable software systems. The interest in providing increasingly sophisticated methods of knowledge representation which provide capabilities of encoding hierarchical structures, concept descriptions, and semantic interoperability between different systems has led to the introduction of knowledge graphs originating from the World Wide Web community [155].

For the scope of this work, a domain ontology is developed to provide convenient and transparent means of capturing relevant concepts from databases, and to enable the inference of existing concepts which are implicitly referenced by corresponding subclasses.

Protégé [154], an ontology editor and framework for building intelligent systems, is used to generate a hierarchical structure of SNOMED-CT, LOINC, and RxNorm concepts representing diagnoses, measurements, treatment options, and outcomes which are identified by the systematic literature review in section 5.1.2 and are present in the database.

The ontology is evaluated using the Ontology Pitfall Scanner [173] to ensure structural and functional integrity, and universal usability.

4.3.3 Study Design

Study Population

The study population is comprised of a cohort of patients with a diagnosis of atrial fibrillation. Patients may be diagnosed in an outpatient or inpatient setting and may be newly diagnosed with atrial fibrillation or have a prior atrial fibrillation diagnosis.

A single patient cohort is defined according to the following inclusion and exclusion criteria.

Inclusion Criteria

- Patients with a diagnostic code indicating atrial fibrillation.
- Patients above the age of 18 are included.

Exclusion Criteria

- Patients with an ICU stay shorter than 24 hours are excluded. [17]

Patient Covariates

As presented in chapter 2, patient outcomes are dependent on biomarkers such as comorbidities, concentrations of different compounds in the blood serum, electrophysiological or imaging measurements, and others. Such markers, while present in standardized databases, lack structure to enable processing by data mining algorithms. To enable efficient processing, patients are represented in the form of a vector consisting of the descriptive variables captured by means of the domain ontology. MIMIC-III was screened for the variables identified in the systematic review in section 4.1. The earliest available record for each variable is used, if more than one value is available. All continuous variables are transformed into z-scores for analysis.

Outcomes

Two primary outcomes are defined: (i) the conversion of AF to SR, and (ii) the achievement of rate control, which is defined as a heart rate below 100 beats per minute (BPM) [70]. In the employed dataset, heart rates and heart rhythms were recorded at regular intervals, and a previous study has confirmed the recordings' accuracy and precision to within 1 hour [46]. For the scope of this analysis, a registered rhythm is assumed to be maintained until a different rhythm is recorded. The secondary outcome of interest is in-hospital mortality. The primary and secondary outcomes are censored at 24 hours and 30 days, respectively [17].

Treatment Group Assignments

Patients are assigned to treatment groups depending on the first administration of a treatment during an AF episode. The treatment groups considered in this study are beta blockers (BBs), potassium channel blockers (PCBs), calcium channel blockers (CCBs), and magnesium sulphate (MgS). The corresponding drug ingredients are shown in table 4.5. Outcomes are evaluated in an intention-to-treat fashion [17].

Table 4.5: Treatment groups with corresponding drug ingredients.

Treatment Group	Drug Ingredients
Beta Blockers	Acebutolol, Esmolol, Labetalol, Metoprolol, Propranolol
Potassium Channel Blockers	Amiodarone, Dofetilide, Dronedarone, Ibutilide, Sotalol
Calcium Channel Blockers	Dilatiazem, Verapamil
Magnesium Sulphate	Magnesium Sulphate

Missing Data Imputation

Biomedical data is often sparse and has missing values for variables that are of interest to the performed analysis. In practice, patients with missing values may either be removed from analysis, or the missing values may be imputed. The removal of patients with missing values is applicable when the number of missing values is sufficiently small that the statistical power of the analysis is not negatively impacted. For particularly sparse datasets, missing value imputation is therefore the preferred method to enable analysis.

Data may be missing due to different types of data collection. Generally, there are three types of mechanisms [55]:

Missing completely at random - Describes the probability of a data entry being missing to be uniform across the entire dataset.

Missing at random - Data is missing, and its probability of being missing is dependent on the already observed data. An example of *Missing at random* is a missing value for a serum biomarker that is usually used to assess myocardial infarction. If there is no indication of a myocardial

infarction, then there is an increased probability that the serum biomarker has not been measured.

Missing not at random - Data is missing due to information that is not available. An example thereof would be a sub-phenotype of patients exhibiting an immediately lethal stroke, and a biomarker not being obtained due to patient expiry before the arrival at the hospital.

A common method for imputation are using the mean value, under the assumption that the imputed variable is not correlated with the remaining data, as is the case when data is *missing completely at random*. More sophisticated methods, such as k nearest neighbors imputation [63], operate on the assumption that variables are, in fact, correlated and impute missing values based on the values of patients with similar properties. Such approach is appropriate where data is *missing at random* [55].

Within the scope of this work, it is assumed that data collection is driven by clinicians' assessment of the added value of collecting said data. For example, a certain test will only be performed, to confirm or rule out a specific diagnosis. This assumption implies that data is *missing at random*. Further, any imputed value carries a degree of uncertainty, which must be accounted for in clinical environments. Multiple imputation [183] is therefore performed. 60 imputed datasets are created using linear regression models and chained equations [25]. The original dataset is resampled with replacement 60 times, and linear regression models are derived from each bootstrapped dataset. These models are used to impute the original dataset for a total of 60 imputed datasets. Given a fraction of missing information of 7.42%, the number of imputed datasets satisfies Bodner's rule [21], which would require at least 8 imputations.

Inverse Probability of Treatment Weighting (IPTW)

In contrast to randomized controlled studies, observational studies are limited by selection bias. Patients may receive different treatments for a variety of reasons, such as underlying pathophysiological conditions. Patients with a worse prognosis may be administered a different drug from patients with a positive prognosis, leading to a biased estimation of the drug's treatment effect. This makes the assessment of treatment effects a non-trivial endeavor.

Inverse probability of treatment weighting (IPTW) is a statistical method that adjusts for this selection bias by weighing each patient with the inverse probability of said patient receiving a treatment. IPTW therefore balances treatment

groups to provide adjusted treatment effects with reduced confounding [180]. To adjust for confounding and minimize selection bias, IPTW is performed for each imputed dataset using the toolkit provided by Ridgeway [179].

Patient Clusters

S-HAC is employed to identify phenotypes in the AF cohort. Given three different outcomes and four different treatments being analyzed, pairwise distance matrices are computed for each combination of treatments and outcomes, resulting in a total of 12 pairwise distance matrices. The largest pairwise distance is used for clustering. During agglomeration, the stopping criterion is evaluated for each treatment and outcome combination. To determine the best set of hyperparameters, a parameter sweep using 10-fold cross-validation is performed as outlined in table 4.6. The final clusters are obtained by performing a clustering using the set of hyperparameters with the highest mean concordance index. Clusters are visualized using a t-distributed stochastic neighbor embedding (T-SNE) incorporating the cluster assignments as variables into the T-SNE algorithm to encourage cluster formation.

Table 4.6: Hyperparameters tested for the case study

Hyperparameter	**Tested Values**
No. of Neighbors	20, 30, 45, 60, 90, 120, 160, 240
Min. Cluster Size	20, 30, 50, 80, 160
Significance Level	0.01, 0.05, 0.10
α	0.5

Treatment Effects

Average treatment effects (ATEs) are approximated using weighted exponential survival models to provide constant event rates. The weights for these models are derived from the previous IPTW. Models are repeatedly fitted to each imputed dataset 100 times using Bayesian bootstrapping [184]. This approach allows for at total of 6000 event rate estimates per ATE, for which the mode and 95% highest density intervals are reported. ATEs are computed for the complete cohort, and for individual clusters. Results are reported as hourly rates for the primary outcomes, and as daily rates for the secondary outcome. Finally, differences in ATEs are assessed using Bayes factors (BFs) [110].

4.4 User Centered Evaluation

Numerous authors have suggested that a clustering based phenotype classification system for AF patients could support clinical decision-making by either providing insights into the heterogeneity of the patient populations, or by providing clinical decision-support directly [95, 206, 217, 223]. Such a CDSS, like any other technology, must be accepted by its end-users, to be useful and improve productivity. To what extent users will accept such a technology, and what reasons contribute to the intention of its use can be determined through a user centered evaluation. Within the scope of this evaluation, the suitability of the previously developed methods to provide guidance in treatment selection in the form of a potential CDSS are evaluated with clinical decision-makers. This evaluation is performed using a visual mock-up of a CDSS, and an explanation of the possible functionality. Feedback from participants is collected and their responses are evaluated.

Different models have been proposed to determine technology acceptance, such as the technology acceptance model (UTAUT2), which measures individual acceptance in terms of end-users' intention to use a technology [43, 216]. Other works propose the evaluation of implementation success on an organizational level [130], or the degree to what extent a technology fits a given task that the technology aims to support, the task technology fit (TTF) [77].

4.4.1 Study Design

A mixed methods study is performed incorporating a qualitative part of interviews and open discussions, and a quantitative part with structural surveys. Qualitative approaches are valuable due to their open-ended nature, and may provide a rich description of complex and multifaceted problems [188].

The study assesses the suitability of the developed methods for treatment selection and their potential adoption in clinical practice. A theoretical implementation of a CDSS incorporating a a user interface with data input, a visualization of the patient population, the assigned patient phenotype, and treatment effects is evaluated. This approach not only investigates the potential acceptance of the developed methods in real-world medical settings but also examines the extent to which the developed methods are perceived as valuable. By doing so, it offers valuable insights for enhancing these methods and guiding future research and development. The objective of the study is to gain an understanding of the perspectives of clinical decision-makers through

a combination of qualitative data obtained from interviews and quantitative data collected through questionnaires.

4.4.2 Setting and Recruitment

The study is performed in collaboration with the University Hospital Freiburg (Germany). Study participants are recruited from the participating institution, with particular importance of including cardiology and critical care specialists. Participants are recruited from the Clinic for Cardiology and Angiology, and are eligible if they are (a) actively involved in the management of AF in the ICU, and (b) have a minimum seniority of physician assistants and c) three years of experience in treating patients with AF.

4.4.3 Data Collection

A total of three session are held with study participants, which are (1) an empathy session, a (2) task-technology-fit session, and (3) a technology-acceptance session.

Empathy Session

The goal of the empathy session is the identification of existing workflows and clinicians' perceptions thereof. The empathy session based on the concept of an empathy map [59, 60], and allows for the understanding of needs and characteristics of users [61]. The existence of internal guidelines and practices is discussed, and potential limitations and obstacles in decision-making are identified, and their consequences and potential solutions are evaluated. A total of two participants are asked questions in one hour long sessions. The sessions were performed with a semi-structured approach with questions initially being general, and increasingly becoming more specific. Interviews were audio-recorded, transcribed, and translated to English.

The following guiding questions are addressed in chronological order:

1. What are the goals of care when a patient develops AF in the ICU?
2. Do internal guidelines for treatment selection exist?
3. What are the strengths and weaknesses in current care?
4. Are any prediction models, or CDSS used for treatment selection?

5. Do you analyze your own data to gain insights into treatment effects?
6. How do you currently address the heterogeneity of the patient population?

Task-Technology-Fit Session

In the task-technology-fit session, the theoretical requirements for choosing the optimal treatment on an individual patient basis are identified. Further, the previously developed methods are discussed in terms of their applicability to the task of selecting treatment on an individual patient level.

The session is based on the theoretical framework of the TTF model, which is portrayed in figure 4.4. The task-technology-fit is defined as the ability of a technology to support a given task. It requires the matching of the technology's capabilities to the task requirements, and is a predictor of technology utilization and the associated performance increase. The fundamental assumption of TTF is that a technology will only be used if its function actually supports the activities performed by the end-user, and technology offering insufficient advantages will remain unused [77].

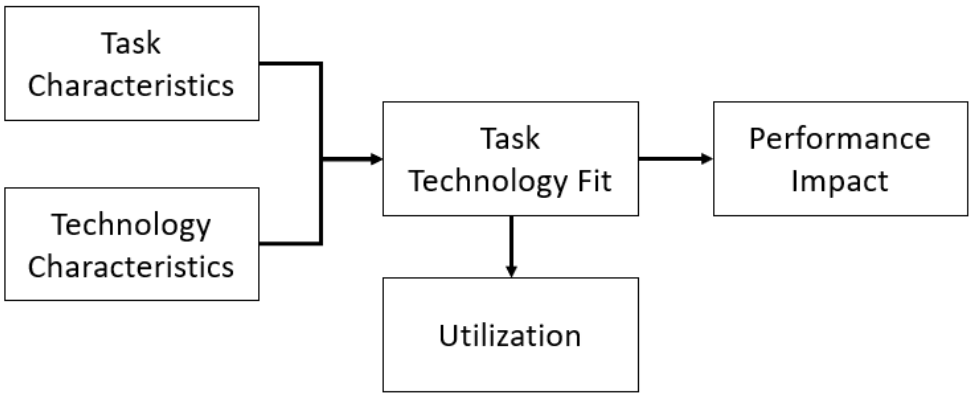

Figure 4.4: Task-Technology Fit Model

The session is initiated with a discussion about clinicians' perceived task characteristics with regards to selecting the best treatment option. A total of five participants are recruited to 30 minute long sessions, which are audio-recorded, transcribed, and translated to English. Participants are asked questions in a semi-structured approach with questions initially being general, and increas-

ingly becoming more specific. The following guiding questions are addressed in chronological order to identify the task characteristics:

1. What makes a treatment choice the "best"?
2. What information is necessary to identify the best treatment?
3. Do you feel like you can confidently make the best treatment selection?
4. To what extent is patient heterogeneity a challenge in treatment selection?

The above questions aim at evaluating several concepts of the CDSS design principles and the five rights of CDSS, which were introduced in section 2.4. Questions 1 and 2 evaluate the adherence to the first principle of CDSS design principles, which states that CDSS should mimic the cognitive process of clinical decision makers. Given that the "best" treatment must be selected out of a range of options, the definition of the "best" treatment is validated. Similarly, the definition of the best treatment options depends on a specific set of information. To what extent the used descriptive variables satisfy clinicians' requirements is evaluated in question 2. Questions 3 and 4 validate the necessity of employing a CDSS to support the task of selecting the best antiarrhythmic drug.

Following the discussion, participants are introduced to the developed methods with the use of a PowerPoint presentation. This Powerpoint presentation describes the data analysis performed in sections 4.3. A conceptualized CDSS is presented to the participants, which would incorporate a communication interface that facilitates the input of patient data, which is imputed using the developed linear regression models. The imputed data is used for a phenotype classification using the nearest neighbor in the development dataset, as described in section 4.2.5. Treatment effects are predicted based on the survival analysis models obtained in section 2.3, and are ranked according the expected utility, which is a predefined linear combination of treatment effects. Both, predicted treatment effects and ranked treatment options are presented to the user in the communication interface.

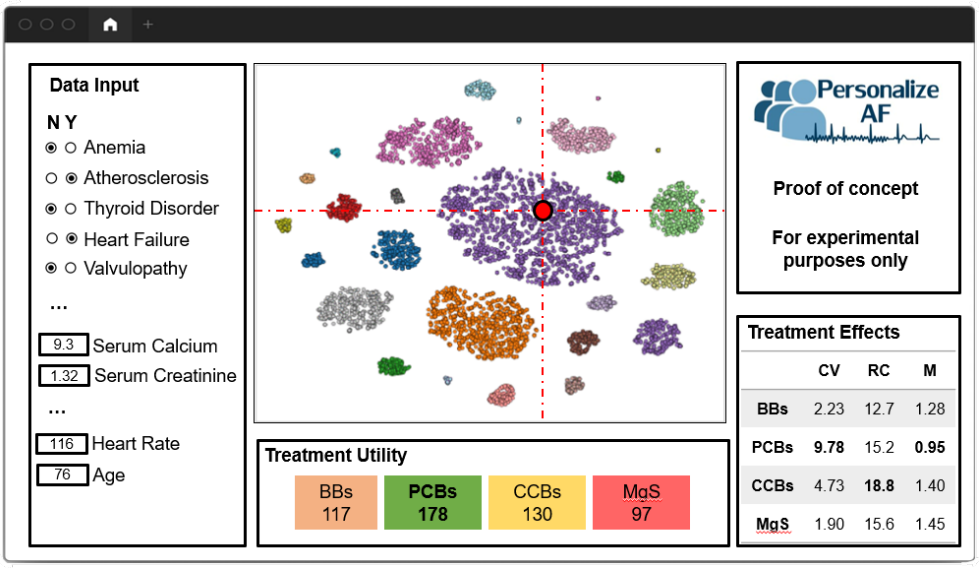

Figure 4.5: Conceptualized CDSS Front End

The applicability of the conceptualized CDSS is discussed, and the participants' feedback is recorded. Finally, participants are asked to complete an online questionnaire, which quantitatively evaluates the two fundamental components of the TTF model:

Task characteristics: It is assumed that clinical decision-makers lack the ability to robustly estimate treatment effects of available treatments, which was revealed by a lack of consensus in surveys among healthcare providers [225]. It is further assumed that the selection of the best treatment option is based on a maximization treatment utility with a simultaneous minimization of adverse effects. To what extent these assumptions capture clinical reality, and correspond to the actual task at hand is evaluated in the questionnaire.

Tool functionality: Based on the assumed task characteristics, tools were developed to identify AF sub-phenotypes with differing treatment responses. The treatment responses were approximated using exponential decay models to aid in the interpretability of results, and used to develop a decision-making algorithm that predicts the best treatment option. These three design decisions are evaluated using the respective questions.

The TTF questionnaire is evaluated using a 5-item Likert scale [132] indicative of clinicians' degree of agreement (strongly disagree, disagree, neutral, agree, strongly agree). The questionnaire is shown in figure 4.6.

Figure 4.6: TTF Questionnaire

Task-Technology Fit Questionnaire: Clinical Decision Support for Critical Care Atrial Fibrillation

Please tell us about yourself.

Specialty	O Cardiology	O Internal Med.	O Intensive Care	O _____
Seniority	O Attending	O Fellow	O Resident	O Student
Years of Practice	_____			
Gender	O Female	O Male	O Other	

Please indicate your level of agreement with the following statements.

	Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	The AF patient population is heterogeneous, and different sub-groups respond differently to available treatments.	O	O	O	O	O
2	The large number of available treatments coupled with patient heterogeneity makes it difficult to know what the "best" treatment is.	O	O	O	O	O
3	The break-down of the population into sub-phenotypes using the demonstrated algorithm appears to be appropriate.	O	O	O	O	O
4	The ranking of treatments using a utility function is useful.	O	O	O	O	O
5	The prediction of a patient's phenotype belongingness using its nearest neighbor is sufficiently transparent.	O	O	O	O	O
6	The approximation of treatment effects as hourly/daily rates provides an appropriate compromise between accuracy and interpretability.	O	O	O	O	O

The TTF questionnaire partially overlaps with the preceding open questions, and quantifies the degree of agreement with

Further, it evaluates several CDSS design choices: Question 3 evaluates to what extent participants agree with the division of the patient populations into clusters. While different methods of predicting treatment effects may be used, a cluster analysis was chosen, which was inspired by previous state of the art approaches. Question 4 evaluates the use of a utility function to rank treatments according to revealed preferences. Question 5 evaluates the use of a k nearest neighbors approach for predicting a patient's cluster belongingness, and, in particular, to what extent such approach is sufficiently transparent to the end user. Another important design choice is the use of exponential survival models for the approximation of treatment effects. This approximation provides easily interpretable results, but reduces the amount of available information. To what extent this is acceptable to the participants is evaluated in question 6.

Technology Acceptance Session

The technology acceptance session is held subsequently to the TTF session. It evaluates users' perceived attitude towards using the conceptualized CDSS. The usability is evaluated by means of the theoretical framework of the unified theory of acceptance and use of technology model (UTAUT2) [216]. The UTAUT2 model suggests several constructs that impact users' behavioral intention to use a technology, as well as the actual technology use: Performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit. These constructs may, in part, be influenced by a user's age, gender, and experience. The UTAUT2 model is shown in figure 4.7.

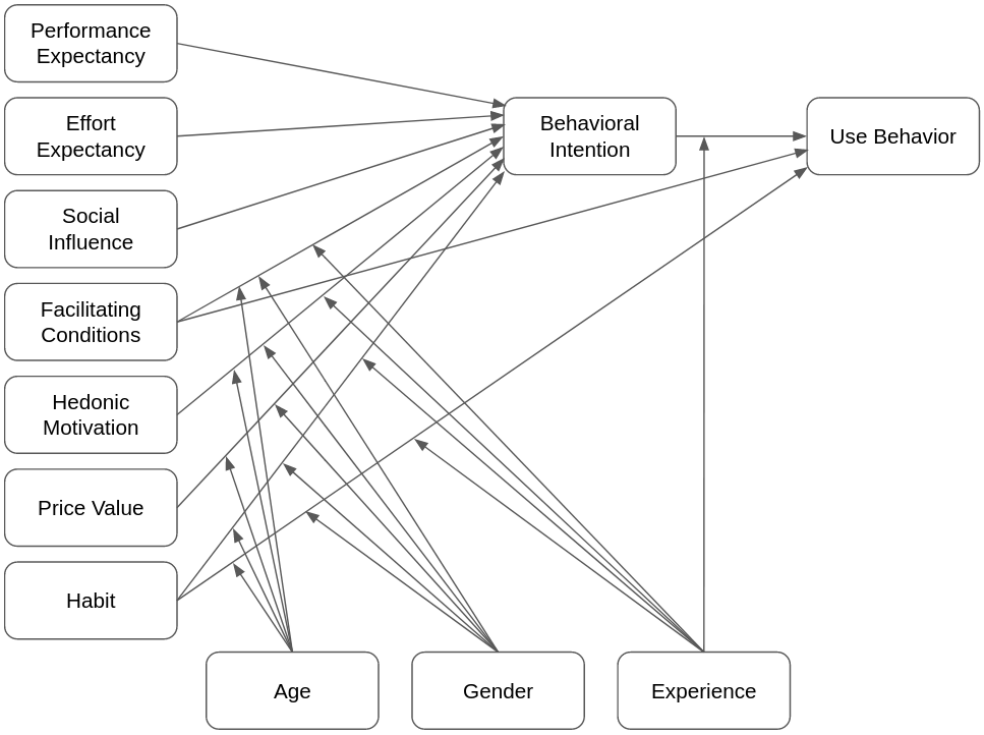

Figure 4.7: UTAUT2 Model

The technology acceptance session evaluates several key elements of the UTAUT2 model:

1. Performance expectancy: Performance expectancy predicts users' behavioral intention to use a technology based on the expected benefits and drawbacks of its use.

2. Effort expectancy: Users often form views about an the time and effort required for using a technology. The expected effort is an important factor that determines technology use.

3. Social influence: Defines the extent to which users believe that other stakeholders may perceive the use of the technology to be important.

4. Facilitating conditions: A user who operates in an environment that facilitates the use of a technology is more likely to use said technology.

5. Hedonic motivation: Relates to the expected pleasure or pain experienced when using a technology. For the evaluation of a conceptualized CDSS, this construct is not applicable.

The acceptance of the developed methods is discussed with a total of five participants in 30 minute long sessions, which are audio-recorded, transcribed, and translated. The sessions are guided by the following questions:

1. Do you see any obstacles or facilitators in the practical application of such CDSS?
2. Would such a CDSS improve treatment efficacy?
3. Would such CDSS be easy to use and require little effort?
4. Do current workflows allow for its integration into practice?
5. Do you think an external entity in your environment, such as your superior or health insurances like you to use such system?

Finally, participants are asked to complete an online questionnaire. Similarly to the TTF questionnaire, the TA questionnaire is evaluated using a 5-item Linkert scale, and is portrayed in figure 4.8.

Figure 4.8: TA Questionnaire

Technology Acceptance Questionnaire:
Clinical Decision Support for Critical Care Atrial Fibrillation

Please tell us about yourself.

Specialty	O Cardiology	O Internal Med.	O Intensive Care	O _____
Seniority	O Attending	O Fellow	O Resident	O Student
Years of Practice	_____			
Gender	O Female	O Male	O Other	

Please indicate your level of agreement with the following statements.

	Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	The presented system could improve my performance.	O	O	O	O	O
2	Use of the presented system requires little effort.	O	O	O	O	O
3	I believe the presented system would be easy to use.	O	O	O	O	O
4	Important people in my work environment would want me to use such system.	O	O	O	O	O
5	I think I would not require extensive training to use the presented system.	O	O	O	O	O
6	I think the presented system can be easily integrated into my current clinical workflow.	O	O	O	O	O
7	I would consider using the presented system.	O	O	O	O	O

4.4.4 Data Analysis

A qualitative analysis of the transcripts resulting from the three sessions is performed. Participants' statements are categorized into recurring themes and subthemes, which are primarily driven by the constructs of the TTF and UTAUT2 models. Comments are further categorized as barriers or facilitators.

For the quantitative analysis, the questionnaire responses are transferred to a spreadsheet for analysis. The responses of the TTF questionnaire are used to evaluate to what degree the assumptions made regarding the task characteristics are correct. This is assessed by computing the scores average and confidence intervals. It is assumed that if the assumed task characteristics are correct, the mean score per questions will reflect agreement and the standard deviation will not contain the middle score. Further, the task-technology fit will be assessed using an interaction approach [48]. The UTAUT2 questionnaire is evaluated in a similar fashion. For questions regarding perceived ease of use and perceived usefulness score averages and standard deviations are computed. The impact of individual statements on the intention to use is evaluated in an interaction approach [48].

Chapter 5

Results

5.1 Systematic Literature Review

The selection process is outlined in figure 5.1. Following a removal of duplicates, 1150 records remained. The unique records were screened using record titles and abstracts, resulting in an elimination of 942 records. 208 records that meet the inclusion criteria were further assessed for eligibility using the exclusion criteria, leading to an exclusion of 46 records. The remaining 162 records were included in the quantitative analysis.

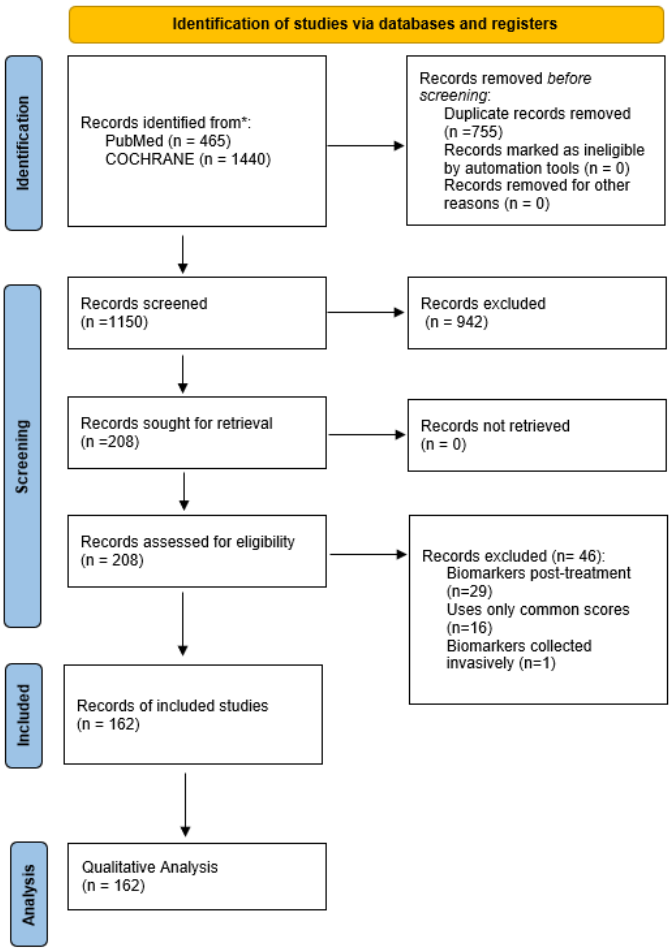

Figure 5.1: PRISMA 2020 flow diagram for of the performed systematic review

For the 162 records considered for analysis, the full text was analyzed to identify biomarkers used for outcome prediction. Biomarkers are grouped according to semantic categories, and the number of publications using each biomarker is identified.

Figure 5.2 presents the number of publications for each considered outcome. Of the 162 publications considered for analysis the majority predicted the onset or incidence of AF, followed by the prediction of AF recurrence after treatments.

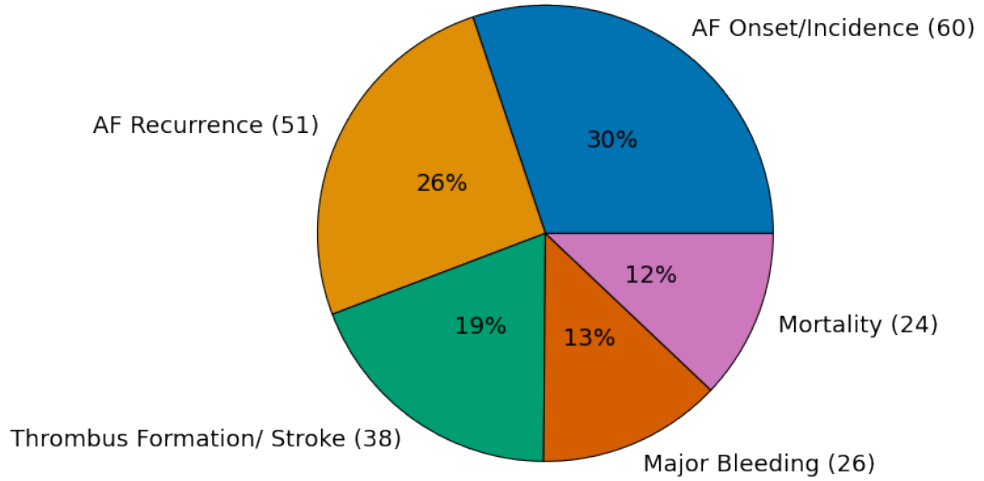

Figure 5.2: Number of publications predicting each outcome

Biomarkers used for prediction are grouped into categories shown in figure 5.3. It can be observed that the majority of biomarkers are obtained through serum tests, as well as imaging measurements.

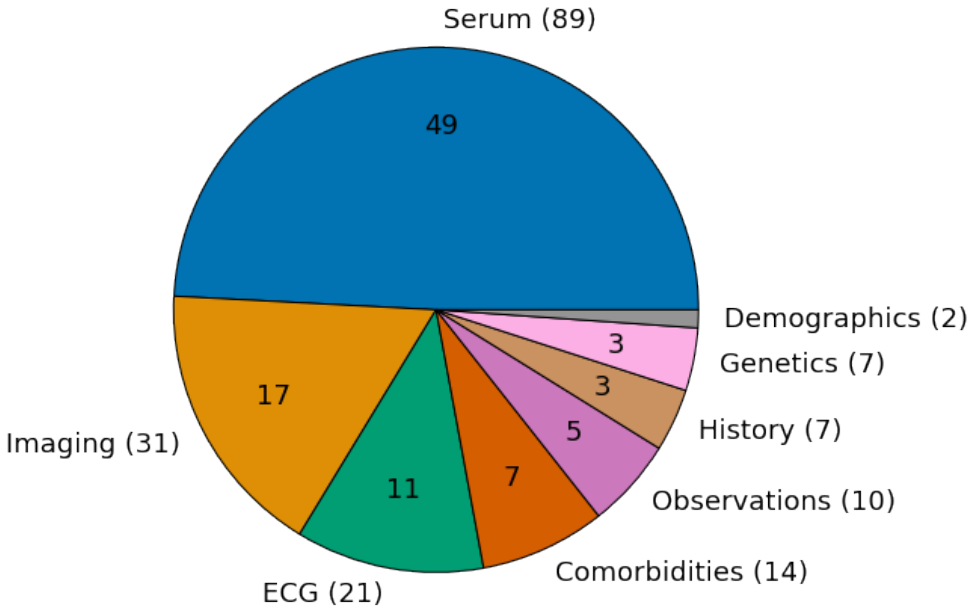

Figure 5.3: Number of biomarkers per biomarker category

5.1.1 Biomarkers

Table 5.1 shows the most common identified biomarkers, grouped by biomarker category, as well as the number of publications in which each biomarker is used. The full table of biomarkers is available in appendix 7.4.2.

Table 5.1: Most commonly used Biomarkers by Category

Category	Biomarker	Count (Prevalence %)
Demographics	Age	49 (31)
	Gender	38 (24)
Comorbidities	Hypertension	33 (21)
	Diabetes	31 (20)
	Heart Failure	27 (17)
	Vascular Disease	24 (15)
	Renal Dysfunction	20 (13)
Observations	BMI	21 (13)
	Systolic BP	18 (11)
	Alcohol Consumption	14 (9)
	Smoking	14 (9)
	Diastolic BP	11 (7)
History	Prev. Stroke	28 (18)
	AF Duration	14 (9)
	AF Type	11 (7)
	Prev. MI	10 (6)
	Prev. Bleed	8 (5)
ECG	Heart Rate	7 (4)
	PR Interval	5 (3)
	LV Hypertrophy	4 (3)
	P-wave Duration	3 (2)
	Premature Atrial Contractions	3 (2)
Serum	NT-proBNT	30 (19)
	GDF-15	18 (11)
	Hemoglobin	14 (9)
	Troponin-T	13 (8)
	IL-6	10 (6)
Imaging	LVEV	28 (18)
	LA Dilation	20 (13)
	LA Volume	17 (11)
	LA Ejection Fraction	9 (6)
	LA Contraction strain	6 (4)
Genetics	CYP 2C9 single-nucleotide polymorphism	2 (1)
	CYP11B2 rs1799998 polymorphism	1 (1)
	GJA1 rs13216675 polymorphism	1 (1)
	FRMD4B	1 (1)
	CAV1	1 (1)

5.1.2 Synthesis

The evaluation of currently used risk scores for AF treatment revealed that 83 biomarkers are used, whereas the systematic review showed that literature considers 182 biomarkers to be potential predictors of AF outcomes. This section compares the two sets to elucidate the discrepancy and discusses potential benefits of adding biomarkers mentioned in literature to prediction models. While the most commonly used biomarkers are evaluated in this section, rare markers are omitted. The complete list of predictive biomarkers in clinical risk scores and systematic review can be found in appendix 7.4.2.

Comorbidities

The three most commonly encountered comorbidities in risk stratification models currently used in clinical practice are hypertension, heart failure, and diabetes.

Hypertension - Hypertension is associated with a dilation of the left atrium due to barometric overload. Further, it is a causal factor of kidney dysfunction and other cardiovascular disorders which are risk factors for AF. Hypertension is used in prediction models for stroke, bleeding events, AF onset, and AF recurrence.

Heart Failure - Heart failure is a comorbidity that is used as a predictor of stroke and AF onset. Heart failure leads to a Ca^{2+} overload in cardiomyocytes, and increases the probability of EADs which are known AF triggers [157]. Heart failure has been associated with chronic inflammatory processes which lead to structural and electrical remodeling that maintains AF [30, 90].

Diabetes - Diabetes is predictive of stroke risk and AF onset. The mechanism by which diabetes causes AF needs further investigation, but it is assumed that inflammatory responses to diabetes drive AF [204].

Several comorbidities that are currently not used in risk stratification models have been proposed in literature.

Valvulopathy - Valvulopathies have been shown to be predictors of AF recurrence following catheter ablation [86] possibly due to a increased barometric pressure or volumetric overload within the atrium leading to a remodeling of atrial tissue. In-vitro studies have demonstrated that

an increase in atrial pressure promotes the maintenance of AF due to a dispersion of ERP facilitating reentrant propagation [92].

Non-alcoholic Fatty Liver Disease - It was demonstrated that patients with non-alcoholic fatty liver disease have a 2.47 times higher probability to experience AF onset [146].

Obstructive Sleep Apnoe - OSA has been shown to improve the predictive performance in models predicting AF recurrence. OSA is associated with an acute decrease in blood oxygen levels which likely lead to a decrease in ERP. Further, OSA leads to a sympathovagal activation creating a vulnerable atrial substrate [133].

Interatrial Block - An interatrial block is a dysfunction of the Bachmann's bundle and leads to a desynchronization of the two atria. It has been shown to be an independent predictor of AF onset [214].

ECG derived Biomarkers

Biomarkers obtained from ECGs have been included in currently used stratification scores. The proposed biomarkers primarily evaluate the morphology of the p-wave, such as its duration, onset and offset time, as well as its terminal force and axis [23, 220]. While p-wave morphology can be measured during SR, it is absent during AF episodes and can therefore not be assessed. During AF, the amplitude of the fibrillatory wave has been shown to be predictor of AF recurrence [1].

Beyond biomarkers obtained from the activity of the atrium itself, further electrophysiological measurements have been shown to be of value. Measures of heart rate variability and AF burden have been shown to be independent predictors of AF outcomes. Similarly, measures of ventricular function such as QRS duration, QT interval duration, t-wave deviations, the Sokolow-Lyon Voltage, and Cornell Product have been proposed.

Serum derived Biomarkers

An abundance of serum biomarkers have been identified as potential predictors of AF outcomes. In particular, recent advances are being made using biomarkers indicative of inflammation, tissue damage , and substrate remodeling.

Even though the mechanisms of inflammation markers are not fully understood, and significant cross-talk between inflammation markers exists, Interleukin-6, Galectin-3, and C-reactive protein have been investigated particularly well and shown to be predictors of stroke, bleeding, AF onset, AF recurrence, and mortality. Further, Growth differentiation factor 15 (GDF-15), a cytokine of the transforming factor beta family has been investigated intensively. It has been shown to be of predictive value in all outcome categories, and is the second most cited serum biomarker referenced in literature, while its use in clinical practice remains to be seen. The exact function of GDF-15 is controversial, but it is understood that its expression is regulated by inflammation, tissue damage, and stress [142].

Several markers have been identified to be representative of damage to cardiac tissue, and may play a role in the atrial remodeling process. Examples are different types of troponin, which are commonly used to assess myocardial infarction, have been identified as potential predictors in AF patients. While cardiac troponin-I has already been included in clinical stratification models, cardiac troponin-T and troponin-T have not, and remain exclusively used in the research environment.

Tumor growth factor-beta 1 (TGF-b1) is a cytokine that has been shown to be predictive in AF recurrence. TGF-b1 was shown to be elevated in AF patients, and of particular value in patients with persistent AF- less so in patients with paroxysmal AF. This indicates that TGF-b1 could be a marker of late-stage atrial remodeling in the form of being a pro-fibrotic marker [99].

Natriuretic peptides such as Brain natriuretic peptide (BNP) concentrations have been proposed by literature as possible predictors of stroke, bleeding, AF onset and recurrence, and mortality. BNP is a hormone that is usually secreted by ventricular cardiomyocytes as a reaction to stress, but its secretion has also been observed in the atria as a result of AF, with patients having higher BNP levels during AF episodes than before or after [112]. Several factors increase the NP level: Renal dysfunction, age, and sex (female). Conversely obesity and flash pulmonary edema decrease NP level. Furthermore, NP levels differ substantially depending on age, ethnicity, gender, and BMI [40].

Cholesterol levels have been associated with AF onset and recurrence. Its role in the facilitation of AF remains, however, controversial. While, in some studies, high cholesterol levels are associated with poor outcomes, other studies observe the opposite effect. The term "cholesterol paradox" has since emerged to describe this phenomenon [205].

Imaging Biomarkers

Clinical risk scores make use of the left atrial dimensions, and the left ventricular ejection fraction. A large number of further imaging biomarkers have been shown to be valuable predictors, but have not found their way into clinical stratification methods.

The most cited biomarkers are the left atrial ejection fraction, and the left atrial contraction strain, which are both markers of mechanical function. While ejection fraction is the result of two volumetric measurements, the systolic and diastolic volume, strain is measured using speckle tracking and an indicator of tissue deformation. Further, a common imaging marker is the E/e' ratio, which is measures the ratio of left ventricular filling pressure (E) and ventricular tissue velocity (e'). It is an indicator of ventricular diastolic dysfunction and a surrogate of atrial pressure [8], even though this relationship remains controversial and only applicable to a subset of patients [65].

Further, late gadolinium enhanced magnetic resonance imaging is commonly used to assess the degree of fibrosis in the atrial tissue which is a marker of local heterogeneity and decreased conduction velocities and facilitates reentries.

5.2 Semi-Supervised Clustering with Survival Data

5.2.1 Simulation Study

The clustering resulting from applying the tested algorithms to simulation data is shown in figure 5.4. The results demonstrate the advantage of incorporating survival data into the clustering domain, which results in non-linear decision boundaries in the covariate domain.

It can be observed that, in Sim-a, K-Means has not been able to capture the ground truth clusters. The lack of structure in the covariate domain does not provided the required information for an unsupervised clustering method to capture the simulated groups. The semi-supervised algorithms, however, suc-

cessfully capture the underlying survival distributions. This is an encouraging result, because unsupervised clustering algorithms assume clusters exist in the form of high-density areas in the covariate domain, which are not present in the simulated Gaussian distribution. Despite differences in decision boundaries among the obtained solutions with semi-supervised methods, each algorithm accurately identifies the presence of the three simulated clusters and their corresponding orientations. Even though SSC-Bair and SSC-Gaynor produce perfectly linear decision boundaries, as present in the ground truth, the alignment is not accurate, leading to dub-optimal accuracies. While LSS-K-Means produces decision boundaries that are not perfectly linear and some outliers are misclassified, it demonstrates the highest classification accuracy.

Sim-b, on the other hand, reveals a significant drawback of SSC-Bair and SSC-Gaynor, as both algorithms are unable to accurately capture the two simulated clusters. The algorithms use K-Means on either selected or weighted covariates, and, as a consequence, cannot produce non-linear decision boundaries, which would be required to correctly identify the two clusters. In contrast, LSS-K-Means and S-HAC are able to successfully capture the clusters with 100% accuracy. This highlights an important advantage of including survival information in the clustering process, as it drives cluster formation, and provides a clustering space in which K-Means is able to separate observations which would have been inseparable relying solely on covariates.

Similar observations can be made in Sim-c, where SSC-Bair and SSC-Gaynor produce a linear vertical split due to their inability to model the decision boundary by a straight line in the covariate domain. While LSS-K-Means and S-HAC are able to model the sine-wave-like boundary.

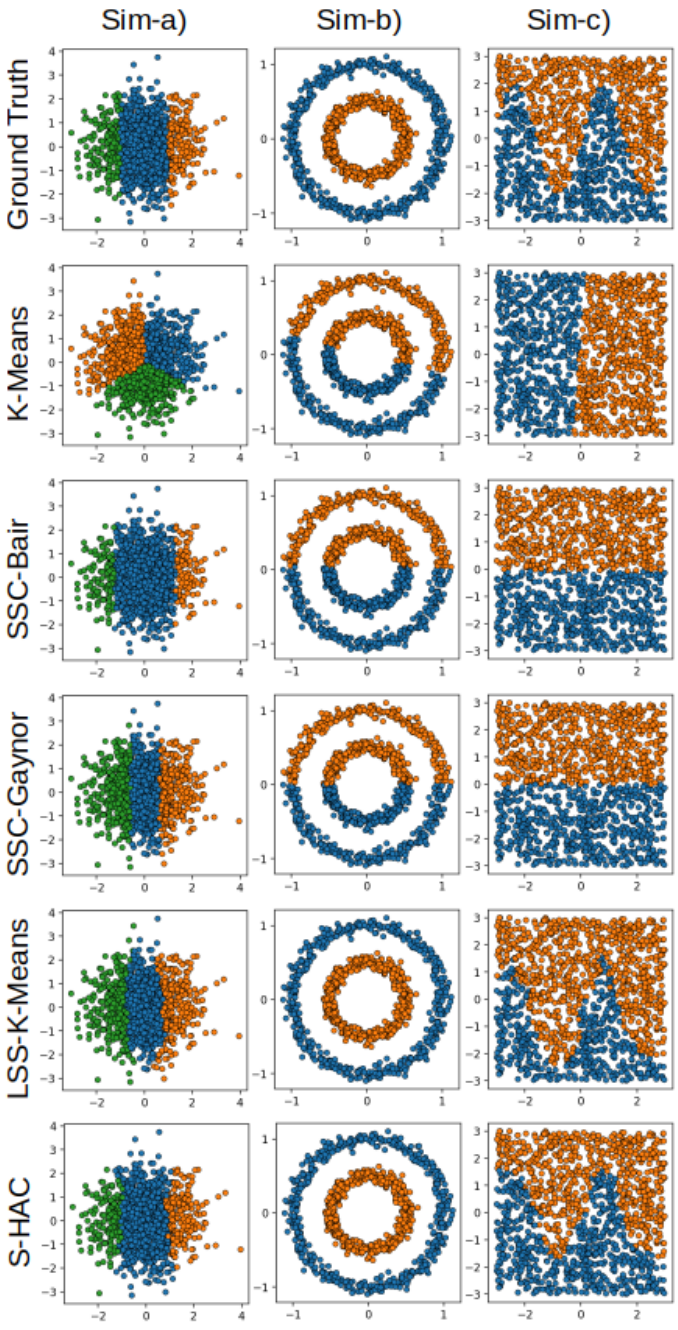

Figure 5.4: Clustering results for different algorithms on the simulated datasets

5.2.2 Benchmark Study

The result of the benchmark study are presented in table 5.2. S-HAC universally outperforms vanilla HAC, highlighting the importance of including survival data into the clustering algorithm, and underlining the fundamental shortcoming of unsupervised clustering algorithms in the capture of predictive patient phenotypes. HAC attempts to identify groups with cohesive covariates, while the inclusion of survival data forms clusters with cohesive survival times.

When compared to the state-of-the-art semi-supervised clustering methods SSC-Bair and SSC-Gaynor, S-HAC and LSS-K-Mean demonstrated superior performance in the majority of datasets. Only in the FLCHAIN dataset, has SSC-Bair demonstrated a higher concordance index. Ultimately, S-HAC has shown the most promising metrics, followed by LSS-K-Means, and the remaining algorithms.

Table 5.2: Benchmark metrics: Concordance index and log-rank test statistic with standard errors in parentheses.

Dataset	Algorithm	Concordance Index	Log-Rank Statistic
FLCHAIN	K-Means	0.670 (0.014)	156 (41.2)
	HAC	0.504 (0.002)	27.9 (26.2)
	SSC-Bair	**0.778 (0.021)**	302 (59.5)
	SSC-Gaynor	0.727 (0.019)	304 (37.4)
	LSS-K-Means (proposed)	0.730 (0.021)	321 (47.1)
	S-HAC (proposed)	0.771 (0.014)	**396 (81.9)**
SUPPORT	K-Means	0.619 (0.012)	131 (34.2)
	HAC	0.587 (0.013)	74.5 (18.9)
	SSC-Bair	0.753 (0.081)	963 (343)
	SSC-Gaynor	0.668 (0.011)	377 (52.3)
	LSS-K-Means (proposed)	0.758 (0.010)	813 (82.8)
	S-HAC (proposed)	**0.812 (0.004)**	**1280 (58.8)**
GBSG2	K-Means	0.560 (0.05)	6.27 (3.92)
	HAC	0.579 (0.023)	5.06 (2.41)
	SSC-Bair	0.574 (0.036)	9.92 (5.52)
	SSC-Gaynor	0.570 (0.047)	9.18 (8.99)
	LSS-K-Means (proposed)	0.602 (0.046)	**22.4 (19.5)**
	S-HAC (proposed)	**0.629 (0.038)**	13.4 (10.5)
METABRIC	K-Means	0.588 (0.023)	16.9 (7.15)
	HAC	0.511 (0.011)	4.94 (6.30)
	SSC-Bair	0.624 (0.053)	16.9 (7.15)
	SSC-Gaynor	0.586 (0.051)	15.4 (7.88)
	LSS-K-Means (proposed)	0.606 (0.047)	18.1 (6.86)
	S-HAC (proposed)	**0.635 (0.035)**	**28.9 (15.5)**

5.3 Atrial Fibrillation Phenotypes

5.3.1 Cohort Characteristics

A total of 46,520 patients were available in the analyzed database. 10,277 were found to have a diagnostic code for AF, corresponding to an incidence rate of 20.2%. After excluding patients with an age below 18 years and a hospital stay below 24 hours, a 9401 patients are included in the analysis.

Of the total 168 biomarkers, which were identified in the systematic literature review, a total of 34 was encountered in the database, and used for analysis. The cohort characteristics in terms of the 34 biomarkers are shown in table 5.3.

Following inverse probability of treatment weighting, patients appeared well-balanced across the treatment groups. The mean-absolute-difference in patient descriptors between treatment groups is provided in tables 7.3 and 7.4 in the appendix.

Table 5.3: Cohort characteristics.

Category	Characteristic	Median (IQR)/ n (%)
Medical History	Anemia	1127.0 (11.99)
	Arrhythmia History	1959 (20.8)
	COPD	1256 (13.4)
	Collagen disease	115 (1.22)
	Cor pulmonale	754 (8.02)
	Coronary artery atherosclerosis	3461 (36.8)
	Diabetes	2147 (22.8)
	Thyroid disorder	1051 (11.2)
	Heart failure	4128 (43.9)
	Hypertension	4562 (48.5)
	Myocardial infarction	933 (9.92)
	OSA	403 (4.29)
	Post-operative condition	3250 (34.6)
	Renal insufficiency	3261 (34.7)
	Respiratory failure	1450 (15.4)
	Rheumatism	403 (4.29)
Laboratory Measurements	Erythrocyte distribution width [ratio]	14.4 (13.5–15.7)
	Erythrocyte count [#/µm]	3.88 (3.38–4.37)
	Serum calcium [mg/dL]	8.60 (8.10–9.10)
	Serum creatinine [mg/dL]	1.10 (0.80–1.50)
	Serum magnesium [mg/dL]	2.00 (1.80–2.20)
	Serum potassium [mmol/L]	4.20 (3.90–4.70)
	Serum sodium [mmol/L]	139 (136–141)
	Hemoglobin [g/dL]	12.1 (10.6–13.4)
	Leukocyte count [#/nL]	10.0 (7.30–13.7)
	Platelet count [#/nL]	215 (163–279)
	Prothrombin time [s]	14.5 (13.1–18.1)
Observation	Heart Rate [BPM]	85.5 (75.0-98.0)
	Left Atrial Dilation	4183 (44.5)
	Right Atrial Dilation	2813 (29.9)
	Sepsis	1144 (12.2)
	Valvulopathy	2865 (30.5)
Demographics	Age [years]	76.5 (67.3-83.6)
	Male sex	5364 (57.1)

5.3.2 Patient Phenotypes

Based on the best performing set of hyperparameters, S-HAC captured a total of 26 patient phenotypes in the analyzed cohort. Figure 5.5 portrays the T-SNE embedding of the clusters. A description of each cluster's characteristics is provided in tables 7.5 and 7.6 in the appendix. The obtained clustering allowed for a prediction of the rhythm control rate, the rate control rate, and the mortality rate with concordance indices of 0.69, 0.56, and 0.64, respectively.

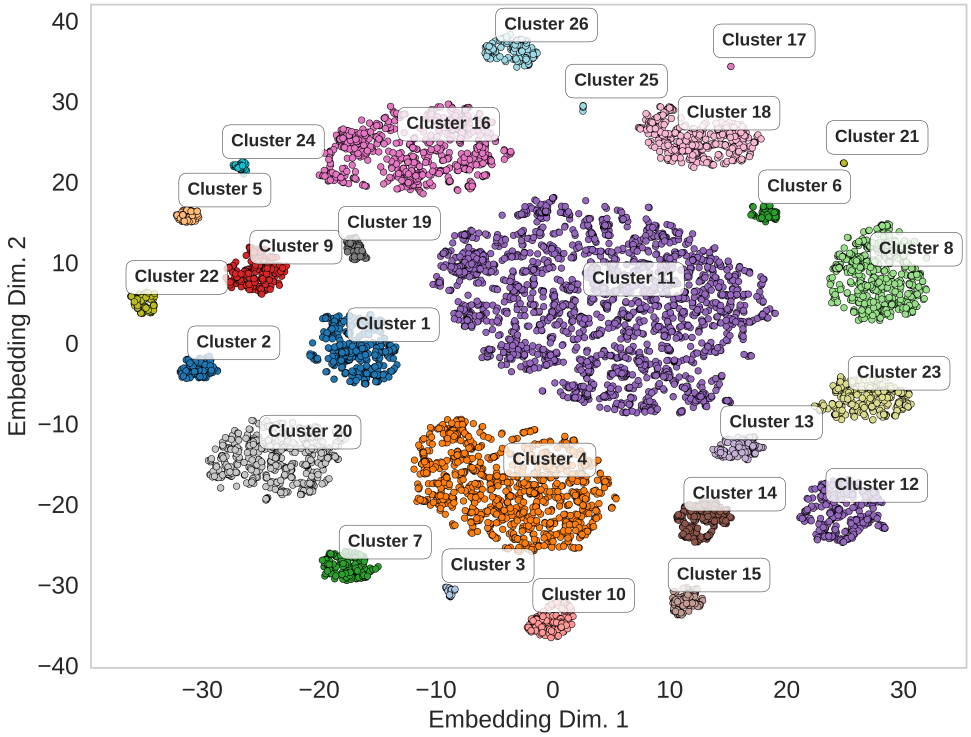

Figure 5.5: T-distributed stochastic neighbor embedding of the obtained clusters [123].

5.3.3 Treatment Responses

Average treatment effects are covered in the following subsections. Due to the large amount of information generated, only a selection of clusters is discussed. A complete enumeration of adjusted treatment effects is presented in tables 7.5 and 7.6 in the appendix.

Rhythm Control

Rhythm control rates were the highest in patients receiving PCBs (9.78%/h [8.74-11.0]), followed by those receiving CCBs (4.73%/h [4.04-5.74]). Lower rhythm control rates were observed in patients treated with BBs (2.28%/h [1.81-2.61]) and MgS (1.90%/h [1.44-2.54]).

Among different phenotypes the cardioversion varied considerably with each treatment showing a complete lack of cardioversion in at least one phenotype. Maximum conversion rates differed across phenotypes with conversion rates of 6.54%/h (3.39-16.8), 15.0%/h (9.61-35.6), 12.2%/h (4.16-45.7), and 3.47%/h (1.20-23.3) for BBs, PCBs, CCBs and MgS, respectively. While PCBs were generally superior to CCBs, CCBs appeared to be more effective in three clusters. In particular, clusters with a high prevalence of thyroid disorders showed higher cardioversion rates with CCBs than PCBs (clusters 15 and 19). Similarly, patients with renal insufficiencies and postoperative conditions appeared to cardiovert better with CCBs than PCBs. This association was, however, only observed in combination with a high prevalence of coronary atherosclerosis.

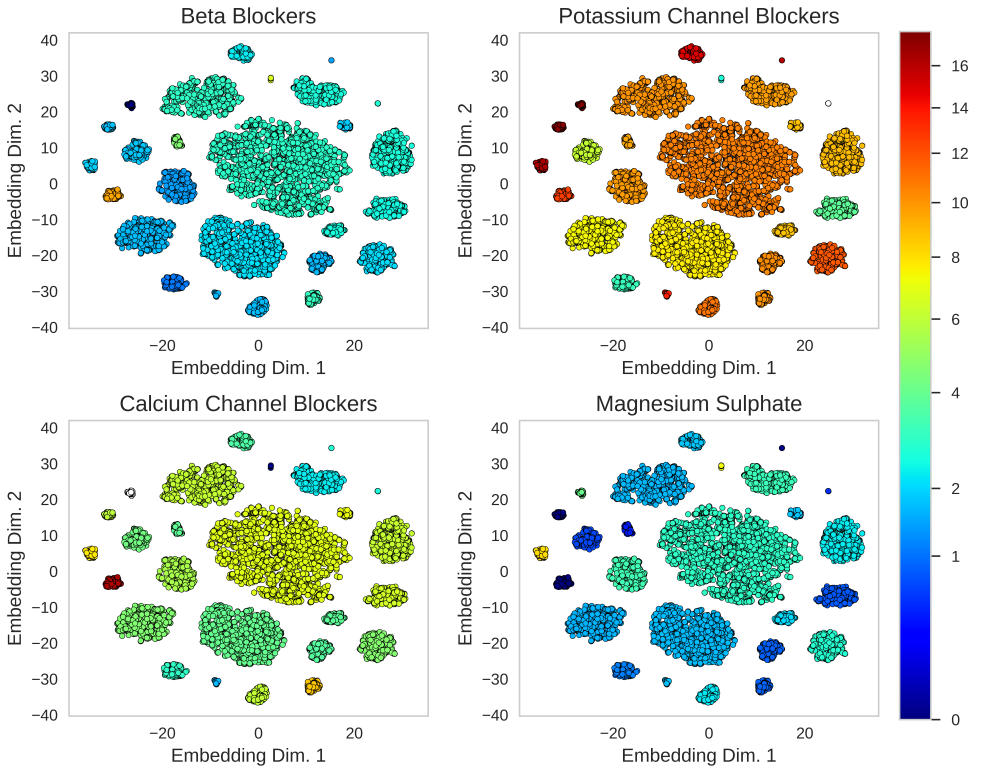

Figure 5.6: Adjusted hourly rhythm control rates for different treatment groups and phenotypes.

Rate Control

Heart rate was best controlled using in patients receiving CCBs (18.8%/h [16.4-22.8]) followed by those receiving MgS (15.6%/h [11.4-19.2). Inferior rate control was observed in patients receiving PCBs and BBs (15.2%/h [13.2-17.5], 15.6%/h [11.4-19.2]). Rate control rates for different treatment groups and phenotypes are portrayed in figure 5.7.

Even though BBs showed the lowest efficacy in the cohort analysis, they appeared to be superior to other treatment groups in several clusters. This observation was made for clusters 3 and 10, which are both characterized by a high share of female patients (97.8%, 78.5%).

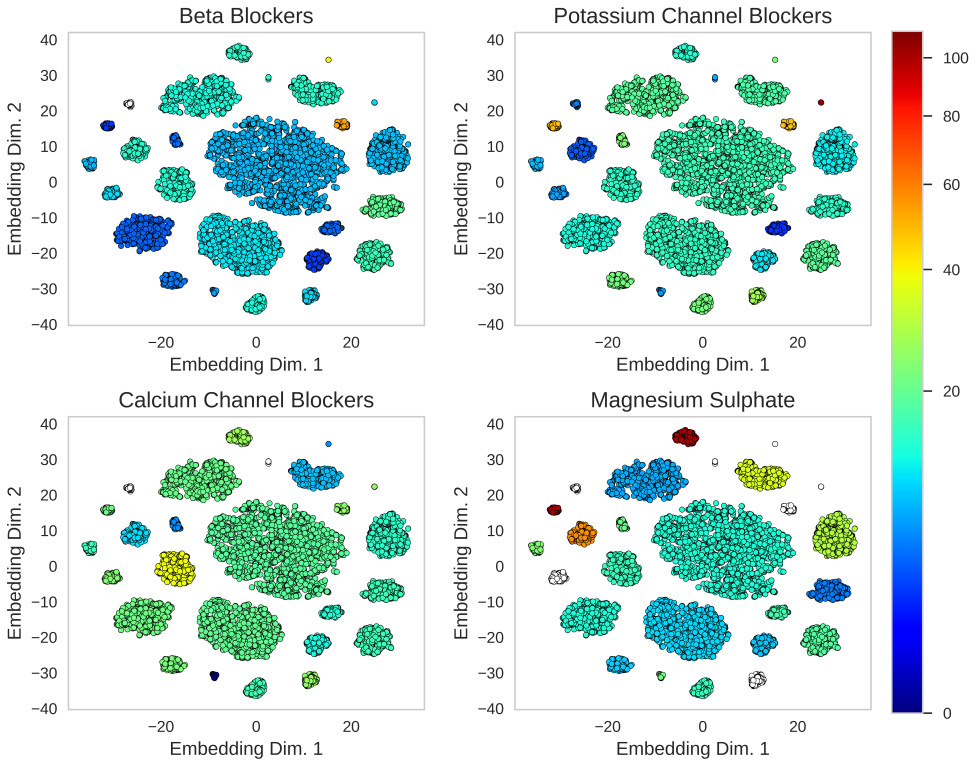

Figure 5.7: Adjusted hourly rate control rates for different treatment groups and phenotypes.

Mortality

Mortality within the cohort varied between treatment groups. The highest mortality was observed in patients receiving MgS (1.45%/d [1.07-1.99]), followed by CCBs and BBs (1.40%/d [1.04-1.88], 1.28%/d [1.00-1.51]). The lowest mortality rates were present in patients receiving PCBs (0.95%/d [0.78-1.18]).

Figure 5.8 portrays the daily mortality rates for different treatment groups and phenotypes. Even though patients in the PCB treatment group showed the lowest mortality in the cohort analysis, patients in cluster 15 demonstrated the highest mortality rates when exposed to PCBs, with modest evidence suggesting mortality higher than in the entire cohort (BF=7.80). Cluster 15 is

primarily characterized by the highest rate of thyroid disorders (90.5%), and is predominantly female (89.0%).

Particularly high mortality rates with MgS can be observed for clusters 7, 10, 16, 17, 24, and 25, which are all characterized by relatively high shares of patients with renal insufficiency. Compared to the population level these clusters had a considerably higher mortality with evidence levels varying from moderate to very strong (BF = 24.6, 1.83, 2.68, 20.0, >100, >100, respectively).

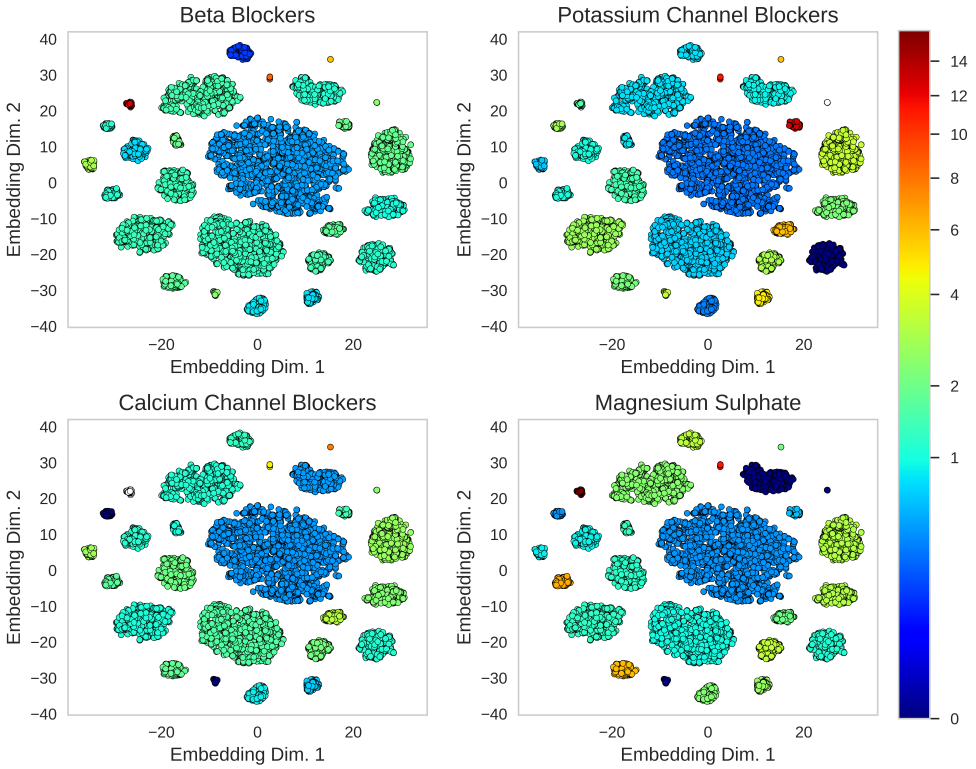

Figure 5.8: Adjusted daily mortality rates for different treatment groups and phenotypes [123].

5.4 User Centered Evaluation

A total of five participants were recruited and completed the user centered evaluation. The participants' characteristics are shown in table 5.4.

Characteristic	P1	P2	P3	P4	P5
Gender	Male	Male	Male	Male	Female
Specialty	Cardio.	Intern. Med.	Cardio.	Cardio.	Cardio.
Seniority	Resident	Resident	Resident	Fellow	Fellow
Years of Experience	5	5	3	6	8

Table 5.4: Characteristics of study participants

5.4.1 Qualitative Analysis

Empathy Session

Several themes were identified in the discussions relating to the current use of decision support tools, advantages and disadvantages of current care, perceived challenges, the number of possible treatment options, as well as patient heterogeneity.

Current use of decision support tools

Care is guided by internal standard operating procedures (SOPs), which provide standardized information for different patient conditions. An SOP specific to AF was, however, not present at the participating institution at the time of the study, as stated by one participant:

> P1: *We have no standard operating procedure, which primarily deals with AF.*

Decision-making was primarily based on the ESC guidelines. As clinicians described:

> P1: *"In many situations we read the guidelines. These have flowchart systems. But I do not extensively use the ESC decision tools, because they are too indiscriminate for the specific situations. We work with scores like the CHA_2DS_2–VASc score, even though it is not validated for the ICU."*

P2: *"The CHA_2DS_2–VASc score is not derived from ICU cohorts. Neither is the HASBLED score. At least to my knowledge, I have not validated that."*

Advantages and Disadvantages of current care

Current care is primarily based on outpatient guidelines provided by the ESC [89], which are interpreted with their limitations in mind. A perceived advantage of such an approach is the simplicity, which was appreciated by one of the participants:

P1: *"A clear strength, to me, is simplicity. Especially in intensive care, we have a lot of beginning doctors, and the flowcharts provided by the guidelines are simple, easy to look up at 3 in the morning, and generally simple to follow."*

Given a lack of guidelines tailored to critical care, and the adaptation being dependent on the physician, the possibility of inconsistencies in care was suggested by one clinician:

P1: *"A very great disadvantage is the inconsistent procedure, I think. That there is no algorithm for how to deal with AF in the ICU in general."*

Similarly, the ESC guidelines were considered too simplistic to reflect the complexity of patients in the ICU, with guidelines considering patients with individual comorbidities. In practice, however, patients would often present with multiple relevant conditions.

P1: *"The primary drawback is that the complexity of intensive care is not sufficiently accounted for. Also, the complexity of the patient is not considered. The guidelines clearly aim for little input information that is required to work through the flowchart [to select a treatment]."*

P2: *"No, its not well understood and its difficult, because ESC assumes that a patient has a single problem. What we find in the ICU, however, especially if patients are here longer, is that they have multiple problems. [...] We adapt the guidelines to the situation, even if they are not developed for our specific situations."*

Challenges

Challenges beyond the simplicity of the employed guidelines were the large number of possible treatment options. This challenge is remedied by the par-

ticipating institution only using a limited number of possible antiarrhythmic agents.

Further, challenges were identified with respect to possible adverse effects of diagnostic procedures, such as transesophageal echocardiography, which one participant described in further detail:

> P1: *"The main challenge is the danger of a thromboembolic event. [...] The question is then if we want to do transesophageal echocardiography, which would expose them to further risk especially when oxygen demand is high. I find this to be a difficult situation, because the risk-benefit ratio between a "blind" cardioversion without excluding a thrombus, or accepting the risk of transesophageal echocardiography is a difficult assessment."*

AF was described as a frequent and complex condition with a heterogeneous patient population. At the time of the study, participants were not aware of any ongoing studies specifically aimed at AF in their institution. One participant described the multifaceted nature of AF as being a possible obstacle in its study:

> P1: *"The problem I see is that AF occurs alongside other conditions, and is seldom specifically addressed. A lot of analyses are made for COVID, cardiogenic shock, STEMI, pneumonia, etc. A lot of things are done, but AF is somewhat cross-sectional. No patient comes to the ICU because of AF, so AF runs alongside, which may be the reason why it is not specifically addressed by us, because nobody looks at it specifically. Especially because the population is so heterogeneous."*

TTF/UTAUT2 Session

Similarly to the empathy session, the TTF/UTAUT2 sessions allowed for the identification of several themes and subthemes, which are presented in table 5.5.

Category	Theme	Subtheme
TTF	Task	Information Requirements Definition of Utility Confidence in Treatment Selection Patient Heterogeneity
	Technology	General Data presentation Explainability Scope of Application
UTAUT2	Effort Expectancy	Situational factors Technology characteristics
	Performance Expectancy	Evidence-based decision-making
	Experience	-
	Habit	-
	Social influence	-

Table 5.5: Categorized themes and subthemes

TTF - Task - Information Requirements

Different types of information were identified as relevant to treatment selection. A particular focus appeared to be on the prevention of stroke, which was mentioned almost universally. This was followed by mentions of rhythm and rate control, as well as complications, mortality and bleeding events. Study participants mentioned different viewpoints with regard to the outcomes:

> P4: *A hard parameter would be the outcome. The prevention of stroke. [...] Death... Bleeding.*

> P2: *For many medications I have experience. For many there are published results, which I remember. The correlation between stroke and CHA_2DS_2–VASc is clear. That NOACS, to some degree, prevent that [stroke]. That beta blockers lead to rate control, and sodium channel blockers control the rhythm is about 50 percent, I estimate.*

> P5: *That depends on the patient's clinical presentations, and how bad they are doing. Based on that I would decide how quickly to treat the AF. That is point one. The second point is that we know what the complications of a*

rhythm stabilizing therapy are, such as stroke risk or success rates. Based on these risks we would decide what the best option is. Stroke is the one that everyone is worried about, and that AF returns.

Further, one clinician mentioned the source of information they would use to select a treatment, which is a randomized study presenting the treatment effect of the treatment in question:

P5: *I would judge that based on a randomized study. With a control group and treatment groups. This measures the treatment effect.*

TTF - Task - Definition of Utility

Several statements regarding treatment utility and the definition of what constitutes the "best" treatment choice were identified.

Participants universally differentiated between positive and negative effects of treatments. Treatment utility was expressed as the conversion and maintenance of SR, control of symptoms, and the prevention of stroke, harm, and side effects.

P4: *Utility, in AF, would ideally be the cardioversion to sinus rhythm. That would be the utility. The risk would be side-effects which arise from the therapy.*

P2: *The best treatment, I think, would be one with which you can restore and maintain the sinus rhythm. The goal would be that the patient is symptom-free.*

P3: *For one, in a symptomatic patient, the symptoms need to be controlled. If rhythm control is desired, then rhythm needs to be controlled. And when the patient has had AF, and has a high CHA_2DS_2–VASc, then the patient needs to be protected against stroke.*

P1: *I mean, treatment can be broken down into rhythm control, rate control, and anticoagulation. The goal is naturally safety for the patient. Ensuring the patient does not have a stroke, this is the most basic one, and is determined using the CHA_2DS_2–VASc score. The second goal would be the restoration of sinus rhythm. Patients are very different. In some patients you can control the rate.*

One participant explicitly expressed treatment utility to be the probability of resolving the problem, while another mentioned the speed with which the

desired effect can be reached. Further, one participant considered the economic aspect of treatment selection an important factor.

> P5: *That the patient's heart rhythms and heart rate is controlled relatively quickly and without complications. That the treatment leads to the stabilization of the patient and heart rhythm and rate are controlled.*

> P2: *The utility is the probability that the underlying problem is resolved. The harm, in terms of AF, are possible side-effects of medications, and the thought of the probability of side-effects. The utility needs to outweigh the side-effects [for a treatment to be considered].*

> P4: *The best treatment is the one that resolves the situation with the best possible outcome and with the lowest harm. One has to honestly say that the aspect of economics needs to be considered.*

TTF - Task - Confidence in Treatment Selection

Participants gave mixed responses when asked if they could select treatments with confidence. Two participants considered optimal treatment selection to be possible, with one participant questioning to what extent different decision-makers would reach the same conclusion.

> P1: *In terms of AF, I think that it can be done. I do not know if different people would come to the same decision, but I think it can generally be done. To what extent everyone makes the same decision- I do not know. [...] Looking back, some patients may not receive the best treatment. If these choices are different, then some will be better, and others worse.*

> P3: *Yes, I think so.*

Further participants considered the optimal treatment selection to be a difficult endeavor and were not sure of their ability to make the optimal choice. One particular challenge expressed by two participants is the number of possible treatments, even within a specific drug group. One of the participants, however, stated that experience is a possible remedy.

> P5: *We cannot know that beforehand. It depends if we obtain rhythm control, if a rhythm control strategy is chosen, or if the patient ends up having complications. We have so many choices, that it is always different what we choose.*

P4: *I cannot be completely sure. It is assumed that they are all equally effective. I cannot tell to what extent one beta blocker is better than another in a specific patient. This is mostly based on clinical experience.*

TTF - Task - Patient Heterogeneity

AF patients were considered heterogeneous. To what extent this heterogeneity was considered an obstacle, differed among participants, with some participants considering different treatment approaches to be sufficient, and said heterogeneity not being a particularly grave problem:

P5: *Yes, sure, that is the main point. If you have someone who just had a stroke or is going to undergo surgery, one will rather choose a rate control strategy, because one cannot anticoagulant these patients, when they are cardioverted. It all depends on the patients: What comorbidities do they have? What their clinical presentation is, how their vitals are. Do they have a heart failure, and how high is the risk of bleeding or thromboembolic events?*

P2: *Heterogeneity leads to different approaches. I do not know if its an obstacle. I think with enough experience, heterogeneity is not a substantial problem.*

Other participants considered the patient population to be sufficiently heterogeneous and complex to consider it an obstacle. In particular one participant described treatment to be somewhat of a "one size fits all" approach, which may be insufficient:

P3: *Yes. Naturally there are very complex patients.*

P4: *We often work with a one-size-fits-all approach. Naturally, doses are adjusted, for example, depending on patient age. Patient heterogeneity may, in fact, constitute a problem, I would say. AF is a very complex clinical picture.*

TTF - Technology - General

Participants generally reacted positively to the presented CDSS concept, stating they see their own decision-making process in the approach, and the results reflecting clinical reality. As some participants stated:

P1: *Naturally, I recognize my own decision-making process in your approach.*

P4: *I think it's a good thing. I think it is interesting how many clusters were identified, because that reflects the clinical reality.*

TTF - Technology - Data Presentation

Participants commented on the information presented in the conceptualized CDSS. Its central element, a point cloud of patients from the development dataset was considered of little use. A bar chart showing the utility of different treatment options was, which was not present in the CDSS front-end, but in the PowerPoint presentation, was considered valuable and its integration into a CDSS was recommended. Further, one participant considered the communication of confidence intervals a necessity.

P1: *To what extent the point cloud is useful- to be honest- I do not think I can make much use of it at 3 in the morning.*

P1: *The bar chart which you showed [in the presentation], with the beta blockers, potassium and calcium antagonists, that I can make use of very well.*

P2: *The treatment effects should include a confidence interval.*

TTF - Technology - Variable Selection

The selection of variables used for patient characterization has been criticized as being incomplete or inaccurate. Further, one participant suggested that the evolution of patient parameters should be considered.

P3: *A drawback is the number of variables, which is limited. We do not have a complete picture of the patients. [...] It is limited to how the patient is "captured" from the data perspective. The patient evolves, for example with a brand-new ECG from the patient, which would assign the patient to a different group. Also, personal preferences or intolerances to certain medications are not considered.*

One participant commented on the explainability aspect of the presented CDSS, suggesting the addition of "red flags" to predictions. Such information would flag the most important features contributing to a predicted treatment being high or low.

P1: *What I think could make it more trustworthy is the addition of "red flags". Such that one could determine which "red flag" parameters led to [a specific treatment] being ranked very last. This could also be useful in situations where someone says "with my 30 years of experience, I think*

magnesium would be the best medication", but someone else says "yes, but our algorithm says, for a really manifested reason, that we should not choose it". I think that would make sense. This does not have to be a text, but it can be a positive-negative list, such that one can see "increased mortality with respiratory insufficiency and magnesium" or "Danger! Thyroid disorder. Increased rate of side effects with amiodarone." Short and concise. What has driven the program to the conclusion? That way one would not just act on the programs recommendations but understand why a decision is recommended. Not all parameters need to be explained, but the ones that have had the highest impact on the result.

TTF - Technology - Application Scope

One participant suggested that the scope of the presented CDSS could be expanded to not only cover the ICU, but also the emergency department.

P1: *Another thing is that the focus on the ICU is good, but I think that it [the CDSS] is something that the emergency department (ED) could extremely benefit from. We have a large ED with 55,000 patients per year. We have a lot of patients who come to the ED with AF, and the treatment is often initiated there. If patients then get admitted to the ICU, because they are unwell, one would have to say that the introduction of such a CDSS needs to happen in both departments, such that the same cognitive approach is taken.*

UTAUT2 - Effort Expectancy - Situational Factors With regards to the ease of use of the presented system, participants mentioned several situational factors that could impact the usability of the CDSS. One participant considered an independent application on PC to be a possible obstruction and recommended that such CDSS should be integrated into existing technological platforms. Another participant suggested the use of an app on a tablet to be an appropriate method. While speed of access was considered an important factor, situations in which the presented CDSS would be used were not particularly time-critical, because AF is not immediately life threatening.

P1: *When we begin work in the morning, we have 3 to 4 programs open in parallel. Technically, this is already too much to be honest. Why do I need 3 programs? I need one. I see the danger that there will be a dedicated program for radiology, AF, etc. And at some point, we will have 17 thousand programs, which slows down the computer and is confusing. At some point one no longer feels like using anything. If it was integrated into existing programs, it would be much easier.*

P3: *There needs to be quick accessibility, ideally as an app on a tablet, or optimally an integration into the data management system, such that information can be extracted directly.*

P5: *Yes, there are only a few cases in which action is required immediately. Generally, one has a few minutes of time.*

UTAUT2 - Effort Expectancy - Technology Characteristics

One aspect that was particularly poorly viewed was the necessity of manually entering patient information for processing. Participants suggested an integration into existing hospital information systems to reduce time and effort.

P1: *Yes, that is what is probably not realistic. There needs to be a data interface to reduce the number of manually entered parameters.*

P2: *Such systems need to be better [than the current standard of care], applicable, and not be time consuming. But this does not seem to be the case.*

P5: *It is only a few clicks.*

UTAUT2 - Performance Expectancy - Evidence-based Decision-Making

Participants agreed that the presented CDSS could increase their performance. In particular, it would offer a standardized evidence-based approach for treatment selection and remove ambiguity.

P5: *A big plus [of using such a CDSS] is that [current] decisions are often not based on clinical facts, but based on gut feeling. Such tool can be a good aid in such situations.*

P2: *Pretty surely, yes. The efficiency in the sense of efficient patient care- I think so.*

UTAUT2 - Experience

The experience level as a modulating factor for CDSS use was mentioned by one participant, who considered its use to be particularly valuable for colleagues with limited experience.

P2: *I think especially for people with little experience this would be helpful. Someone who has already seen a thousand AF patients should be able to judge based on their experience. Surely, time savings may exist.*

I think physicians are used to work with clinical randomized studies before they would really use anything. Many may use it for a recommendation, but not definitely follow that recommendation.

With regards to the social influence of technology acceptance, different viewpoints were taken by participants. While one participant stated that the use of novel technologies is primarily suggested by "middle level" clinicians, another stated that possible increases in efficiency would likely be positively received by senior management, as well as other staff due to freed up capacities:

P2: *In general, novel technologies have difficulties to be used. It usually takes some time. I imagine such changes do not necessarily come from the "top", but maybe even from the "middle level". Digitalization was previously initiated from normal doctors, not from the "top".*

P1: *The clinic director and the chief financial officer would be happy. And, especially in times when we have limited beds or a staff shortage, we would be happy with every patient that can be dismissed earlier, especially because the freed-up capacities can be used for other patients.*

5.4.2 Quantitative Analysis

The results from the TTF questionnaire are shown in figure 5.9. Participants generally agreed with the statements in the TTF questionnaire with average responses exceeding the middle value of 2.5 points on the Likert scale. Participants uniformly agreed that the patient population exhibits heterogeneity and sub-groups exhibited different responses to treatments (mean 5.00, SD 0.00). While this heterogeneity in combination with the large number of available treatments was perceived as a difficulty when selecting treatments, not all participants fully agreed and one participant stating that they neither agree nor disagree (mean 4.00, SD 0.63).

In terms of the CDSS design choices, participants agreed that a division of patients into sub-groups was a useful approach (mean 4.60, SD 0.49), and the use of a utility function to rank treatment options received similarly favorable opinions (mean 4.60, SD 0.49). The lowest agreements were recorded in the two final statements, which evaluated the use of a k nearest neighbors approach to assign patients to a phenotype (mean 3.80, SD 0.40), as well as the reduction of treatment effects to hourly and daily rates (mean 3.80, SD 0.75). While both statements showed high agreement, the results hint at room for improvement.

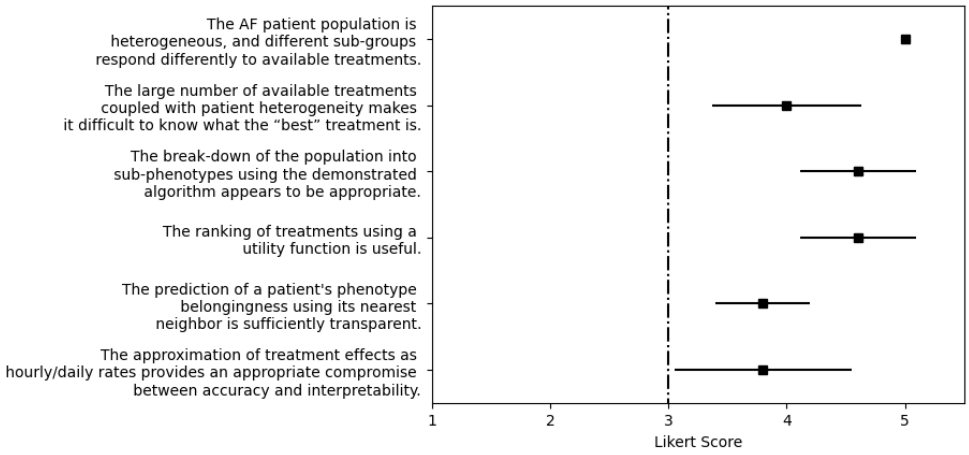

Figure 5.9: Responses to the TTF questionnaire. Markers and solid lines indicate the mean response and standard deviation, respectively.

The results from the TA questionnaire are shown in figure 5.10. Participants uniformly agreed that the presented system could improve their performance (mean 4.00, SD 0.00). With regards to the expected effort, however, participants showed concern of increased workload (mean 3.40, SD 0.80), which may be attributed to the necessity of manually registering patients' descriptive variables. Nonetheless, participants agreed that the presented system would be easy to use (mean 4.00, SD 0.63). With regards to the social influence, participants generally agreed that people in their environment may expect them to use the presented system (mean 4.00, SD 0.63). Further, the requirement for extensive training to use the presented system was not anticipated by the study participants (mean 4.00, SD 0.63). The last statement, expressing the intention to use the presented system, received the highest degree of agreement with 3 participants agreeing and 2 participants completely agreeing with the statement (mean 4.40 SD 0.50).

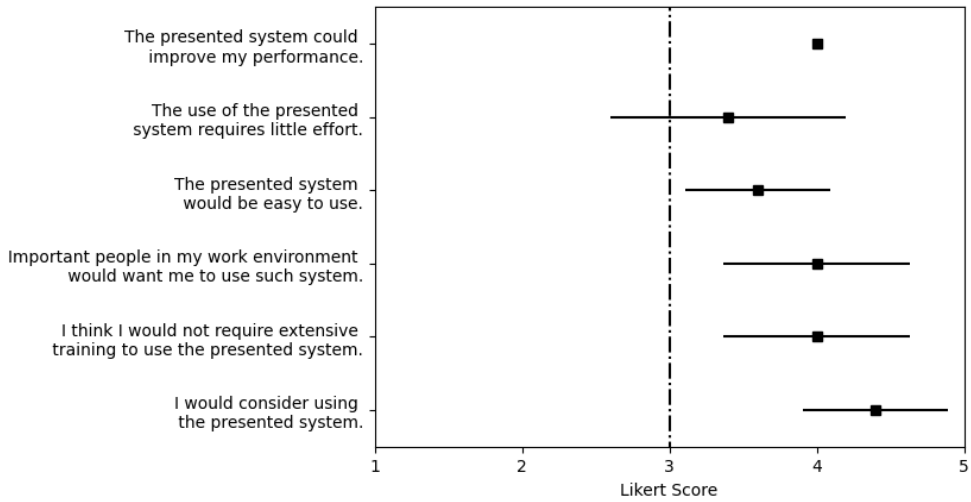

Figure 5.10: Responses to the TA questionnaire. Markers and solid lines indicate the mean response and standard deviation, respectively.

Chapter 6

Discussion

6.1 Systematic Literature Review

A total of 154 predictive biomarkers were found. Of these, less than half are currently employed in clinical risk scores. Further, studies performing subgroup analyses indicate that relationships between biomarkers are complex and require careful modeling. As such, [219] have shown that some genetic risk factors are only predictive for patients of particular geographical regions, but not others, and, further, that several genetic risk factors are only predictive in the presence of certain comorbidities. In a similar fashion, [22] showed that BNP levels are a stronger predictor of AF recurrence in paroxysmal AF patients than in the general AF population. On the other hand, [40] demonstrated that NP levels depend on the ethnicity, sex, and gender. Such insights reveal the existence of complex relationships between predictive variables and call for the application of more complex predictive models than currently used point scores or linear models.

The obtained results need to be interpreted with several limitations in mind. Firstly, the study is limited by the number of databases, and the requirement for studies being published in English language. Further studies introducing additional biomarkers may exist in other databases or in further languages.

Further, the quality of the incorporated publications was not assessed, leading to potentially biased results.

6.2 Semi-Supervised Clustering with Survival Data

6.2.1 Simulation Study

The results of the simulation study highlight a limitation of applying K-Means exclusively to the covariates of the observations. When decision boundaries between clusters are non-linear, there is no subset of covariates, and no set of weights which could transform the covariate space such that decision-boundaries become linear. In fact, a clustering space with non-linear decision boundaries will lead to incorrect classifications and degrade the accuracy of K-Means.

Further, the results reveal an intriguing characteristic of LSS-K-Means, wherein the incorporation of survival data into the clustering process leads to the creation of non-linear decision boundaries in the covariate domain that can closely approximate the boundaries of the underlying survival distributions. In particular, in Sim-b and Sim-c the inclusion of survival data provided encouraging results, and demonstrated the fundamental problem of the K-Means algorithm. The assumption that clusters are spherical and of similar variance, and, ideally, separable with a reasonable margin, may not hold. Like its competitors, LSS-K-Means will create its decision boundaries along a hyperplane equidistant to the two closest cluster centroids. These hyperplanes are, however, created in the combined space of covariates and survival data, which results in a non-linear decision boundary in the covariate domain, which can closely resemble the ground truth distributions between clusters of different survival distributions.

Even though LSS-K-Means generally demonstrates favorable test metrics, the performance of different clustering algorithms may depend on different variables, and no algorithm is guaranteed to provide the best results. However, the results of the benchmark study suggest that LSS-K-Means may be a promising method for clustering with survival data, often outperforming competing methods.

6.2.2 Benchmark Study

The benchmark study has shown that the developed algorithms, LSS-K-Means and S-HAC, generally outperformed not only unsupervised KMeans and HAC, but also both semi-supervised contenders, SSC-Bair and SSC-Gaynor. Neither SSC-Bair nor SSC-Gaynor universally outperformed the unsupervised clustering algorithms. Similar observations have been made in a previous study [32], which, among others, compared K-Means to SSC-Bair.

This phenomenon can be attributed to the assumption of linear independence between covariates, which is made by both algorithms. SSC-Bair and SSC-Gaynor both use Cox regression models to quantify the importance of individual covariates. This approach assumes a strictly increasing/decreasing impact of the covariate on the measured outcome, and cannot account for covariates that result in an increased risk on the extremes of the covariate range. Further, complex covariate interactions cannot be accounted for, making the application of a Cox regression model a questionable choice. In a similar fashion, both SSC-Bair and SSC-Gaynor use the K-Means algorithms to achieve clustering. This may result in inaccurate decision boundaries. Neither limitation are true for S-HAC, which uses the underlying data structure to define a decision boundary resulting in superior test metrics.

6.3 Atrial Fibrillation Phenotypes

The application of S-HAC to the patient cohort resulted in clusters with different treatment responses. 10-fold cross validation showed high concordance for the outcomes rhythm control and mortality. The prediction of rate control has, however, showed low performance. This may be attributed to the fact that none of the biomarkers identified in the systematic review, and, hence, the employed patient descriptors were predictive of the success of a rate control strategy.

The identified phenotypes partially aligned with existing knowledge of treatment effects on patient outcomes. The use of PCBs has previously been associated with potentially adverse effects on thyroid function, resulting in a contraindication of PCBs for patients with hyper and hypothyroidism [16]. Notably, cluster 15, where PCBs showed the highest mortality rate, was primarily characterized by a high share of female patients, in line with the prevalence of thyroid disorders being approximately 10 times higher in women than in men [74]. Similarly, PCB administration have been associated with acute respiratory failure, and a consequent increased mortality [9]. Such observations could

also be made in the results, where cluster 13 demonstrated a particularly high share of respiratory disorders and a relatively high mortality when exposed to PCBs.

Similarly, prophylactic MgS administration has previously been shown to reduce the incidence of AF following cardiac surgery [41], and improved rhythm control and rate control rates [145]. MgS is, however, contraindicated if severe renal failure is present. The obtained results similarly show an increased mortality following MgS exposure in clusters characterized by a high rate of renal insufficiency. MgS did not outperform other treatment groups in terms of cardioversion in any cluster. Several phenotypes showed high rate control rates when exposed to MgS. In particular, patients with thyroid disorders appeared to have increased rate control rates, as well as the lowest mortality rates following MgS administration. While the mechanisms governing thyroid disease and the magnesium metabolism remain to be elucidated, recent work suggests a normalization of thyroid hormones following magnesium supplementation in community cohorts [149], which may explain the observation made in these phenotypes.

Several studies have reported similar conversion rates for BBs and PCBs [53]. This finding could only be confirmed for a single phenotype, characterized by myocardial infarction, renal insufficiency, and particularly advanced age. Similar observations have been made in several community studies, which demonstrated a reduction in mortality in patients receiving BBs following MI [114].

Conflicting evidence exists over the superiority of PCBs over CCBs in rhythm control [17, 109]. While the presented analysis showed a general superiority of PCBs over CCBs, the opposite was true for phenotypes characterized by a high prevalence of thyroid disorders and share of female patients. Similar observations can be made for postoperative patients with renal failure and coronary atherosclerosis, who further show a high mortality when treated with BBs.

Several phenotypes and treatment effects identified which demonstrated ATEs in accordance with existing clinical knowledge. These results support the conclusion that S-HAC not only forms predictive clusters, as demonstrated in the benchmark study, but also clusters which are biologically meaningful, and clinically useful. The extent to which the obtained results may provide insight into treatment effects that are currently unknown, drive hypothesis generation, or provide a basis for clinical decision-making remains to be determined. Ultimately, such endeavors warrant a validation of the obtained results using a secondary database.

Further, the obtained results should be interpreted with several limitations in mind. First, while the systematic review revealed many biomarkers to be viable predictors, only a small subset were available in the database. Given the crucial role of these variables in clustering algorithms, the inclusion of further biomarkers may provide different results. Further, treatment groups were assigned based on the first exposure to an AAD in the investigated groups. Further AADs are often employed, but the sparsity did not allow for their inclusion in the analysis. Also, treatments may be used in combination or incrementally escalated until the desired effect is achieved. Ultimately, a larger database may provide the basis for a more in-depth analysis with further treatment groups, and allow for the analysis of treatment combinations, while also providing more robust ATE estimates. Treatments were, further, considered as binary events in which a treatment was administered or not. The mode of administration, such as bolus or drip administration may have different effects, while the dosing also plays a crucial role [171].

6.4 User Centered Evaluation

The user centered evaluation generally showed positive results and supports the idea of using the developed methods in clinical practice. Several important topics were identified in the interviews with regards to the current treatment of AF in the ICU, specific requirements for AF treatment selection, as well as the validity of the developed methods and other CDSS of similar nature.

6.4.1 Current Treatment Approaches

Treatment of AF in the participating institution was primarily driven by ESC guidelines, which were adapted to the specific situation based on clinicians' discretion. Similarly to the majority of ICUs [225], internal guidelines for AF were not present in the participating institution. Decisions on anticoagulation were often guided by risk scores, which were derived from community cohorts, and not validated for use in the ICU. The use of such risk scores has, however, been shown to be common practice among critical care physicians [225] even though their applicability has been questioned [218].

The ESC guidelines were considered to be simple and easy to follow. While this simplicity was appreciated and viewed positively, especially with regards to facilitating quick decision-making, it was also considered to be a limiting factor, because patient complexity was not accounted for, which was of particular concern given ICU patients' multicomborbid nature. While this patient com-

plexity was considered problematic by some participants, in particular coupled with the large number of treatment options, others considered it a manageable challenge given sufficient experience. It is this patient complexity that the methods developed in this thesis aim to unravel, to provide estimates of treatment effects on a patient level basis.

6.4.2 Requirements for Treatment Selection

The selection of AF treatment was considered an act of balancing positive and negative impacts of possible treatments. The main treatment goal was to maximize the probability of positive effects while minimizing the probability of negative consequences. Treatments would only be considered if their expected benefit would outweigh their risks. Some participants mentioned specific ways to quantify outcomes, referring to "probability" and "speed", which are concepts that are explicitly captured by the survival models in the conceptualized CDSS.

Participants were particularly concerned with preventing stroke, bleeding events, restoring and maintaining SR, preventing mortality, controlling heart rate and patients' symptoms, while minimizing side effects and other potential harm. These suggested outcomes were significantly broader than the outcomes covered by the presented CDSS, which only covers the restoration of SR, the restoration of a normal heart rate, and mortality. Table 6.1 provides an overview of the outcomes suggested by the study participants and the outcomes covered in the CDSS. Even though many outcomes are not covered, overlap may exist between patients' mortality and ischemic stroke, bleeding events, and side effects. To what extent such an assumption can be made, is, however, uncertain. The gaps in outcomes reveal the complexity of decision-making in clinical practice, which is not fully accounted for, and must be resolved if such CDSS is to be used in practice.

Suggested by Participants	Covered by CDSS
Ischemic Stroke	No
Bleeding	No
Mortality	Yes
Restoration of SR	Yes
Maintenance of SR	No
Control of Heart Rate	Yes
Symptoms	No
Side Effects	No

Table 6.1: Outcomes suggested by study participants vs. covered by CDSS

Participants' confidence in being able to select the best treatment option was mixed, with the level of experience being cited as a key factor. The proposed CDSS was therefore of particular interest to less experienced participants, even though their more experienced counterparts suggested that it may be useful to resolve disagreements between more experienced clinicians. Similar associations have been reported by other studies, which observed an increased perceived usefulness of CDSS in users with limited clinical experience [125, 12, 129].

6.4.3 CDSS Evaluation

The presented CDSS was generally well received, with one participant stating they felt like it reminded them of their own though process. This statement supported the fulfillment of the first design principle of CDSS, stating that CDSS should mimic the cognitive processes of decision-makers. Participants agreed that the developed methods could support their decision-making and increase their efficiency, with quantitative results agreeing with participant statements.

The use of the proposed CDSS was considered plausible in practice. Immediate action was seldom required when patients developed AF. Nonetheless, quick accessibility was considered an important factor, with one participant suggesting the CDSS be deployed in the form of an application on a tablet. The use of tablets is a popular choice due to the immediate accessibility of the CDSS at the bedside [235, 111]. A major obstacle was the necessity for users to input data manually, which resulted in a high perceived effort. Further, the idea of a standalone system was questioned due to an already high number of independent technological platforms being used. It was therefore recom-

mended to implement CDSS as components of already employed information systems, similarly to other studies evaluating CDSS usability [153].

The specific requirements for such integration need to be elucidated in future works.

Further, gaps were identified in the variables used for patient description, which were incomplete or incorrectly captured. While variable selection was performed based on a systematic literature review, the user centered evaluation has revealed that the involvement of domain experts is indispensable to provide quality decision support. At the same time, required patient variables are often not present or sparsely available in EHR databases. An example thereof is the condition of systolic heart failure, which is a contraindication for CCBs. Within the database, a distinction between systolic and diastolic heart failure is, however, seldom made, and surrogate variables for the distinction between the two conditions are sparse. This makes the distinction between the two conditions difficult and highlights a limitation of using EHR databases for research.

As another participant noted, patients' evolution was not considered. In practice, patients evolve as new information is obtained and patient parameters are updated over the course of their ICU stay. Within the scope of this work, patients were described using the first available record, likely resulting in patient descriptions being outdated, suffering from look-ahead bias, or both. While this limitation may partially be remedied with more precise data extraction, limitations of the used database will remain. In particular, diagnostic codes suffer from inexact time resolution due to their creation at the time of patients' discharge.

Further limitations were identified in the presentation of information. As one participant noted, the central element of the CDSS, the point cloud visualizing the patients in the training database, added no informational value. While it may be considered a centerpiece of the previous analyses, it does not contribute in the context of providing decision support. On the other hand, a visualization of each treatment's utility in the form of a bar chart was desired. Another participant noted that treatment effects needed to be presented including their respective confidence intervals, which may easily be provided. Participants' comments regarding the presentation of information suggest that data visualization is a relevant task that requires the input of stakeholders and end-users in the design process.

The centered design evaluation has resulted in positive responses from participants, while highlighting key limitations of the proposed CDSS. Key patient outcomes were not covered, and patients' descriptive variables showed potential for refinement. While some patient variables and outcomes could be easily introduced, others may not be available in the used database. The developed methods assumed patients to be static and characterized them based on the first available record for each variable, patient evolution needs to be considered, requiring a database with improved temporal resolution. Ideally, such a database would contain the necessary descriptive variables to fully capture patient characteristics as required in clinical practice. In spite of the many limitations, the general clinician consensus was positive, supporting the idea of a data-driven phenotype classification approach for clinical decision support. These results should, however, be interpreted considering that this study was performed at a single institution and with a limited number of study participants. A similar study with multiple institutions and more participants may provide more conclusive results.

Chapter 7

Conclusion

7.1 Revision of Study Objectives

The aim of this doctoral thesis was the investigation of a phenotype classification based CDSS for the multiparametric stratification of AF patients in critical care. The fundamental problem that was addressed was the heterogeneity of the patient population, and lack of consensus among healthcare providers in which treatment would be most appropriate on a patient-level basis.

Prior works have suggested the use of data-driven phenotype classifications as a decision support tool for AF treatment. This thesis has contributed to this idea in several ways:

- Previous works investigating data-driven AF phenotypes have considered only outcome rates as relevant information. This thesis has demonstrated that beyond outcome rates, average treatment effects vary across phenotypes.
- AF phenotypes have previously exclusively been identified using unsupervised clustering algorithms, which have been criticized for forming clusters with little clinical significance. This thesis provided a novel semi-supervised clustering algorithm and the first application of such to a

cohort of AF patients. The resulting clusters have demonstrated treatment effects that were in line with existing clinical knowledge, and have uncovered phenotypes with unexpected average treatment effects.

- The use of a data-driven phenotype classification for guiding treatment selection was evaluated in a user centered evaluation. The results support the potential of a phenotype classification approach to provide decision support in clinical practice.

A novel semi-supervised clustering algorithm, S-HAC, was developed that incorporated survival data into the clustering process. This algorithm outperformed existing solutions on several real-world datasets, and elucidated the composition of the AF population in a clinically useful fashion. The identified phenotypes often aligned with existing knowledge of treatment effects such as contraindications and adverse effects, demonstrating that the developed algorithm does not only outperform existing methods quantitatively, but generates clusters which can readily be interpreted and that are of clinical utility.

The evaluation of the proposed CDSS in a user centered evaluation showed favorable outcomes for the application of the developed methods in clinical practice. Nonetheless, many limitations were identified. Partially, these limitations may be easily remedied, while others likely require substantial effort. It can be concluded that the development of CDSS for the multi-parametric stratification of AF patients require the involvement from domain experts from the initial stages. To what extent EHR databases provide data of sufficient quality to develop sufficiently accurate models remains questionable.

7.2 Limitations

The conclusions of this work should be interpreted in the context of several limitations, of which some have already been stated. The systematic literature review performed in section 4.1 may provide limited information, as critical biomarkers predictive of AF outcomes may not have been investigated in previous works. Similarly, the quality of the identified studies has not been assessed, and the strength of the evidence for biomarkers being predictive has not been determined. While some biomarkers may be predictive in certain socioeconomic or ethnic groups, they may not be valid in the analyzed population.

The developed algorithms S-HAC and LSS-K-Means have shown promising performance in benchmarks studies, but their performance needs to be val-

idated in further datasets to allow for more conclusive results, as well as a better understanding of the advantages and limitations of each method. Further, the benchmark tests were performed on datasets with a single outcome, while the analysis of AF phenotypes used a combination of outcomes for cluster formation.

S-HAC has identified AF phenotypes that corresponded to current knowledge on the treatment effects of antiarrhythmic drugs. Further databases Nonetheless, the resulting solution should not be considered in any way final, nor can it be considered a proposal for a formal classification. The presented study has only considered a limited number of patient variables, and the inclusion of further variables will likely lead to the discovery of different patient phenotypes. Similarly, only a limited number of treatments were considered, and treatment groups were considered to be mutually exclusive. In reality treatments may be combined or changed until the desired effect is achieved.

The user centered evaluation has only evaluated a mock-up of a CDSS, while no functioning tool has been tested. The results of the study must therefore be treated with caution, as the evaluation of a working CDSS may uncover additional facilitators and obstacles in the user acceptance of such tool. More critically, potential impacts on clinicians' performance must be measured to provide conclusive results on the applicability of such a CDSS in clinical practice. Such studies, however, would require a prior validation of the phenotype classification results to ensure patient safety. Finally, it must be considered that the results of the user centered evaluation may suffer from bias due to the single-center design, and a relatively small sample size.

7.3 Future Work

This thesis has provided a foundation for further research. In particular, the phenotypes obtained using S-HAC are sufficiently encouraging to consider a validation study using a secondary dataset. Such validation could shed further light on the quality of the results obtained, and identify which observed treatment effects are worth pursuing in further research endeavors, which may be rooted in a variety of disciplines, and provide opportunity for multidisciplinary cooperation. Unexpected patient phenotypes with significantly poorer outcomes may be validated with secondary datasets, and provide sufficient empirical support to further refine treatment selection in AF patients.

The user research analysis has shown favorable results, but several obstacles need to be considered before a usable prototype can be implemented and eval-

uated. Critical biomarkers used for treatment selection were not present in the employed dataset, likely resulting in residual confounding and biased results. The analysis has further shown a significant lack of coverage in relevant outcomes. Future work should therefore consider the inclusion of critical variables from the first stages of the analysis, and the evaluation of further patient outcomes. A standalone system has been criticized as a potential obstruction in clinical workflows, making an integration of the CDSS into existing IT infrastructure a fundamental necessity for clinical usability.

Further improvements to the developed methods may ultimately provide a CDSS, which could be evaluated in a prospective trial. Such a prospective trial could be executed in the form of a single center randomized controlled fashion, in which the control arm of the study group receives standard care, while the intervention arm receives standard care supplemented with the CDSS. Several outcomes may be measured to assess the impact of the CDSS, such as the time until cardioversion or rate control, or mortality. These direct indicators of treatment effects would be expected to be more favorable in the intervention arm than the control arm. Further, adherence to system recommendations, the total antiarrhythmic costs, or the number of different antiarrhythmics administered per patient may be evaluated to assess the impact of CDSS use.

7.4 Scientific Contributions

7.4.1 Journal Publications

Lacki A, Martinez-Millana A. A Comparison of the Impact of Pharmacological Treatments on Cardioversion, Rate Control, and Mortality in Data-Driven Atrial Fibrillation Phenotypes in Critical Care. Bioengineering. 2024; 11(3):199. https://doi.org/10.3390/bioengineering11030199

7.4.2 International Conferences

Lacki, A., Martinez-Millana, A. (2023). Survival Hierarchical Agglomerative Clustering: A Semi-Supervised Clustering Method Incorporating Survival Data. In: Juarez, J.M., Marcos, M., Stiglic, G., Tucker, A. (eds) Artificial Intelligence in Medicine. AIME 2023. Lecture Notes in Computer Science(), vol 13897. Springer, Cham. https://doi.org/10.1007/978-3-031-34344-5_1

Lacki, A., Boscá Tomás D., Martínez-Millana A. (2022, September). Probabilistic Inference of Comorbidities from Symptoms in Patients with Atrial

Fibrillation: An Ontology-Driven Hybrid Clinical Decision Support System. In 2022 Computing in Cardiology (Vol. 498, pp. 1-4). IEEE.

Lacki, A., Hernández-Romero, I., Guillem, M. S., Climent, A. M. (2021, September). ECGI Periodicity Unraveled: A Deep Learning Approach for the Visualization of Periodic Spatiotemporal Patterns in Atrial Fibrillation Patients. In 2021 Computing in Cardiology (Vol. 48, pp. 1-4). IEEE.

Bibliography

[1] Niv Ad et al. "The Need for Consistent Predictors of Success for Surgical Ablation of Atrial Fibrillation". In: *Innovations: Technology and Techniques in Cardiothoracic and Vascular Surgery* 12 (6 Nov. 2017), pp. 421–429. ISSN: 1556-9845. DOI: 10.1097/imi.0000000000000426 (cit. on p. 91).

[2] Nazem Akoum et al. "Atrial Fibrosis Helps Select the Appropriate Patient and Strategy in Catheter Ablation of Atrial Fibrillation: A DE-MRI Guided Approach". In: *Journal of Cardiovascular Electrophysiology* 22 (1 Jan. 2011), pp. 16–22. ISSN: 10453873. DOI: 10.1111/j.1540-8167.2010.01876.x (cit. on p. 13).

[3] B Aldhoon et al. "New insights into mechanisms of atrial fibrillation". In: *Physiological Research* (Feb. 2010), pp. 1–12. ISSN: 1802-9973. DOI: 10.33549/physiolres.931651 (cit. on p. 14).

[4] Maurits A. Allessie et al. "Electropathological Substrate of Long-Standing Persistent Atrial Fibrillation in Patients With Structural Heart Disease". In: *Circulation: Arrhythmia and Electrophysiology* 3 (6 Dec. 2010), pp. 606–615. ISSN: 1941-3149. DOI: 10.1161/CIRCEP.109.910125 (cit. on p. 13).

[5] Alvaro Alonso et al. "Simple Risk Model Predicts Incidence of Atrial Fibrillation in a Racially and Geographically Diverse Population: the

CHARGE-AF Consortium". In: *Journal of the American Heart Association* 2.2 (Mar. 2013). ISSN: 2047-9980. DOI: 10.1161/JAHA.112.000102 (cit. on p. 25).

[6] P. Arabie, L.J. Hubert, and G. de Soete. *Clustering and Classification*. World Scientific, 1996. ISBN: 9789810212872 (cit. on p. 28).

[7] Doron Aronson et al. "Risk Score for Prediction of 10-Year Atrial Fibrillation: A Community-Based Study". In: *Thrombosis and Haemostasis* 118.09 (Sept. 2018), pp. 1556–1563. ISSN: 0340-6245. DOI: 10.1055/s-0038-1668522 (cit. on p. 25).

[8] Stephane Arques, Emmanuel Roux, and Roger Luccioni. "Current clinical applications of spectral tissue Doppler echocardiography (E/E' ratio) as a noninvasive surrogate for left ventricular diastolic pressures in the diagnosis of heart failure with preserved left ventricular systolic function". In: *Cardiovascular Ultrasound* 5 (1 Dec. 2007), p. 16. ISSN: 1476-7120. DOI: 10.1186/1476-7120-5-16 (cit. on p. 93).

[9] Houman Ashrafian and Patrick Davey. "Is Amiodarone an Underrecognized Cause of Acute Respiratory Failure in the ICU?" In: *Chest* 120 (1 July 2001), pp. 275–282. ISSN: 00123692. DOI: 10.1378/chest.120.1.275 (cit. on p. 121).

[10] Hideki Asoh et al. "An Application of Inverse Reinforcement Learning to Medical Records of Diabetes Treatment". In: 2013 (cit. on p. 35).

[11] P. S. Badase, G. P. Deshbhratar, and A. P. Bhagat. "Classification and analysis of clustering algorithms for large datasets". In: IEEE, Mar. 2015, pp. 1–5. ISBN: 978-1-4799-6817-6. DOI: 10.1109/ICIIECS.2015.7193191 (cit. on pp. 29, 30).

[12] Mirza Mansoor Baig et al. "Clinical decision support systems in hospital care using ubiquitous devices: Current issues and challenges". In: *Health Informatics Journal* 25 (3 Sept. 2019), pp. 1091–1104. ISSN: 1460-4582. DOI: 10.1177/1460458217740722 (cit. on p. 125).

[13] Eric Bair. "Semi-supervised clustering methods". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 5 (5 Sept. 2013), pp. 349–361. ISSN: 19395108 (cit. on pp. 30, 56).

[14] Eric Bair and Robert Tibshirani. "Semi-Supervised Methods to Predict Patient Survival from Gene Expression Data". In: *PLoS Biology* 2 (4 Apr. 2004), e108. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.0020108 (cit. on pp. 32, 33, 56, 64).

[15] Adam S. Barnett et al. "Treatment of Atrial Fibrillation and Concordance With the American Heart Association/American College of Cardiology/Heart Rhythm Society Guidelines". In: *Circulation: Arrhythmia and Electrophysiology* 10 (11 Nov. 2017). ISSN: 1941-3149. DOI: 10.1161/CIRCEP.117.005051 (cit. on p. 39).

[16] Shehzad Basaria and David S. Cooper. "Amiodarone and the thyroid". In: *The American Journal of Medicine* 118 (7 July 2005), pp. 706–714. ISSN: 00029343. DOI: 10.1016/j.amjmed.2004.11.028 (cit. on p. 121).

[17] Jonathan P. Bedford et al. "Comparative effectiveness of common treatments for new-onset atrial fibrillation within the ICU: Accounting for physiological status". In: *Journal of Critical Care* 67 (Feb. 2022), pp. 149–156. ISSN: 08839441. DOI: 10.1016/j.jcrc.2021.11.005 (cit. on pp. 68, 69, 122).

[18] Sahar Behpour and Avi Udash. "Understanding Machine Learning Through Data-Oriented and Human Learning Approaches". In: 2022, pp. 3–29. DOI: 10.1007/978-3-030-84729-6_1 (cit. on p. 27).

[19] Emelia J. Benjamin et al. "Impact of Atrial Fibrillation on the Risk of Death". In: *Circulation* 98 (10 Sept. 1998), pp. 946–952. ISSN: 0009-7322. DOI: 10.1161/01.CIR.98.10.946 (cit. on p. 11).

[20] Mikhail Bilenko, Sugato Basu, and Raymond J. Mooney. "Integrating constraints and metric learning in semi-supervised clustering". In: ACM Press, 2004, p. 11. DOI: 10.1145/1015330.1015360 (cit. on p. 53).

[21] Todd E. Bodner. "What Improves with Increased Missing Data Imputations?" In: *Structural Equation Modeling: A Multidisciplinary Journal* 15 (4 Oct. 2008), pp. 651–675. ISSN: 1070-5511. DOI: 10.1080/10705510802339072 (cit. on p. 70).

[22] Vennela Boyalla et al. "Biomarkers as predictors of recurrence of atrial fibrillation post ablation: an updated and expanded systematic review

and meta-analysis". en. In: *Clin. Res. Cardiol.* (Jan. 2022) (cit. on p. 119).

[23] N. Brasier et al. "The quest for indicators of paroxysmal atrial fibrillation in sinus rhythm – the DETECT AF trial". In: *Acta Cardiologica* 74 (4 July 2019), pp. 301–307. ISSN: 0001-5385. DOI: 10.1080/00015385.2018.1493248 (cit. on p. 91).

[24] Ramon Brugada et al. "Identification of a Genetic Locus for Familial Atrial Fibrillation". In: *New England Journal of Medicine* 336 (13 Mar. 1997), pp. 905–911. ISSN: 0028-4793. DOI: 10.1056/NEJM199703273361302 (cit. on p. 11).

[25] Stef van Buuren and Karin Groothuis-Oudshoorn. "mice: Multivariate Imputation by Chained Equations in R". In: *Journal of Statistical Software* 45 (3 2011). ISSN: 1548-7660. DOI: 10.18637/jss.v045.i03 (cit. on p. 70).

[26] Patrizia Camelliti et al. "Fibroblast Network in Rabbit Sinoatrial Node". In: *Circulation Research* 94 (6 Apr. 2004), pp. 828–835. ISSN: 0009-7330. DOI: 10.1161/01.RES.0000122382.19400.14 (cit. on p. 7).

[27] A. J. Camm et al. "Guidelines for the management of atrial fibrillation: The Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC)". In: *European Heart Journal* 31 (19 Oct. 2010), pp. 2369–2429. ISSN: 0195-668X. DOI: 10.1093/eurheartj/ehq278 (cit. on p. 1).

[28] Uğur Canpolat et al. "A proposal for a new scoring system in the prediction of catheter ablation outcomes: Promising results from the Turkish Cryoablation Registry". In: *International Journal of Cardiology* 169.3 (Nov. 2013), pp. 201–206. ISSN: 01675273. DOI: 10.1016/j.ijcard.2013.08.097 (cit. on p. 26).

[29] Jaime G. Carbonell, Ryszard S. Michalski, and Tom M. Mitchell. "AN OVERVIEW OF MACHINE LEARNING". In: Elsevier, 1983, pp. 3–23. DOI: 10.1016/B978-0-08-051054-5.50005-4 (cit. on pp. 27, 30).

[30] Matthew A. Carlisle et al. "Heart Failure and Atrial Fibrillation, Like Fire and Fury". In: *JACC: Heart Failure* 7 (6 June 2019), pp. 447–456. ISSN: 22131779. DOI: 10.1016/j.jchf.2019.03.005 (cit. on p. 90).

[31] Gerard M. Castro, Lisa Buczkowski, and Joanne M. Hafner. "The Contribution of Sociotechnical Factors to Health Information Technology–Related Sentinel Events". In: *The Joint Commission Journal on Quality and Patient Safety* 42 (2 Feb. 2016), 70–AP3. ISSN: 15537250. DOI: 10.1016/S1553-7250(16)42008-8 (cit. on p. 35).

[32] Paidamoyo Chapfuwa et al. "Survival cluster analysis". In: *Proceedings of the ACM Conference on Health, Inference, and Learning* (Apr. 2020), pp. 60–68. DOI: 10.1145/3368555.3384465 (cit. on pp. 56, 121).

[33] Sumeet S. Chugh et al. "Worldwide Epidemiology of Atrial Fibrillation". In: *Circulation* 129 (8 Feb. 2014), pp. 837–847. ISSN: 0009-7322. DOI: 10.1161/CIRCULATIONAHA.113.005119 (cit. on pp. 1, 11).

[34] Benjamin Clayton et al. "Risk of thromboembolism in patients developing critical illness-associated atrial fibrillation". In: *Clinical Medicine* 18 (4 Aug. 2018), pp. 282–287. ISSN: 1470-2118. DOI: 10.7861/clinmedicine.18-4-282 (cit. on p. 1).

[35] Jamie J Coleman et al. "On the alert: future priorities for alerts in clinical decision support for computerized physician order entry identified from a European workshop". In: *BMC Medical Informatics and Decision Making* 13 (1 Dec. 2013), p. 111. ISSN: 1472-6947. DOI: 10.1186/1472-6947-13-111 (cit. on p. 34).

[36] Francis S. Collins and Harold Varmus. "A New Initiative on Precision Medicine". In: *New England Journal of Medicine* 372 (9 Feb. 2015), pp. 793–795. ISSN: 0028-4793. DOI: 10.1056/NEJMp1500523 (cit. on p. 33).

[37] D. R. Cox. "Regression Models and Life-Tables". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 34 (2 Jan. 1972), pp. 187–202. ISSN: 00359246. DOI: 10.1111/j.2517-6161.1972.tb00899.x (cit. on p. 56).

[38] Jafna L. Cox et al. "Optimizing primary care management of atrial fibrillation: The rationale and methods of the Integrated Management Program Advancing Community Treatment of Atrial Fibrillation (IMPACT-AF) study". In: *American Heart Journal* 201 (July 2018), pp. 149–157. ISSN: 00028703. DOI: 10.1016/j.ahj.2018.04.008 (cit. on pp. 44, 46).

[39] Christina Curtis et al. "The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups". In: *Nature* 486 (7403 June 2012), pp. 346–352. ISSN: 0028-0836. DOI: 10.1038/nature10983 (cit. on p. 63).

[40] Lori B. Daniels et al. "Influence of age, race, sex, and body mass index on interpretation of midregional pro atrial natriuretic peptide for the diagnosis of acute heart failure: results from the BACH multinational study". In: *European Journal of Heart Failure* 14 (1 Jan. 2012), pp. 22–31. ISSN: 13889842. DOI: 10.1093/eurjhf/hfr157 (cit. on pp. 92, 119).

[41] Michael John Davey and David Teubner. "A Randomized Controlled Trial of Magnesium Sulfate, in Addition to Usual Care, for Rate Control in Atrial Fibrillation". In: *Annals of Emergency Medicine* 45 (4 Apr. 2005), pp. 347–353. ISSN: 01960644. DOI: 10.1016/j.annemergmed.2004.09.013 (cit. on p. 122).

[42] Ian Davidson and S. S. Ravi. "Clustering With Constraints: Feasibility Issues and the k -Means Algorithm". In: Society for Industrial and Applied Mathematics, Apr. 2005, pp. 138–149. ISBN: 978-0-89871-593-4. DOI: 10.1137/1.9781611972757.13 (cit. on p. 53).

[43] Fred D. Davis. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology". In: *MIS Quarterly* 13 (3 Sept. 1989), p. 319. ISSN: 02767783. DOI: 10.2307/249008 (cit. on p. 72).

[44] Thomas A. Dewland et al. "Inflammation as a Mediator of the Association Between Race and Atrial Fibrillation". In: *JACC: Clinical Electrophysiology* 1 (4 Aug. 2015), pp. 248–255. ISSN: 2405500X. DOI: 10.1016/j.jacep.2015.04.014 (cit. on p. 14).

[45] D Dick. "Mechanical modulation of stretch-induced premature ventricular beats: induction of a mechanoelectric adaptation period". In: *Car-*

diovascular Research 38 (1 Apr. 1998), pp. 181–191. ISSN: 00086363. DOI: 10.1016/S0008-6363(97)00314-3 (cit. on p. 13).

[46] Eric Y. Ding et al. "Novel Method of Atrial Fibrillation Case Identification and Burden Estimation Using the MIMIC-III Electronic Health Data Set". In: *Journal of Intensive Care Medicine* 34 (10 Oct. 2019), pp. 851–857. ISSN: 0885-0666. DOI: 10.1177/0885066619866172 (cit. on p. 68).

[47] L Ding et al. "Incidence of atrial fibrillation and its risk prediction model based on a prospective urban Han Chinese cohort". In: *Journal of Human Hypertension* 31.9 (Sept. 2017), pp. 574–579. ISSN: 0950-9240. DOI: 10.1038/jhh.2017.23 (cit. on p. 25).

[48] Mark T. Dishaw and Diane M. Strong. "Supporting software maintenance with software engineering tools: A Computed task–technology fit analysis". In: *Journal of Systems and Software* 44 (2 Dec. 1998), pp. 107–120. ISSN: 01641212. DOI: 10.1016/S0164-1212(98)10048-1 (cit. on p. 83).

[49] Angela Dispenzieri et al. "Use of Nonclonal Serum Immunoglobulin Free Light Chains to Predict Overall Survival in the General Population". In: *Mayo Clinic Proceedings* 87 (6 June 2012), pp. 517–523. ISSN: 00256196 (cit. on p. 63).

[50] Ugljesa Djuric et al. "Precision histology: how deep learning is poised to revitalize histomorphology for personalized cancer care". In: *npj Precision Oncology* 1 (1 Dec. 2017), p. 22. ISSN: 2397-768X. DOI: 10.1038/s41698-017-0022-1 (cit. on p. 33).

[51] S. van Doorn et al. "Effectiveness of CHA2DS2-VASc based decision support on stroke prevention in atrial fibrillation: A cluster randomised trial in general practice". In: *International Journal of Cardiology* 273 (Dec. 2018), pp. 123–129. ISSN: 01675273. DOI: 10.1016/j.ijcard.2018.08.096 (cit. on pp. 39, 46).

[52] Derek J. Dosdall, Vladimir G. Fast, and Raymond E. Ideker. "Mechanisms of Defibrillation". In: *Annual Review of Biomedical Engineering* 12 (1 July 2010), pp. 233–258. ISSN: 1523-9829. DOI: 10.1146/annurev-bioeng-070909-105305 (cit. on p. 17).

[53] Laura Drikite et al. "Treatment strategies for new onset atrial fibrillation in patients treated on an intensive care unit: a systematic scoping review". In: *Critical Care* 25 (1 Dec. 2021), p. 257. ISSN: 1364-8535. DOI: 10.1186/s13054-021-03684-5 (cit. on pp. 2, 122).

[54] Arita Dubnika et al. "Cytokines as therapeutic agents and targets in heart disease". In: *Cytokine & Growth Factor Reviews* 43 (Oct. 2018), pp. 54–68. ISSN: 13596101. DOI: 10.1016/j.cytogfr.2018.08.003 (cit. on p. 7).

[55] Tlamelo Emmanuel et al. "A survey on missing data in machine learning". In: *Journal of Big Data* 8 (1 Dec. 2021), p. 140. ISSN: 2196-1115. DOI: 10.1186/s40537-021-00516-9 (cit. on pp. 69, 70).

[56] Jesper E. van Engelen and Holger H. Hoos. "A survey on semi-supervised learning". In: *Machine Learning* 109 (2 Feb. 2020), pp. 373–440. ISSN: 0885-6125. DOI: 10.1007/s10994-019-05855-6 (cit. on p. 30).

[57] Ricardo J Esper et al. "Endothelial dysfunction: a comprehensive appraisal". In: *Cardiovascular Diabetology* 5 (1 2006), p. 4. ISSN: 14752840. DOI: 10.1186/1475-2840-5-4 (cit. on p. 6).

[58] Brendan M. Everett et al. "Novel genetic markers improve measures of atrial fibrillation risk prediction". In: *European Heart Journal* 34.29 (Aug. 2013), pp. 2243–2251. ISSN: 0195-668X. DOI: 10.1093/eurheartj/eht033 (cit. on p. 25).

[59] Bruna Ferreira, Tayana Conte, and Simone Diniz Junqueira Barbosa. "Eliciting Requirements Using Personas and Empathy Map to Enhance the User Experience". In: *2015 29th Brazilian Symposium on Software Engineering*. IEEE, Sept. 2015, pp. 80–89. ISBN: 978-1-4673-9272-3. DOI: 10.1109/SBES.2015.14 (cit. on p. 73).

[60] Bruna Ferreira et al. "Designing Personas with Empathy Map". In: July 2015, pp. 501–505. DOI: 10.18293/SEKE2015-152 (cit. on p. 73).

[61] Giuseppe Fico et al. "What do healthcare professionals need to turn risk models for type 2 diabetes into usable computerized clinical decision support systems? Lessons learned from the MOSAIC project". In: *BMC*

Medical Informatics and Decision Making 19 (1 Dec. 2019), p. 163. ISSN: 1472-6947. DOI: 10.1186/s12911-019-0887-8 (cit. on p. 73).

[62] Florin Mihail Filipoiu. *Atlas of Heart Anatomy and Development*. Springer London, 2014. ISBN: 978-1-4471-5381-8. DOI: 10.1007/978-1-4471-5382-5 (cit. on p. 7).

[63] Evelyn Fix and J. L. Hodges. "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties". In: *International Statistical Review / Revue Internationale de Statistique* 57 (3 Dec. 1989), p. 238. ISSN: 03067734. DOI: 10.2307/1403797 (cit. on p. 70).

[64] Greg C. Flaker et al. "Antiarrhythmic drug therapy and cardiac mortality in atrial fibrillation". In: *Journal of the American College of Cardiology* 20.3 (Sept. 1992), pp. 527–532. ISSN: 07351097. DOI: 10.1016/0735-1097(92)90003-6 (cit. on p. 2).

[65] A J Fletcher, S Robinson, and B S Rana. "Echocardiographic RV-E/e' for predicting right atrial pressure: a review". In: *Echo Research and Practice* 7 (4 Dec. 2020), R11–R20. ISSN: 2055-0464. DOI: 10.1530/ERP-19-0057 (cit. on p. 93).

[66] Keith A A Fox et al. "Improved risk stratification of patients with atrial fibrillation: an integrated GARFIELD-AF tool for the prediction of mortality, stroke and bleed in patients with and without anticoagulation". In: *BMJ Open* 7.12 (Dec. 2017), e017157. ISSN: 2044-6055. DOI: 10.1136/bmjopen-2017-017157 (cit. on p. 22).

[67] Liana Fraenkel, Richard L Street, and Terri R Fried. "Development of a tool to improve the quality of decision making in atrial fibrillation". In: *BMC Medical Informatics and Decision Making* 11 (1 Dec. 2011), p. 59. ISSN: 1472-6947. DOI: 10.1186/1472-6947-11-59 (cit. on pp. 38, 46).

[68] Liana Fraenkel et al. "A Pilot Randomized Controlled Trial of a Decision Support Tool to Improve the Quality of Communication and Decision-Making in Individuals with Atrial Fibrillation". In: *Journal of the American Geriatrics Society* 60 (8 Aug. 2012), pp. 1434–1441. ISSN: 00028614. DOI: 10.1111/j.1532-5415.2012.04080.x (cit. on p. 39).

[69] Piero Colli Franzone, Luca F. Pavarino, and Simone Scacchi. "Basic Cardiac Anatomy and Electrocardiology". In: 2014, pp. 1–19. DOI: 10.1007/978-3-319-04801-7_1 (cit. on p. 4).

[70] Christian Fromm et al. "Diltiazem vs. Metoprolol in the Management of Atrial Fibrillation or Flutter with Rapid Ventricular Rate in the Emergency Department". In: *The Journal of Emergency Medicine* 49 (2 Aug. 2015), pp. 175–182. ISSN: 07364679. DOI: 10.1016/j.jemermed.2015.01.014 (cit. on p. 68).

[71] Brian F. Gage et al. "Clinical classification schemes for predicting hemorrhage: Results from the National Registry of Atrial Fibrillation (NRAF)". In: *American Heart Journal* 151 (3 Mar. 2006), pp. 713–719. ISSN: 00028703. DOI: 10.1016/j.ahj.2005.04.017 (cit. on p. 24).

[72] Sheila Gaynor and Eric Bair. "Identification of biologically relevant subtypes via preweighted sparse clustering". In: *CoRR* abs/1304.3760 (2013). arXiv: 1304.3760 (cit. on pp. 32, 56, 57, 64).

[73] ISABELLE C. GELDER et al. "Alterations in Gene Expression of Proteins Involved in the Calcium Handling in Patients with Atrial Fibrillation". In: *Journal of Cardiovascular Electrophysiology* 10 (4 Apr. 1999), pp. 552–560. ISSN: 1045-3873. DOI: 10.1111/j.1540-8167.1999.tb00712.x (cit. on p. 15).

[74] Alois Gessl, Rosa Lemmens-Gruber, and Alexandra Kautzky-Willer. "Thyroid Disorders". In: 2013, pp. 361–386 (cit. on p. 121).

[75] Zahra Ghasemi, Hadi Akbarzadeh Khorshidi, and Uwe Aickelin. "Multi-objective Semi-supervised clustering for finding predictive clusters". In: *Expert Systems with Applications* 195 (June 2022), p. 116551. ISSN: 09574174. DOI: 10.1016/j.eswa.2022.116551 (cit. on p. 56).

[76] A. M. Gilroy et al. *Atlas of Anatomy*. Thieme, 2008 (cit. on pp. 3, 5).

[77] Dale L. Goodhue and Ronald L. Thompson. "Task-Technology Fit and Individual Performance". In: *MIS Quarterly* 19 (2 June 1995), p. 213. ISSN: 02767783. DOI: 10.2307/249689 (cit. on pp. 72, 74).

[78] A.T.Marcel Gosselink et al. "Changes in left and right atrial size after cardioversion of atrial fibrillation: Role of mitral valve disease". In: *Journal of the American College of Cardiology* 22 (6 Nov. 1993), pp. 1666–1672. ISSN: 07351097. DOI: 10.1016/0735-1097(93)90593-P (cit. on p. 15).

[79] Kevin G Graves et al. "Improving CHA ₂ DS ₂ -VASc stratification of non-fatal stroke and mortality risk using the Intermountain Mortality Risk Score among patients with atrial fibrillation". In: *Open Heart* 5.2 (Nov. 2018), e000907. ISSN: 2053-3624. DOI: 10.1136/openhrt-2018-000907 (cit. on p. 22).

[80] Robert A. Greenes et al. "Clinical decision support models and frameworks: Seeking to address research issues underlying implementation successes and failures". In: *Journal of Biomedical Informatics* 78 (Feb. 2018), pp. 134–143. ISSN: 15320464. DOI: 10.1016/j.jbi.2017.12.005 (cit. on p. 33).

[81] Yutao Guo et al. "Mobile Health (mHealth) technology for improved screening, patient involvement and optimising integrated care in atrial fibrillation: The mAFA (mAF-App) II randomised trial". In: *International Journal of Clinical Practice* 73 (7 July 2019). ISSN: 1368-5031. DOI: 10.1111/ijcp.13352 (cit. on pp. 43, 44, 46).

[82] Yutao Guo et al. "Mobile health technology-supported atrial fibrillation screening and integrated care: A report from the mAFA-II trial Long-term Extension Cohort". In: *European Journal of Internal Medicine* 82 (Dec. 2020), pp. 105–111. ISSN: 09536205. DOI: 10.1016/j.ejim.2020.09.024 (cit. on p. 43).

[83] Michel Haïssaguerre et al. "Spontaneous Initiation of Atrial Fibrillation by Ectopic Beats Originating in the Pulmonary Veins". In: *New England Journal of Medicine* 339 (10 Sept. 1998), pp. 659–666. ISSN: 0028-4793. DOI: 10.1056/NEJM199809033391003 (cit. on p. 11).

[84] Richiro Hamada and Shigeki Muto. "Simple risk model and score for predicting of incident atrial fibrillation in Japanese". In: *Journal of Cardiology* 73.1 (Jan. 2019), pp. 65–72. ISSN: 09145087. DOI: 10.1016/j.jjcc.2018.06.005 (cit. on p. 25).

[85] W L Henry et al. "Relation between echocardiographically determined left atrial size and atrial fibrillation." In: *Circulation* 53 (2 Feb. 1976), pp. 273–279. ISSN: 0009-7322. DOI: 10.1161/01.CIR.53.2.273 (cit. on p. 15).

[86] Ziad Hijazi et al. "A biomarker-based risk score to predict death in patients with atrial fibrillation: the ABC (age, biomarkers, clinical history) death risk score". en. In: *Eur. Heart J.* 39.6 (Feb. 2018), pp. 477–485 (cit. on p. 90).

[87] Ziad Hijazi et al. "The ABC (age, biomarkers, clinical history) stroke risk score: a biomarker-based risk score for predicting stroke in atrial fibrillation". In: *European Heart Journal* 37.20 (May 2016), pp. 1582–1590. ISSN: 0195-668X. DOI: 10.1093/eurheartj/ehw054 (cit. on p. 22).

[88] Ziad Hijazi et al. "The novel biomarker-based ABC (age, biomarkers, clinical history)-bleeding risk score for patients with atrial fibrillation: a derivation and validation study". In: *The Lancet* 387.10035 (June 2016), pp. 2302–2311. ISSN: 01406736. DOI: 10.1016/S0140-6736(16)00741-8 (cit. on p. 24).

[89] Gerhard Hindricks et al. "2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS)". In: *European Heart Journal* 42.5 (Feb. 2021), pp. 373–498. ISSN: 0195-668X. DOI: 10.1093/eurheartj/ehaa612 (cit. on pp. 11, 16–19, 21, 23, 24, 27, 31, 43, 106).

[90] Yu-Feng Hu et al. "Inflammation and the pathogenesis of atrial fibrillation". In: *Nature Reviews Cardiology* 12 (4 Apr. 2015), pp. 230–243. ISSN: 1759-5002. DOI: 10.1038/nrcardio.2015.2 (cit. on p. 90).

[91] CX Huang et al. "Current knowledge and management recommendations of atrial fibrillation: 2018". In: *Chin J Cardiac Arrhyth* 22.4 (2018), pp. 279–346 (cit. on p. 40).

[92] Jin-Long Huang et al. "Effect of atrial dilatation on electrophysiologic properties and inducibility of atrial fibrillation". In: *Basic Research in Cardiology* 98 (1 Jan. 2003), pp. 16–24. ISSN: 0300-8428. DOI: 10.1007/s00395-003-0385-z (cit. on p. 91).

[93] Paul A. Iaizzo. *Handbook of Cardiac Anatomy, Physiology, and Devices*. Ed. by Paul A. Iaizzo. Humana Press, 2009. ISBN: 978-1-60327-371-8. DOI: 10.1007/978-1-60327-372-5 (cit. on pp. 3, 4).

[94] Taku Inohara et al. "A Cluster Analysis of the Japanese Multicenter Outpatient Registry of Patients With Atrial Fibrillation". In: *The American Journal of Cardiology* 124 (6 Sept. 2019), pp. 871–878. ISSN: 00029149. DOI: 10.1016/j.amjcard.2019.05.071 (cit. on p. 32).

[95] Taku Inohara et al. "Association of of Atrial Fibrillation Clinical Phenotypes With Treatment Patterns and Outcomes". In: *JAMA Cardiology* 3 (1 Jan. 2018), p. 54. ISSN: 2380-6583. DOI: 10.1001/jamacardio.2017.4665 (cit. on pp. 31, 72).

[96] Ziad F. Issa, John M. Miller, and Douglas P. Zipes. "Molecular Mechanisms of Cardiac Electrical Activity". In: Elsevier, 2019, pp. 1–14. DOI: 10.1016/B978-0-323-52356-1.00001-3 (cit. on pp. 9–11, 13–15).

[97] Theodore J. Iwashyna et al. "Implications of Heterogeneity of Treatment Effect for Reporting and Analysis of Randomized Trials in Critical Care". In: *American Journal of Respiratory and Critical Care Medicine* 192 (9 Nov. 2015), pp. 1045–1051. ISSN: 1073-449X. DOI: 10.1164/rccm.201411-2125CP (cit. on p. 1).

[98] Jos Jalife et al. *Basic Cardiac Electrophysiology for the Clinician*. Wiley-Blackwell, May 2009. ISBN: 9781444316940. DOI: 10.1002/9781444316940 (cit. on p. 12).

[99] Hui Jiang et al. "Association of pre-ablation level of potential blood markers with atrial fibrillation recurrence after catheter ablation: a meta-analysis". In: *Europace* (July 2016), euw088. ISSN: 1099-5129. DOI: 10.1093/europace/euw088 (cit. on p. 92).

[100] Xin Jin and Jiawei Han. "K-Means Clustering". In: Springer US, 2011, pp. 563–564. DOI: 10.1007/978-0-387-30164-8_425 (cit. on p. 28).

[101] Alistair E. W. Johnson et al. "MIMIC-IV, a freely accessible electronic health record dataset". In: *Scientific Data* 10 (1 Jan. 2023), p. 1. ISSN: 2052-4463. DOI: 10.1038/s41597-022-01899-x (cit. on p. 65).

[102] Alistair E.W. Johnson et al. "MIMIC-III, a freely accessible critical care database". In: *Scientific Data* 3 (1 Dec. 2016), p. 160035. ISSN: 2052-4463. DOI: 10.1038/sdata.2016.35 (cit. on p. 65).

[103] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. "Reinforcement learning: A survey". In: *Journal of artificial intelligence research* 4 (1996), pp. 237–285 (cit. on p. 27).

[104] Rahul Kakkar and Richard T. Lee. "Intramyocardial Fibroblast Myocyte Communication". In: *Circulation Research* 106 (1 Jan. 2010), pp. 47–57. ISSN: 0009-7330. DOI: 10.1161/CIRCRESAHA.109.207456 (cit. on p. 7).

[105] A Kamkin. "Stretch-activated currents in ventricular myocytes: amplitude and arrhythmogenic effects increase with hypertrophy". In: *Cardiovascular Research* 48 (3 Dec. 2000), pp. 409–420. ISSN: 00086363. DOI: 10.1016/S0008-6363(00)00208-X (cit. on p. 13).

[106] W.B Kannel et al. "Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates 11Reprints are not available." In: *The American Journal of Cardiology* 82 (7 Oct. 1998), 2N–9N. ISSN: 00029149. DOI: 10.1016/S0002-9149(98)00583-9 (cit. on p. 11).

[107] Lars O. Karlsson et al. "A clinical decision support tool for improving adherence to guidelines on anticoagulant therapy in patients with atrial fibrillation at risk of stroke: A cluster-randomized trial in a Swedish primary care setting (the CDS-AF study)". In: *PLOS Medicine* 15 (3 Mar. 2018), e1002528. ISSN: 1549-1676. DOI: 10.1371/journal.pmed.1002528 (cit. on p. 39).

[108] Lars O. Karlsson et al. "Clinical decision support for stroke prevention in atrial fibrillation (CDS-AF): Rationale and design of a cluster randomized trial in the primary care setting". In: *American Heart Journal* 187 (May 2017), pp. 45–52. ISSN: 00028703. DOI: 10.1016/j.ahj.2017.02.009 (cit. on pp. 39, 46).

[109] Georg Delle Karth et al. "Amiodarone versus diltiazem for rate control in critically ill patients with atrial tachyarrhythmias". In: *Critical Care Medicine* 29 (6 June 2001), pp. 1149–1153. ISSN: 0090-3493. DOI: 10.1097/00003246-200106000-00011 (cit. on p. 122).

[110] Robert E. Kass and Adrian E. Raftery. "Bayes Factors". In: *Journal of the American Statistical Association* 90 (430 June 1995), pp. 773–795. ISSN: 0162-1459. DOI: 10.1080/01621459.1995.10476572 (cit. on p. 71).

[111] Monika Kastner et al. "Usability evaluation of a clinical decision support tool for osteoporosis disease management". In: *Implementation Science* 5 (1 Dec. 2010), p. 96. ISSN: 1748-5908. DOI: 10.1186/1748-5908-5-96 (cit. on p. 125).

[112] Tsuchida Keizo and Kazuhiko Tanabe. "Influence of paroxysmal atrial fibrillation attack on brain natriuretic peptide secretion". In: *Journal of cardiology* 44 (1 2004), pp. 1–11 (cit. on p. 92).

[113] Gad Keren et al. "Atrial fibrillation and atrial enlargement in patients with mitral stenosis". In: *American Heart Journal* 114 (5 Nov. 1987), pp. 1146–1155. ISSN: 00028703. DOI: 10.1016/0002-8703(87)90190-6 (cit. on p. 15).

[114] Anna Kezerashvili, Kevin Marzo, and Joshua De Leon. "Beta Blocker Use After Acute Myocardial Infarction in the Patient with Normal Systolic Function: When is it "Ok" to Discontinue?" In: *Current Cardiology Reviews* 8 (1 June 2012), pp. 77–84. ISSN: 1573403X. DOI: 10.2174/157340312801215764 (cit. on p. 122).

[115] Paulus Kirchhof et al. "2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS". In: *European Heart Journal* 37 (38 Oct. 2016), pp. 2893–2962. ISSN: 0195-668X. DOI: 10.1093/eurheartj/ehw210 (cit. on p. 31).

[116] William A. Knaus. "The SUPPORT Prognostic Model: Objective Estimates of Survival for Seriously Ill Hospitalized Adults". In: *Annals of Internal Medicine* 122 (3 Feb. 1995), p. 191. ISSN: 0003-4819 (cit. on p. 63).

[117] Jelena Kornej et al. "The APPLE score: a novel and simple score for the prediction of rhythm outcomes after catheter ablation of atrial fibrillation". In: *Clinical Research in Cardiology* 104.10 (Oct. 2015), pp. 871–876. ISSN: 1861-0684. DOI: 10.1007/s00392-015-0856-x (cit. on p. 26).

[118] Jedrzej Kosiuk et al. "Prospective, multicenter validation of a clinical risk score for left atrial arrhythmogenic substrate based on voltage analysis: DR-FLASH score". In: *Heart Rhythm* 12.11 (Nov. 2015), pp. 2207–2212. ISSN: 15475271. DOI: 10.1016/j.hrthm.2015.07.003 (cit. on p. 26).

[119] Agnieszka Kotalczyk, Gregory YH Lip, and Hugh Calkins. "The 2020 ESC Guidelines on the Diagnosis and Management of Atrial Fibrillation". In: *Arrhythmia & Electrophysiology Review* 10 (2 July 2021), pp. 65–67. ISSN: 20503369. DOI: 10.15420/aer.2021.07 (cit. on p. 15).

[120] Dipak Kotecha et al. "European Society of Cardiology smartphone and tablet applications for patients with atrial fibrillation and their health care providers". In: *EP Europace* 20 (2 Feb. 2018), pp. 225–233. ISSN: 1099-5129. DOI: 10.1093/europace/eux299 (cit. on pp. 43, 46).

[121] N. Kouvas et al. "The complex crosstalk between inflammatory cytokines and ventricular arrhythmias". In: *Cytokine* 111 (Nov. 2018), pp. 171–177. ISSN: 10434666. DOI: 10.1016/j.cyto.2018.08.007 (cit. on p. 7).

[122] Alexander Lacki and Antonio Martinez-Millana. "A Comparison of the Impact of Pharmacological Treatments on Cardioversion, Rate Control, and Mortality in Data-Driven Atrial Fibrillation Phenotypes in Critical Care". In: *Bioengineering* 11 (3 Feb. 2024), p. 199. ISSN: 2306-5354. DOI: 10.3390/bioengineering11030199 (cit. on pp. 65, 66).

[123] Alexander Lacki and Antonio Martinez-Millana. "Survival Hierarchical Agglomerative Clustering: A Semi-Supervised Clustering Method Incorporating Survival Data". In: *Artificial Intelligence in Medicine*. Ed. by Jose M. Juarez et al. Cham: Springer Nature Switzerland, 2023, pp. 3–12. ISBN: 978-3-031-34344-5 (cit. on pp. 57, 58, 66, 100, 104).

[124] Stephen Andrew LaHaye et al. "A clinical decision aid for the selection of antithrombotic therapy for the prevention of stroke due to atrial fibrillation". In: *European Heart Journal* 33 (17 Sept. 2012), pp. 2163–2171. ISSN: 0195-668X. DOI: 10.1093/eurheartj/ehs167 (cit. on pp. 39, 46).

[125] Mah Laka, Adriana Milazzo, and Tracy Merlin. "Factors That Impact the Adoption of Clinical Decision Support Systems (CDSS) for Antibiotic Management". In: *International Journal of Environmental Research and Public Health* 18 (4 Feb. 2021), p. 1901. ISSN: 1660-4601. DOI: 10.3390/ijerph18041901 (cit. on p. 125).

[126] T. Lange et al. "Learning with Constrained and Unlabelled Data". In: IEEE, 2005, pp. 731–738. ISBN: 0-7695-2372-2. DOI: 10.1109/CVPR.2005.210 (cit. on p. 30).

[127] Semi-Supervised Learning. "Semi-supervised learning". In: *CSZ2006. html* 5 (2006) (cit. on p. 28).

[128] Changhee Lee, Jinsung Yoon, and Mihaela van der Schaar. "Dynamic-DeepHit: A Deep Learning Approach for Dynamic Survival Analysis With Competing Risks Based on Longitudinal Data". In: *IEEE Transactions on Biomedical Engineering* 67 (1 Jan. 2020), pp. 122–133. ISSN: 0018-9294. DOI: 10.1109/TBME.2019.2909027 (cit. on p. 56).

[129] Hopin Lee et al. "Smartphone and tablet apps for concussion road warriors (team clinicians): a systematic review for practical users". In: *British Journal of Sports Medicine* 49 (8 Apr. 2015), pp. 499–505. ISSN: 0306-3674. DOI: 10.1136/bjsports-2013-092930 (cit. on p. 125).

[130] Dorothy Leonard-Barton and Isabelle Deschamps. "Managerial Influence in the Implementation of New Technology". In: *Management Science* 34 (10 Oct. 1988), pp. 1252–1265. ISSN: 0025-1909. DOI: 10.1287/mnsc.34.10.1252 (cit. on p. 72).

[131] Yan-Guang Li et al. "A Simple Clinical Risk Score (C2HEST) for Predicting Incident Atrial Fibrillation in Asian Subjects". In: *Chest* 155.3 (Mar. 2019), pp. 510–518. ISSN: 00123692. DOI: 10.1016/j.chest.2018.09.011 (cit. on p. 25).

[132] R Likert. "A technique for the measurement of attitudes." In: *Archives of Psychology* 22 140 (1932), p. 55 (cit. on p. 77).

[133] Dominik Linz et al. "Associations of Obstructive Sleep Apnea With Atrial Fibrillation and Continuous Positive Airway Pressure Treatment".

In: *JAMA Cardiology* 3 (6 June 2018), p. 532. ISSN: 2380-6583. DOI: 10.1001/jamacardio.2018.0095 (cit. on p. 91).

[134] Dominik Linz et al. "Role of autonomic nervous system in atrial fibrillation". In: *International Journal of Cardiology* 287 (July 2019), pp. 181–188. ISSN: 01675273. DOI: 10.1016/j.ijcard.2018.11.091 (cit. on p. 14).

[135] Gregory Y H Lip et al. "Clinical factors related to successful or unsuccessful cardioversion in the EdoxabaN versus warfarin in subjectS UndeRgoing cardiovErsion of Atrial Fibrillation (ENSURE-AF) randomized trial". en. In: *J. Arrhythm.* 36.3 (June 2020), pp. 430–438 (cit. on p. 24).

[136] Gregory Y.H. Lip et al. "Refining Clinical Risk Stratification for Predicting Stroke and Thromboembolism in Atrial Fibrillation Using a Novel Risk Factor-Based Approach". In: *Chest* 137.2 (Feb. 2010), pp. 263–272. ISSN: 00123692. DOI: 10.1378/chest.09-1584 (cit. on p. 22).

[137] L. Liu and S. Nattel. "Differing sympathetic and vagal effects on atrial fibrillation in dogs: role of refractoriness heterogeneity". In: *American Journal of Physiology-Heart and Circulatory Physiology* 273 (2 Aug. 1997), H805–H816. ISSN: 0363-6135. DOI: 10.1152/ajpheart.1997.273.2.H805 (cit. on p. 14).

[138] Marie Lockhart et al. "Extracellular matrix and heart development". In: *Birth Defects Research Part A: Clinical and Molecular Teratology* 91 (6 June 2011), pp. 535–550. ISSN: 15420752. DOI: 10.1002/bdra.20810 (cit. on p. 7).

[139] Laurens van der Maaten and Geoffrey Hinton. "Visualizing Data using t-SNE". In: *Journal of Machine Learning Research* 9.86 (2008), pp. 2579–2605 (cit. on p. 59).

[140] Ravi Mandapati et al. "Stable Microreentrant Sources as a Mechanism of Atrial Fibrillation in the Isolated Sheep Heart". In: *Circulation* 101 (2 Jan. 2000), pp. 194–199. ISSN: 0009-7322. DOI: 10.1161/01.CIR.101.2.194 (cit. on p. 13).

[141] Laura Manduchi et al. "A Deep Variational Approach to Clustering Survival Data". In: *International Conference on Learning Representations*. 2022 (cit. on p. 56).

[142] Bruna Miers May et al. "GDF-15 como Biomarcador em Doenças Cardiovasculares". In: *Arquivos Brasileiros de Cardiologia* (Feb. 2021). ISSN: 0066-782X. DOI: 10.36660/abc.20200426 (cit. on p. 92).

[143] João Mesquita et al. "Development and validation of a risk score for predicting atrial fibrillation recurrence after a first catheter ablation procedure - ATLAS score". In: *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology* 20.FI3 (2018), f428–f435. ISSN: 15322092. DOI: 10.1093/europace/eux265 (cit. on p. 26).

[144] M. Michalowski et al. "AFGuide System to Support Personalized Management of Atrial Fibrillation". In: *AAAI Workshops*. 2017 (cit. on pp. 40, 46).

[145] S Miller. "Effects of magnesium on atrial fibrillation after cardiac surgery: a meta-analysis". In: *Heart* 91 (5 May 2005), pp. 618–623. ISSN: 1355-6037. DOI: 10.1136/hrt.2004.033811 (cit. on p. 122).

[146] Abdul M Minhas et al. "Link Between Non-Alcoholic Fatty Liver Disease and Atrial Fibrillation: A Systematic Review and Meta-Analysis". In: *Cureus* (Apr. 2017). ISSN: 2168-8184. DOI: 10.7759/cureus.1142 (cit. on p. 91).

[147] L. Brent Mitchell. "Bradyarrhythmias/Heart Blocks". In: 2017, pp. 219–238. DOI: 10.1007/978-3-319-58000-5_18 (cit. on p. 5).

[148] G.K. Moe and J.A. Abildskov. "Atrial fibrillation as a self-sustaining arrhythmia independent of focal discharge". In: *American Heart Journal* 58 (1 July 1959), pp. 59–70. ISSN: 00028703. DOI: 10.1016/0002-8703(59)90274-1 (cit. on p. 13).

[149] Roy Moncayo and Helga Moncayo. "The WOMED model of benign thyroid disease: Acquired magnesium deficiency due to physical and psychological stressors relates to dysfunction of oxidative phosphoryla-

tion". In: *BBA Clinical* 3 (June 2015), pp. 44–64. ISSN: 22146474. DOI: 10.1016/j.bbacli.2014.11.002 (cit. on p. 122).

[150] Carlos A. Morillo et al. "Chronic Rapid Atrial Pacing". In: *Circulation* 91 (5 Mar. 1995), pp. 1588–1595. ISSN: 0009-7322. DOI: 10.1161/01.CIR.91.5.1588 (cit. on p. 15).

[151] Ravil Muhamedyev. "Machine learning methods: An overview". In: *Computer modelling & new technologies* 19.6 (2015), pp. 14–29 (cit. on p. 27).

[152] Nebojša Mujović et al. "Prediction of very late arrhythmia recurrence after radiofrequency catheter ablation of atrial fibrillation: The MB-LATER clinical score". In: *Scientific Reports* 7.1 (Feb. 2017), p. 40828. ISSN: 2045-2322. DOI: 10.1038/srep40828 (cit. on p. 26).

[153] Marcel Lucas Müller et al. "Towards integration of clinical decision support in commercial hospital information systems using distributed, reusable software and knowledge components". In: *International Journal of Medical Informatics* 64 (2-3 Dec. 2001), pp. 369–377. ISSN: 13865056. DOI: 10.1016/S1386-5056(01)00218-0 (cit. on p. 126).

[154] Mark A. Musen. "The protégé project". In: *AI Matters* 1 (4 June 2015), pp. 4–12. ISSN: 2372-3483. DOI: 10.1145/2757001.2757003 (cit. on p. 67).

[155] Mark A. Musen, Blackford Middleton, and Robert A. Greenes. "Clinical Decision-Support Systems". In: Springer International Publishing, 2021, pp. 795–840. DOI: 10.1007/978-3-030-58721-5_24 (cit. on p. 67).

[156] Yasuyuki Nasuhara et al. "Physicians' responses to computerized drug interaction alerts with password overrides". In: *BMC Medical Informatics and Decision Making* 15 (1 Dec. 2015), p. 74. ISSN: 1472-6947. DOI: 10.1186/s12911-015-0194-y (cit. on p. 35).

[157] Stanley Nattel. "Atrial Electrophysiology and Mechanisms of Atrial Fibrillation". In: *Journal of Cardiovascular Pharmacology and Therapeutics* 8 (1_suppl Mar. 2003), S5–S11. ISSN: 1074-2484. DOI: 10.1177/107424840300800102 (cit. on pp. 9, 90).

[158] Stanley Nattel, Brett Burstein, and Dobromir Dobrev. "Atrial Remodeling and Atrial Fibrillation". In: *Circulation: Arrhythmia and Electrophysiology* 1 (1 Apr. 2008), pp. 62–73. ISSN: 1941-3149. DOI: 10.1161/CIRCEP.107.754564 (cit. on pp. 11, 13, 14).

[159] Peter Brønnum Nielsen et al. "Female Sex Is a Risk Modifier Rather Than a Risk Factor for Stroke in Atrial Fibrillation". In: *Circulation* 137 (8 Feb. 2018), pp. 832–840. ISSN: 0009-7322. DOI: 10.1161/CIRCULATIONAHA.117.029081 (cit. on p. 29).

[160] OHDSI. *The Book of OHDSI: Observational Health Data Sciences and Informatics*. OHDSI, 2019. ISBN: 9781088855195 (cit. on p. 66).

[161] Olufisayo Olusegun Olakotan and Maryati Mohd Yusof. "The appropriateness of clinical decision support systems alerts in supporting clinical workflows: A systematic review". In: *Health Informatics Journal* 27 (2 Apr. 2021), p. 146045822110075. ISSN: 1460-4582. DOI: 10.1177/14604582211007536 (cit. on pp. 33, 34).

[162] Jerome A Osheroff et al. "Improving Medication Use and Outcomes with Clinical Decision Support: A Step by Step Guide". In: HIMSS. 2009 (cit. on p. 33).

[163] Nicolas Paris, Antoine Lamer, and Adrien Parrot. "Transformation and Evaluation of the MIMIC Database in the OMOP Common Data Model: Development and Usability Study". In: *JMIR Medical Informatics* 9 (12 Dec. 2021), e30970. ISSN: 2291-9694. DOI: 10.2196/30970 (cit. on p. 66).

[164] M. D. Paterno et al. "Tiering Drug-Drug Interaction Alerts by Severity Increases Compliance Rates". In: *Journal of the American Medical Informatics Association* 16 (1 Jan. 2009), pp. 40–46. ISSN: 1067-5027. DOI: 10.1197/jamia.M2808 (cit. on p. 35).

[165] Emily S. Patterson et al. "Identifying barriers to the effective use of clinical reminders: Bootstrapping multiple methods". In: *Journal of Biomedical Informatics* 38 (3 June 2005), pp. 189–199. ISSN: 15320464. DOI: 10.1016/j.jbi.2004.11.015 (cit. on p. 34).

[166] Giuseppe Patti et al. "Safety and efficacy of nonvitamin K antagonist oral anticoagulants versus warfarin in diabetic patients with atrial fibril-

lation: A study-level meta-analysis of phase III randomized trials". In: *Diabetes/Metabolism Research and Reviews* 33 (3 Mar. 2017), e2876. ISSN: 15207552. DOI: 10.1002/dmrr.2876 (cit. on p. 17).

[167] Thomas H Payne et al. "Recommendations to improve the usability of drug-drug interaction clinical decision support alerts". In: *Journal of the American Medical Informatics Association* 22 (6 Nov. 2015), pp. 1243–1250. ISSN: 1527-974X. DOI: 10.1093/jamia/ocv011 (cit. on p. 33).

[168] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cit. on p. 61).

[169] Alejandro Perez-Lugones et al. "Evidence of Specialized Conduction Cells in Human Pulmonary Veins of Patients with Atrial Fibrillation". In: *Journal of Cardiovascular Electrophysiology* 14 (8 Aug. 2003), pp. 803–809. ISSN: 1045-3873. DOI: 10.1046/j.1540-8167.2003.03075.x (cit. on p. 12).

[170] Jeremy Petch, Shuang Di, and Walter Nelson. "Opening the Black Box: The Promise and Limitations of Explainable Machine Learning in Cardiology". In: *Canadian Journal of Cardiology* 38 (2 Feb. 2022), pp. 204–213. ISSN: 0828282X (cit. on p. 56).

[171] Arnold Pinter and Paul Dorian. "Intravenous antiarrhythmic agents". In: *Current opinion in cardiology* 16.1 (2001), pp. 17–22 (cit. on p. 123).

[172] Ron Pisters et al. "A Novel User-Friendly Score (HAS-BLED) To Assess 1-Year Risk of Major Bleeding in Patients With Atrial Fibrillation". In: *Chest* 138.5 (Nov. 2010), pp. 1093–1100. ISSN: 00123692. DOI: 10.1378/chest.10-0134 (cit. on p. 24).

[173] María Poveda-Villalón, Asunción Gómez-Pérez, and Mari Carmen Suárez-Figueroa. "OOPS! (OntOlogy Pitfall Scanner!)" In: *International Journal on Semantic Web and Information Systems* 10 (2 Apr. 2014), pp. 7–34. ISSN: 1552-6283. DOI: 10.4018/ijswis.2014040102 (cit. on p. 67).

[174] Yue Qin et al. "Research Progress on Semi-Supervised Clustering". In: *Cognitive Computation* 11 (5 Oct. 2019), pp. 599–612. ISSN: 1866-9956. DOI: 10.1007/s12559-019-09664-w (cit. on p. 53).

[175] Faisal Rahman, Gene F. Kwan, and Emelia J. Benjamin. "Global epidemiology of atrial fibrillation". In: *Nature Reviews Cardiology* 11 (11 Nov. 2014), pp. 639–654. ISSN: 1759-5002. DOI: 10.1038/nrcardio.2014.118 (cit. on p. 1).

[176] Kiran Reddy et al. "Subphenotypes in critical care: translation into clinical practice". In: *The Lancet Respiratory Medicine* 8 (6 June 2020), pp. 631–643. ISSN: 22132600. DOI: 10.1016/S2213-2600(20)30124-7 (cit. on p. 31).

[177] James A. Reiffel. "Atrial Fibrillation and Stroke: Epidemiology". In: *The American Journal of Medicine* 127 (4 Apr. 2014), e15–e16. ISSN: 00029343. DOI: 10.1016/j.amjmed.2013.06.002 (cit. on p. 11).

[178] Muthukrishnan Renganathan, Haiyang Wei, and Yong Zhao. "Cardiac Action Potential Measurement in Human Embryonic Stem Cell Cardiomyocytes for Cardiac Safety Studies Using Manual Patch-Clamp Electrophysiology". In: 2017, pp. 37–56. DOI: 10.1007/978-1-4939-6661-5_3 (cit. on pp. 8, 9).

[179] Greg Ridgeway et al. "Twang: Toolkit for Weighting and Analysis of Nonequivalent Groups. R package version 1.0–1". In: *R Foundation for Statistical Computing, 2006. (http://www.cran.r-project.org). (Accessed.* 2008 (cit. on p. 71).

[180] Paul R. Rosenbaum and Donald B. Rubin. "The central role of the propensity score in observational studies for causal effects". In: *Biometrika* 70 (1 1983), pp. 41–55. ISSN: 0006-3444. DOI: 10.1093/biomet/70.1.41 (cit. on p. 71).

[181] Bijit Roy et al. "Survival analysis with semi-supervised predictive clustering trees". In: *Computers in Biology and Medicine* 141 (Feb. 2022), p. 105001. ISSN: 00104825. DOI: 10.1016/j.compbiomed.2021.105001 (cit. on p. 56).

[182] Xueying Ru et al. "Using a Clinical Decision Support System to Improve Anticoagulation in Patients with Nonvalve Atrial Fibrillation in China's Primary Care Settings: A Feasibility Study". In: *International Journal of Clinical Practice* 2023 (Jan. 2023), pp. 1–10. ISSN: 1742-1241. DOI: 10.1155/2023/2136922 (cit. on pp. 40, 46).

[183] Donald Rubin. *Multiple Imputation for Nonresponse in Surveys*. Ed. by Donald B. Rubin. John Wiley & Sons, Inc., June 1987. ISBN: 9780470316696. DOI: 10.1002/9780470316696 (cit. on p. 70).

[184] Donald B. Rubin. "The Bayesian Bootstrap". In: *The Annals of Statistics* 9 (1 Jan. 1981). ISSN: 0090-5364. DOI: 10.1214/aos/1176345338 (cit. on p. 71).

[185] Alissa L. Russ et al. "Prescribers' interactions with medication alerts at the point of prescribing: A multi-method, in situ investigation of the human–computer interaction". In: *International Journal of Medical Informatics* 81 (4 Apr. 2012), pp. 232–243. ISSN: 13865056. DOI: 10.1016/j.ijmedinf.2012.01.002 (cit. on p. 34).

[186] Ryota Sakurai and Kazuhiko Ohe. "Effects of Computerized Guideline-Oriented Clinical Decision Support System on Antithrombotic Therapy in Patients with Atrial Fibrillation: A Systematic Review and Meta-Analysis". In: *MEDINFO 2019: Health and Wellbeing e-Networks for All* 264 (2019), pp. 768–772 (cit. on p. 40).

[187] A. L. Samuel. "Some Studies in Machine Learning Using the Game of Checkers". In: *IBM Journal of Research and Development* 3.3 (1959), pp. 210–229. DOI: 10.1147/rd.33.0210 (cit. on p. 27).

[188] Margarete Sandelowski. "Whatever happened to qualitative description?" In: *Research in Nursing & Health* 23 (4 Aug. 2000), pp. 334–340. ISSN: 0160-6891. DOI: 10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G (cit. on p. 72).

[189] Renate B Schnabel et al. "Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study". In: *The Lancet* 373.9665 (Feb. 2009), pp. 739–745. ISSN: 01406736. DOI: 10.1016/S0140-6736(09)60443-8 (cit. on p. 25).

[190] Ulrich Schotten et al. "Electrical and Contractile Remodeling During the First Days of Atrial Fibrillation Go Hand in Hand". In: *Circulation* 107 (10 Mar. 2003), pp. 1433–1439. ISSN: 0009-7322. DOI: 10.1161/01.CIR.0000055314.10801.4F (cit. on pp. 14, 15).

[191] M Schumacher et al. "Randomized 2 x 2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. German Breast Cancer Study Group." In: *Journal of Clinical Oncology* 12 (10 Oct. 1994), pp. 2086–2093. ISSN: 0732-183X (cit. on p. 63).

[192] Peter J. Schwartz and Raymond L. Woosley. "Predicting the Unpredictable". In: *Journal of the American College of Cardiology* 67 (13 Apr. 2016), pp. 1639–1650. ISSN: 07351097. DOI: 10.1016/j.jacc.2015.12.063 (cit. on p. 34).

[193] Simon Sedej and Burkert Pieske. *Mechanisms of Ca2+-Triggered Arrhythmias*. InTech, Mar. 2012. DOI: 10.5772/25906 (cit. on p. 12).

[194] Naqash J. Sethi et al. "The effects of rhythm control strategies versus rate control strategies for atrial fibrillation and atrial flutter: A systematic review with meta-analysis and Trial Sequential Analysis". In: *PLOS ONE* 12 (10 Oct. 2017), e0186856. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0186856 (cit. on p. 19).

[195] Ciara M. Shaver et al. "Atrial Fibrillation Is an Independent Predictor of Mortality in Critically Ill Patients*". In: *Critical Care Medicine* 43 (10 Oct. 2015), pp. 2104–2111. ISSN: 0090-3493. DOI: 10.1097/CCM.0000000000001166 (cit. on p. 1).

[196] Reza Sheibani et al. "Effects of Computerized Decision Support Systems on Management of Atrial Fibrillation: A Scoping Review". In: *Journal of Atrial Fibrillation* 10 (1 June 2017). ISSN: 1941-6911. DOI: 10.4022/jafib.1579 (cit. on p. 41).

[197] Reza Sheibani et al. "The Effect of a Clinical Decision Support System on Improving Adherence to Guideline in the Treatment of Atrial Fibrillation: An Interrupted Time Series Study". In: *Journal of Medical Systems* 42 (2 Feb. 2018), p. 26. ISSN: 0148-5598. DOI: 10.1007/s10916-017-0881-6 (cit. on pp. 39, 46).

[198] Shruthi H. Shetty et al. "Supervised Machine Learning: Algorithms and Applications". In: Wiley, Feb. 2022, pp. 1–16. DOI: 10.1002/9781119821908.ch1 (cit. on p. 28).

[199] Günther Silbernagel et al. "Electronic Alert System for Improving Stroke Prevention Among Hospitalized Oral-Anticoagulation-Naïve Patients With Atrial Fibrillation: A Randomized Trial". In: *Journal of the American Heart Association* 5 (7 July 2016). ISSN: 2047-9980. DOI: 10.1161/JAHA.116.003776 (cit. on pp. 39, 46).

[200] Daniel E. Singer et al. "A New Risk Scheme to Predict Ischemic Stroke and Other Thromboembolism in Atrial Fibrillation: The ATRIA Study Stroke Risk Score". In: *Journal of the American Heart Association* 2.3 (May 2013). ISSN: 2047-9980. DOI: 10.1161/JAHA.113.000250 (cit. on p. 22).

[201] Deepti Sisodia, Lokesh Singh, and Sheetal Sisodia. "Clustering Techniques: A Brief Survey of Different Clustering Algorithms". In: 2012 (cit. on pp. 28, 30).

[202] Gopal Sivagangabalan et al. "Regional Ion Channel Gene Expression Heterogeneity and Ventricular Fibrillation Dynamics in Human Hearts". In: *PLoS ONE* 9 (1 Jan. 2014), e82179. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0082179 (cit. on p. 9).

[203] John Stacy et al. "Qualitative Evaluation of an Artificial Intelligence–Based Clinical Decision Support System to Guide Rhythm Management of Atrial Fibrillation: Survey Study". In: *JMIR Formative Research* 6 (8 Aug. 2022), e36443. ISSN: 2561-326X. DOI: 10.2196/36443 (cit. on pp. 41, 42, 46).

[204] Yihong Sun and Dayi Hu. "The link between diabetes and atrial fibrillation: cause or correlation?" In: *Journal of Cardiovascular Disease Research* 1 (1 Jan. 2010), pp. 9–11. ISSN: 09753583. DOI: 10.4103/0975-3583.59978 (cit. on p. 90).

[205] Shinya Suzuki. ""Cholesterol Paradox" in Atrial Fibrillation". In: *Circulation Journal* 75 (12 2011), pp. 2749–2750. ISSN: 1346-9843. DOI: 10.1253/circj.CJ-11-1134 (cit. on p. 93).

[206] Shinya Suzuki et al. "Identifying risk patterns in older adults with atrial fibrillation by hierarchical cluster analysis: A retrospective approach based on the risk probability for clinical events". In: *IJC Heart & Vas-*

culature 37 (Dec. 2021), p. 100883. ISSN: 23529067. DOI: 10.1016/j.ijcha.2021.100883 (cit. on pp. 32, 72).

[207] Michelle Sweidan et al. "Quality of drug interaction alerts in prescribing and dispensing software". In: *Medical Journal of Australia* 191 (6 Sept. 2009), pp. 358–359. ISSN: 0025-729X. DOI: 10.5694/j.1326-5377.2009.tb02832.x (cit. on p. 34).

[208] WA Tacker. *Defibrillation of the heart: ICDs, AEDs, and Manual*. Mosby Incorporated, 1994 (cit. on p. 17).

[209] Daiji Takeuchi et al. "Prominent atrial fibrosis and its correlation with atrial tachyarrhythmia and dilated atrium long after classic Fontan surgery". In: *Journal of Cardiology* 79 (5 May 2022), pp. 671–677. ISSN: 09145087. DOI: 10.1016/j.jjcc.2021.12.005 (cit. on p. 15).

[210] Morten B. Thomsen and Kirstine Calloe. "Human atrial fibroblasts and their contribution to supraventricular arrhythmia". In: *Physiological Reports* 4 (3 Feb. 2016), e12711. ISSN: 2051817X. DOI: 10.14814/phy2.12711 (cit. on p. 7).

[211] R. Thomson. "Development and description of a decision analysis based decision support tool for stroke prevention in atrial fibrillation". In: *Quality and Safety in Health Care* 11 (1 Mar. 2002), pp. 25–31. ISSN: 14753898. DOI: 10.1136/qhc.11.1.25 (cit. on p. 37).

[212] R. G Thomson et al. "A patient decision aid to support shared decision-making on anti-thrombotic treatment of patients with atrial fibrillation: randomised controlled trial". In: *Quality and Safety in Health Care* 16 (3 June 2007), pp. 216–223. ISSN: 1475-3898. DOI: 10.1136/qshc.2006.018481 (cit. on pp. 38, 46).

[213] Madhukar H Trivedi et al. "Barriers to implementation of a computerized decision support system for depression: an observational report on lessons learned in "real world" clinical settings". In: *BMC Medical Informatics and Decision Making* 9 (1 Dec. 2009), p. 6. ISSN: 1472-6947. DOI: 10.1186/1472-6947-9-6 (cit. on p. 35).

[214] Gary Tse et al. "Predictive value of inter-atrial block for new onset or recurrent atrial fibrillation: A systematic review and meta-analysis". In:

International Journal of Cardiology 250 (Jan. 2018), pp. 152–156. ISSN: 01675273. DOI: 10.1016/j.ijcard.2017.09.176 (cit. on p. 91).

[215] H van der Velden. "Gap junctional remodeling in relation to stabilization of atrial fibrillation in the goat". In: *Cardiovascular Research* 46 (3 June 2000), pp. 476–486. ISSN: 00086363. DOI: 10.1016/S0008-6363(00)00026-2 (cit. on p. 15).

[216] Venkatesh, Thong, and Xu. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology". In: *MIS Quarterly* 36 (1 2012), p. 157. ISSN: 02767783. DOI: 10.2307/41410412 (cit. on pp. 72, 79).

[217] Marco Vitolo et al. "Clinical Phenotype Classification of Atrial Fibrillation Patients Using Cluster Analysis and Associations with Trial-Adjudicated Outcomes". In: *Biomedicines* 9 (7 July 2021), p. 843. ISSN: 2227-9059. DOI: 10.3390/biomedicines9070843 (cit. on pp. 32, 72).

[218] Allan J. Walkey et al. "Practice Patterns and Outcomes of Treatments for Atrial Fibrillation During Sepsis". In: *Chest* 149 (1 Jan. 2016), pp. 74–83. ISSN: 00123692 (cit. on p. 123).

[219] Xiaoxu Wang, Yingzheng Li, and Qiulai Li. "A comprehensive meta-analysis on relationship between CYP11B2 rs1799998 polymorphism and atrial fibrillation". en. In: *J. Electrocardiol.* 52 (Jan. 2019), pp. 101–105 (cit. on p. 119).

[220] Yao-Sheng Wang et al. "Prolonged P-wave duration is associated with atrial fibrillation recurrence after radiofrequency catheter ablation: A systematic review and meta-analysis". In: *International Journal of Cardiology* 227 (Jan. 2017), pp. 355–359. ISSN: 01675273. DOI: 10.1016/j.ijcard.2016.11.058 (cit. on p. 91).

[221] Joe H. Ward. "Hierarchical Grouping to Optimize an Objective Function". In: *Journal of the American Statistical Association* 58 (301 Mar. 1963), pp. 236–244. ISSN: 0162-1459. DOI: 10.1080/01621459.1963.10500845 (cit. on p. 31).

[222] A. T. M. Wasylewicz and A. M. J. W. Scheepers-Hoeks. "Clinical Decision Support Systems". In: *Fundamentals of Clinical Data Science.*

Ed. by Pieter Kubben, Michel Dumontier, and Andre Dekker. Cham: Springer International Publishing, 2019, pp. 153–169. ISBN: 978-3-319-99713-1. DOI: 10.1007/978-3-319-99713-1_11 (cit. on p. 33).

[223] Eiichi Watanabe et al. "Clinical phenotypes of patients with non-valvular atrial fibrillation as defined by a cluster analysis: A report from the J-RHYTHM registry". In: *IJC Heart & Vasculature* 37 (Dec. 2021), p. 100885. ISSN: 23529067. DOI: 10.1016/j.ijcha.2021.100885 (cit. on pp. 32, 72).

[224] Chi-Jen Weng et al. "Rhythm control better prevents stroke and mortality than rate control strategies in patients with atrial fibrillation - A nationwide cohort study". In: *International Journal of Cardiology* 270 (Nov. 2018), pp. 154–159. ISSN: 01675273. DOI: 10.1016/j.ijcard.2018.06.090 (cit. on p. 19).

[225] Mik Wetterslev et al. "Management of acute atrial fibrillation in the intensive care unit: An international survey". In: *Acta Anaesthesiologica Scandinavica* 66 (3 Mar. 2022), pp. 375–385. ISSN: 0001-5172. DOI: 10.1111/aas.14007 (cit. on pp. 2, 76, 123).

[226] Rohan S Wijesurendra and Barbara Casadei. "Mechanisms of atrial fibrillation". In: *Heart* 105 (24 Dec. 2019), pp. 1860–1867. ISSN: 1355-6037. DOI: 10.1136/heartjnl-2018-314267 (cit. on pp. 1, 15).

[227] Maurits C.E.F. Wijffels et al. "Atrial Fibrillation Begets Atrial Fibrillation". In: *Circulation* 92 (7 Oct. 1995), pp. 1954–1968. ISSN: 0009-7322. DOI: 10.1161/01.CIR.92.7.1954 (cit. on p. 15).

[228] E. M. VAUGHAN WILLIAMS. "A Classification of Antiarrhythmic Actions Reassessed After a Decade of New Drugs". In: *The Journal of Clinical Pharmacology* 24 (4 Apr. 1984), pp. 129–147. ISSN: 00912700. DOI: 10.1002/j.1552-4604.1984.tb01822.x (cit. on p. 18).

[229] Roger A. Winkle et al. "Predicting atrial fibrillation ablation outcome: The CAAP-AF score". In: *Heart Rhythm* 13.11 (Nov. 2016), pp. 2119–2125. ISSN: 15475271. DOI: 10.1016/j.hrthm.2016.07.018 (cit. on p. 26).

[230] Daniela M. Witten and Robert Tibshirani. "A Framework for Feature Selection in Clustering". In: *Journal of the American Statistical Association* 105 (490 June 2010), pp. 713–726. ISSN: 0162-1459. DOI: 10.1198/jasa.2010.tm09415 (cit. on p. 57).

[231] Maciej Wójcik et al. "Repeated catheter ablation of atrial fibrillation: How to predict outcome?" In: *Circulation Journal* 77.9 (2013), pp. 2271–2279. ISSN: 13469843. DOI: 10.1253/circj.CJ-13-0308 (cit. on p. 26).

[232] P A Wolf et al. "Probability of stroke: a risk profile from the Framingham Study." In: *Stroke* 22 (3 Mar. 1991), pp. 312–318. ISSN: 0039-2499. DOI: 10.1161/01.STR.22.3.312 (cit. on p. 37).

[233] RL Woosley et al. "Medical decision support systems and therapeutics: The role of autopilots". In: *Clinical Pharmacology & Therapeutics* 99 (2 Feb. 2016), pp. 161–164. ISSN: 00099236. DOI: 10.1002/cpt.259 (cit. on p. 33).

[234] Yung-Hsin Yeh et al. "Calcium-Handling Abnormalities Underlying Atrial Arrhythmogenesis and Contractile Dysfunction in Dogs With Congestive Heart Failure". In: *Circulation: Arrhythmia and Electrophysiology* 1 (2 June 2008), pp. 93–102. ISSN: 1941-3149. DOI: 10.1161/CIRCEP.107.754788 (cit. on p. 13).

[235] Michael Juntao Yuan et al. "Evaluation of User Interface and Workflow Design of a Bedside Nursing Clinical Decision Support System". In: *Interactive Journal of Medical Research* 2 (1 Jan. 2013), e4. ISSN: 1929-073X. DOI: 10.2196/ijmr.2402 (cit. on p. 125).

[236] Maedeh Zamani, Esra Karaca, and Ngan F. Huang. "Multicellular Interactions in 3D Engineered Myocardial Tissue". In: *Frontiers in Cardiovascular Medicine* 5 (Oct. 2018). ISSN: 2297-055X. DOI: 10.3389/fcvm.2018.00147 (cit. on p. 6).

[237] Philip Zazove et al. "To act or not to act: responses to electronic health record prompts by family medicine clinicians". In: *Journal of the American Medical Informatics Association* 24 (2 Mar. 2017), pp. 275–280. ISSN: 1067-5027. DOI: 10.1093/jamia/ocw178 (cit. on p. 35).

[238] Xiaojin Zhu. "Semi-Supervised Learning Literature Survey". In: *Comput Sci, University of Wisconsin-Madison* 2 (July 2008) (cit. on p. 30).

[239] Dimitrios Zikos. "A Framework to Design Successful Clinical Decision Support Systems". In: ACM, June 2017, pp. 185–188. ISBN: 9781450352277. DOI: 10.1145/3056540.3064960 (cit. on p. 35).

[240] Douglas P. Zipes et al. "Termination of ventricular fibrillation in dogs by depolarizing a critical amount of myocardium". In: *The American Journal of Cardiology* 36 (1 July 1975), pp. 37–44. ISSN: 00029149. DOI: 10.1016/0002-9149(75)90865-6 (cit. on p. 17).

Appendix

Appendix 1 - Predictive biomarkers identified through systematic review

Table 7.1: Biomarkers and Counts

Category	Biomarker	Count (Prevalence %)
Demographics	Age	49 (31)
	Gender	38 (24)
Comorbidities	Hypertension	33 (21)
	Diabetes	31 (20)
	Heart Failure	27 (17)
	Vascular Disease	24 (15)
	Renal Dysfunction	20 (13)
	COPD	6 (4)
	Liver Dysfunction	6 (4)
	Thyroid Dysfunction	5 (3)
	Cardiomyopathy	4 (3)
	OSA	3 (2)
	Metabolic Syndrome	3 (2)
	Rheumatism	3 (2)
	Valvulopathy	1 (1)

	Non-alcoholic Fatty Liver Disease	1 (1)
Observations	BMI	21 (13)
	SysBP	18 (11)
	Alcohol	14 (9)
	Smoking	14 (9)
	DiaBP	11 (7)
	Drug Abuse	7 (4)
	Weight	7 (4)
	Height	6 (4)
	BSA	2 (1)
	BP Variability	1 (1)
History	Prev. Stroke	28 (18)
	AF duration	14 (9)
	AF type	11 (7)
	Prev. MI	10 (6)
	Prev. Bleed	8 (5)
	Prev. infarct location	1 (1)
	Prev. Infarct Vol.	1 (1)
ECG	Heart Rate	7 (4)
	PR interval	5 (3)
	LV Hypertrophy	4 (3)
	P-wave Duration	3 (2)
	Premature Atrial Contractions	3 (2)
	QTc	3 (2)
	P-Wave Axis	2 (1)
	Cornell product	2 (1)
	tP+	1 (1)
	tP-	1 (1)
	HRV	1 (1)
	Strain	1 (1)
	P-wave Terminal force V1	1 (1)
	AF Burden	1 (1)
	F-wave Voltage	1 (1)
	F-wave Amplitude	1 (1)
	T-wave Devation	1 (1)
	QRS duration	1 (1)
	Bundle branch bock	1 (1)

	IA Block	1 (1)
	Sokolow-Lyon voltage	1 (1)
Serum	NT-proBNT	30 (19)
	GDF-15	18 (11)
	Hemoglobin	14 (9)
	Troponin-T	13 (8)
	IL-6	10 (6)
	C-Troponin-T	10 (6)
	Total Cholesterol	8 (5)
	BNP	7 (4)
	Creatinine	7 (4)
	INR	7 (4)
	CRP	7 (4)
	Triglycerides	6 (4)
	HDL-C	6 (4)
	White Blood Cell Count	5 (3)
	LDL-C	5 (3)
	Galectin-3	5 (3)
	D-dimer	5 (3)
	Von Willebrand factor (vWF)	4 (3)
	TGF-beta1	4 (3)
	Uric acid	4 (3)
	FGF23	3 (2)
	Red Blood Cell Distribution Width	3 (2)
	Dyslipidemia	3 (2)
	Fibrinogen	3 (2)
	Mid-regional pro atrial natriuretic peptide (MR-proANP)	2 (1)
	ANP	2 (1)
	TNF-R1	2 (1)
	ApoA1	2 (1)
	ApoB	2 (1)
	Albuminuria	2 (1)
	Urea Nitrogen	2 (1)
	Neutrophil-to-lymphocyte ratio	2 (1)
	Platelet count	2 (1)
	Thyrotropin (TSH)	2 (1)

C-Troponin-I	2 (1)
C-Reactive Protein	2 (1)
Neopterin	1 (1)
Ang2	1 (1)
BMP10	1 (1)
DKK3	1 (1)
ESM1	1 (1)
FABP^3	1 (1)
IGFBP7	1 (1)
MYBPC3	1 (1)
Mid-regional pro-atrial natriuretic peptide	1 (1)
Cancer Antigen-125 (Ca-125)	1 (1)
TNF-alpha	1 (1)
TNF-R2	1 (1)
Trefoil Factor 3 (TFF3)	1 (1)
ST2	1 (1)
Cortisol	1 (1)
Albumin	1 (1)
Red Blood Cell Count	1 (1)
Mean Platelet Volume	1 (1)
Platelet Distribution Width	1 (1)
Cystatin C	1 (1)
CITP	1 (1)
TIMP	1 (1)
TIMP2	1 (1)
Apolipoprotein CIII (ApoC-III)	1 (1)
ADAMTS13	1 (1)
Urokinase plasminogen activator surface receptor (uPAR)	1 (1)
Urokinase plasminogen activator (uPA)	1 (1)
Ephrin type-B receptor 4 (EPHB4),	1 (1)
IL-8	1 (1)
Plasmin-antiplasmin	1 (1)
Fibrinopeptide-A	1 (1)
Antithrombin III	1 (1)

	Plasminogen activator inhibitor	1 (1)
	Prothrombin F1+2	1 (1)
	Free thyroxine (fT4)	1 (1)
	Inflammatory markers	1 (1)
	OPG	1 (1)
	OPN	1 (1)
	suPAR	1 (1)
	EphB4	1 (1)
	TRAIL-R2	1 (1)
	ST2	1 (1)
	C-C motif chemokine 16 (CCL16)	1 (1)
	Troponin-I	1 (1)
	Transferrin receptor protein 1 (TfR1)	1 (1)
	Osteopontin (OPN)	1 (1)
	Tissue plasminogen activator antigen (tPA ag)	1 (1)
	ADMA/SDMA	1 (1)
	Aspartate aminotransferase	1 (1)
	Alanine aminotransferase	1 (1)
	Glutamyltransferase (GGT)	1 (1)
	Telomerre length	1 (1)
	Thrombomodulin (sTM)	1 (1)
Imaging	LVEV	28 (18)
	LAD	20 (13)
	LA-Vol	17 (11)
	LA-EF	9 (6)
	LA contraction strain	6 (4)
	LA strain	4 (3)
	LA-Expansion	3 (2)
	Fibrosis Quantity	3 (2)
	E/e'	3 (2)
	LVDDis	2 (1)
	LA reservoir strain	2 (1)
	TAPSE	1 (1)
	RAD	1 (1)
	Peak E-wave velociy	1 (1)
	Peak A-wave velocity	1 (1)

	Passive emptying index	1 (1)
	LV Diastolic Diameter	1 (1)
	LV Volume	1 (1)
	LV end-Sys Volume	1 (1)
	LV end-Dia Volume	1 (1)
	LAA Shape	1 (1)
	LAA Flow Velocity	1 (1)
	LA Stiffness	1 (1)
	LA Sphericity	1 (1)
	LA Conduit Strain	1 (1)
	Global Longitudinal Strain	1 (1)
	Epicardial Adipose Tissue Thickness	1 (1)
	E/Em	1 (1)
	E/A ratio	1 (1)
	Atrial Conduction Time	1 (1)
	Active Emptying Index	1 (1)
Genetics	CYP 2C9 single-nucleotide polymorphisms	2 (1)
	CYP11B2 rs1799998 polymorphism	1 (1)
	GJA1 rs13216675 Polymorphism	1 (1)
	FRMD4B	1 (1)
	CAV1	1 (1)
	rs10033464	1 (1)
	rs2200733	1 (1)
	FRMD4B	1 (1)
	CAV1	1 (1)
	rs10033464	1 (1)
	rs2200733	1 (1)

Table 7.3: Covariate means before and after inverse probability treatment weighting for rhythm control and in-hospital mortality.

	Unweighted Means				Weighted Means					
Variable	BBs	PCBs	CCBs	MgS	BBs	PCBs	CCBs	MgS	Maximum absolute pairwise SMD unweighted	Maximum pairwise SMD weighted
Anemia	0.13	0.13	0.11	0.12	0.13	0.12	0.12	0.11	0.02	0.04
Serum calcium	8.66	8.56	8.51	8.52	8.6	8.55	8.52	8.56	0.09	0.06
COPD	0.14	0.12	0.16	0.14	0.14	0.13	0.15	0.13	0.07	0.05
Collagen disease	0.01	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.03
Cor pulmonale	0.11	0.09	0.07	0.1	0.1	0.09	0.09	0.12	0.05	0.11
Coronary artery atherosclerosis	0.29	0.52	0.18	0.23	0.32	0.38	0.3	0.3	0.38	0.09
Serum creatinine	1.45	1.39	1.45	1.5	1.43	1.48	1.4	1.42	0.05	0.04
Diabetes	0.25	0.21	0.22	0.23	0.25	0.21	0.23	0.21	0.05	0.05
Thyroid disorder	0.13	0.08	0.13	0.11	0.11	0.10	0.13	0.1	0.06	0.06
Erythrocyte distribution width	15.1	14.5	15.0	15.2	14.9	14.8	14.9	14.9	0.15	0.03
Erythrocyte count	3.96	3.88	3.94	3.84	3.93	3.9	3.93	3.94	0.06	0.04
Heart failure	0.54	0.42	0.42	0.54	0.5	0.46	0.45	0.51	0.14	0.07
Hemoglobin	12.01	12.26	11.94	11.6	12.01	12.08	12	12.07	0.11	0.02
Hypertension	0.48	0.51	0.45	0.42	0.49	0.46	0.48	0.42	0.07	0.11
Left Atrial Dilation	0.55	0.52	0.5	0.55	0.53	0.54	0.51	0.56	0.05	0.05
Leukocyte count	11.8	11.2	13.3	12.5	12.1	11.7	12.2	11.9	0.12	0.03
Serum magnesium	2.02	2.09	1.94	1.86	2.00	2.03	1.96	1.94	0.14	0.10
Myocardial infarction	0.08	0.16	0.08	0.08	0.09	0.13	0.11	0.11	0.16	0.07
OSA	0.05	0.03	0.05	0.05	0.05	0.03	0.05	0.05	0.05	0.07
Platelet count	237	222	253	239	232	234	241	235	0.16	0.05
Post-operative condition	0.21	0.64	0.08	0.11	0.29	0.36	0.25	0.24	0.68	0.18
Serum Potassium	4.35	4.29	4.36	4.39	4.32	4.32	4.33	4.32	0.06	0.01
Prothrombin time	20.4	15.6	17.8	21.8	18.8	17.1	18.0	18.6	0.27	0.10
Renal insufficiency	0.41	0.37	0.39	0.41	0.39	0.43	0.38	0.38	0.04	0.07
Respiratory Failure	0.21	0.15	0.27	0.23	0.2	0.21	0.24	0.21	0.17	0.09
Rheumatism	0.06	0.05	0.02	0.04	0.05	0.05	0.03	0.05	0.08	0.09
Right Atrial Dilation	0.44	0.31	0.34	0.45	0.41	0.35	0.35	0.45	0.16	0.14
Sepsis	0.14	0.13	0.19	0.23	0.14	0.18	0.16	0.19	0.20	0.09
Serum sodium	138	138	138	138	138	138	138	138	0.04	0.02
Valvulopathy	0.31	0.47	0.17	0.24	0.32	0.38	0.27	0.28	0.30	0.12
Age	77.6	73.1	73.8	77.3	75.7	74.5	74.8	76.2	0.21	0.08
Sex	0.55	0.6	0.48	0.58	0.58	0.56	0.53	0.6	0.09	0.08
Arrhythmia History	0.3	0.14	0.24	0.3	0.25	0.2	0.23	0.28	0.16	0.12
Heart Rate	93.0	89.9	100.9	90.8	93.0	92.8	94.8	93.0	0.35	0.07

Table 7.4: Covariate means before and after inverse probability treatment weighting for rate control.

	Unweighted Means				Weighted Means					
Variable	BBs	PCBs	CCBs	MgS	BBs	PCBs	CCBs	MgS	Maximum absolute pairwise SMD unweighted	Maximum pairwise SMD weighted
Anemia	0.11	0.12	0.12	0.12	0.13	0.14	0.13	0.11	0.05	0.06
Serum calcium	8.62	8.50	8.56	8.50	8.55	8.45	8.52	8.54	0.11	0.05
COPD	0.14	0.16	0.13	0.16	0.13	0.15	0.14	0.15	0.07	0.05
Collagen disease	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.04	0.08
Cor pulmonale	0.09	0.07	0.07	0.09	0.09	0.09	0.09	0.10	0.05	0.06
Coronary artery atherosclerosis	0.27	0.16	0.30	0.29	0.52	0.23	0.37	0.32	0.38	0.10
Serum creatinine	1.49	1.45	1.44	1.39	1.41	1.56	1.47	1.41	0.08	0.04
Diabetes	0.23	0.23	0.23	0.24	0.21	0.24	0.22	0.25	0.03	0.05
Thyriod disorder	0.12	0.12	0.11	0.11	0.08	0.11	0.10	0.11	0.07	0.04
Erythrocyte distribution width	15.2	15.0	15.0	14.9	14.6	15.4	14.9	14.8	0.25	0.06
Erythrocytes count	3.94	3.94	3.91	3.94	3.87	3.81	3.89	3.99	0.13	0.11
Heart failure	0.57	0.41	0.48	0.43	0.41	0.55	0.44	0.42	0.22	0.09
Hemoglobin	11.9	12.0	12.0	12.0	12.2	11.5	12.1	12.2	0.24	0.05
Hypertension	0.46	0.46	0.47	0.50	0.50	0.39	0.46	0.48	0.16	0.05
Left Atrial Dilation	0.58	0.53	0.56	0.54	0.51	0.62	0.54	0.56	0.15	0.03
Leukocyte count	11.6	13.8	11.9	12.5	11.2	12.0	11.9	11.9	0.18	0.04
Serum magnesium	2.01	1.95	1.99	1.97	2.08	1.86	2.03	1.92	0.21	0.13
Myocardial infarction	0.08	0.08	0.10	0.11	0.16	0.11	0.14	0.17	0.14	0.17
OSA	0.05	0.05	0.04	0.05	0.03	0.01	0.03	0.01	0.17	0.17
Platelet count	236	251	235	242	225	237	234	231	0.13	0.05
Post-operative condition	0.17	0.09	0.28	0.25	0.63	0.15	0.38	0.29	0.62	0.17
Serum potassium	4.33	4.38	4.33	4.33	4.30	4.38	4.31	4.32	0.05	0.02
Prothrombin time	20.3	17.8	18.1	17.5	15.7	20.0	16.6	17.6	0.23	0.10
Renal insufficiency	0.43	0.41	0.39	0.39	0.39	0.52	0.44	0.39	0.22	0.05
Respiratory failure	0.25	0.30	0.22	0.26	0.16	0.32	0.22	0.25	0.21	0.08
Rheumatism	0.06	0.03	0.05	0.02	0.04	0.06	0.04	0.05	0.08	0.09
Right Atrial Dilation	0.45	0.38	0.40	0.37	0.31	0.52	0.35	0.46	0.31	0.17
Sepsis	0.17	0.20	0.16	0.18	0.14	0.29	0.20	0.20	0.32	0.08
Serum sodium	138	138	138	138	139	138	138	138	0.10	0.03
Valvulopathy	0.30	0.16	0.32	0.25	0.44	0.22	0.36	0.25	0.33	0.13
Age	77.1	74.0	75.3	74.1	72.9	76.0	73.9	74.8	0.24	0.08
Sex	0.53	0.49	0.55	0.51	0.60	0.52	0.57	0.57	0.12	0.07
Arrhythmia history	0.28	0.26	0.23	0.24	0.13	0.25	0.17	0.26	0.20	0.13
Heart Rate	98.7	101.5	96.3	97.8	90.0	100.9	94.8	97.4	0.26	0.08

Table 7.5: Phenotype characteristics

	Clusters												
Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
Anemia	39 (9.31)	48 (34.8)	12 (25.5)	134 (11.2)	7 (9.33)	26 (35.6)	22 (11.6)	79 (13.8)	11 (4.44)	14 (7.33)	344 (13.2)	27 (7.05)	22 (14.5)
Serum Calcium	8.50 (8.00-9.00)	8.60 (8.18-9.30)	8.40 (7.90-8.80)	8.70 (8.20-9.10)	8.60 (7.90-9.10)	8.70 (8.20-9.00)	8.70 (8.20-9.10)	8.40 (7.80-9.00)	8.70 (8.10-9.10)	8.80 (8.30-9.20)	8.56 (8.20-9.00)	8.60 (8.20-9.00)	8.50 (7.67-9.00)
COPD	68 (16.2)	35 (25.4)	5 (10.6)	220 (18.4)	12 (16.0)	24 (32.9)	34 (18.0)	101 (17.6)	19 (7.66)	39 (20.4)	223 (8.56)	51 (13.3)	7 (4.61)
Collagen Disease	3 (0.72)	10 (7.25)	5 (10.6)	15 (1.26)	1 (1.33)	0 (0.00)	1 (0.53)	10 (1.74)	6 (2.42)	2 (1.05)	22 (0.84)	5 (1.31)	0 (0.00)
Cor Pulmonale	5 (1.19)	26 (18.8)	8 (17.0)	114 (9.56)	12 (16.0)	26 (35.6)	36 (19.1)	26 (4.53)	35 (14.1)	40 (20.9)	234 (8.98)	16 (4.18)	8 (5.26)
Coronary artery atherosclerosis	37 (8.83)	113 (81.9)	11 (23.4)	361 (30.3)	26 (34.7)	27 (37.0)	30 (15.9)	53 (9.23)	14 (5.65)	94 (49.2)	1873 (71.9)	47 (12.3)	19 (12.5)
Creatinine	1.10 (0.90-1.40)	1.50 (1.10-2.00)	0.90 (0.70-1.30)	1.30 (1.00-2.00)	1.10 (0.90-1.50)	0.90 (0.70-1.30)	1.50 (1.00-2.40)	1.70 (1.10-3.00)	1.10 (0.87-1.50)	1.00 (0.80-1.30)	0.90 (0.80-1.10)	1.00 (0.80-1.20)	1.80 (1.20-2.60)
Diabetes	34 (8.11)	27 (19.6)	6 (12.8)	418 (35.0)	11 (14.7)	26 (35.6)	38 (20.1)	87 (15.2)	91 (36.7)	62 (32.5)	552 (21.2)	20 (5.22)	31 (20.4)

Left Atrial Dilation	Hypertension	Hemoglobin	Heart Failure	Erythrocyte Count	Erythrocyte distribution width	Thyroid Disorder
113 (27.0)	312 (74.5)	12.2 (10.6-13.9)	17 (4.06)	3.92 (3.45-4.47)	14.4 (13.5-15.7)	16 (3.82)
88 (63.8)	42 (30.4)	11.3 (10.1-12.6)	60 (43.5)	3.69 (3.29-4.21)	14.6 (13.6-15.8)	19 (13.8)
43 (91.5)	16 (34.0)	11.7 (10.3-12.9)	12 (25.5)	3.7 (3.03-4.28)	14.4 (13.2-15.9)	14 (29.8)
302 (25.3)	450 (37.7)	11.5 (10.0-13.0)	749 (62.8)	3.84 (3.37-4.33)	15.0 (14.1-16.4)	167 (14.0)
56 (74.7)	72 (96.0)	12.9 (10.9-14.5)	33 (44.0)	4.14 (3.61-4.48)	14.1 (13.4-15.4)	8 (10.7)
54 (74.0)	46 (63.0)	11.2 (10.4-12.1)	59 (80.8)	3.83 (3.39-4.25)	15.1 (14.2-16.4)	15 (20.6)
94 (49.7)	27 (14.3)	11.1 (9.7-12.6)	174 (92.1)	3.68 (3.25-4.19)	15.6 (14.2-17.4)	20 (10.6)
102 (17.8)	128 (22.3)	11.1 (9.8-12.8)	155 (27.0)	3.65 (3.23-4.15)	15.1 (14.0-16.6)	57 (9.93)
242 (97.6)	81 (32.7)	11.8 (10.57-13.2)	180 (72.6)	3.97 (3.55-4.41)	14.7 (13.8-16.1)	30 (12.1)
161 (84.3)	158 (82.7)	12.0 (10.7-13.7)	156 (81.7)	4.03 (3.55-4.5)	14.3 (13.6-15.6)	29 (15.2)
1125 (43.2)	1710 (65.6)	12.6 (11.4-13.7)	840 (32.3)	3.85 (3.3-4.34)	13.8 (13.2-14.6)	214 (8.21)
374 (97.7)	204 (53.3)	12.4 (10.8-13.7)	70 (18.3)	4.03 (3.58-4.51)	14.1 (13.5-15.2)	34 (8.88)
149 (98.0)	29 (19.1)	11.1 (9.97-13.0)	109 (71.7)	3.74 (3.31-4.24)	15.3 (14.1-17.3)	16 (10.5)

Serum Potassium	Postoperative Condition	Platelet Count	OSA	Myocardial infarction	Serum Magnesium	Leukocyte Count
4.20 (3.80-4.60)	13 (3.10)	214 (167-283)	17 (4.06)	9 (2.15)	1.90 (1.70-2.10)	10.7 (8.10-15.2)
4.40 (3.80-4.90)	90 (65.2)	234 (178-298)	6 (4.35)	45 (32.6)	2.10 (1.80-2.30)	10.3 (7.10-13.4)
4.10 (3.80-4.50)	11 (23.4)	224 (169-290)	1 (2.13)	8 (17.0)	1.90 (1.70-2.00)	11.2 (7.50-16.1)
4.30 (3.90-4.80)	58 (4.86)	222 (170-287)	83 (6.96)	58 (4.86)	2.00 (1.80-2.20)	10.1 (7.50-13.3)
4.20 (3.90-4.60)	10 (13.3)	211 (156-250)	17 (22.7)	14 (18.7)	1.90 (1.70-2.10)	8.40 (6.80-14.3)
4.28 (3.70-4.80)	21 (28.8)	260 (196-348)	3 (4.11)	18 (24.7)	2.00 (1.80-2.30)	10.7 (8.20-16.8)
4.40 (3.90-5.10)	15 (7.94)	225 (174-307)	13 (6.88)	54 (28.6)	2.10 (1.80-2.40)	11.1 (8.10-15.5)
4.40 (3.90-5.00)	20 (3.48)	222 (157-324)	22 (3.83)	41 (7.14)	2.00 (1.70-2.20)	11.5 (7.70-17.0)
4.25 (3.90-4.80)	13 (5.24)	234 (186-300)	18 (7.26)	20 (8.06)	1.95 (1.70-2.20)	10.4 (7.20-13.7)
4.10 (3.90-4.60)	70 (36.7)	245 (191-320)	6 (3.14)	68 (35.6)	1.90 (1.80-2.10)	11.1 (8.00-13.8)
4.20 (3.90-4.50)	2407 (92.4)	195 (150-243)	77 (2.96)	326 (12.5)	2.08 (1.90-2.30)	9.10 (6.90-12.4)
4.10 (3.80-4.50)	26 (6.79)	221 (173-291)	12 (3.13)	11 (2.87)	1.90 (1.70-2.10)	10.0 (7.10-13.1)
4.40 (3.90-4.92)	8 (5.26)	196 (147-291)	8 (5.26)	18 (11.8)	2.00 (1.70-2.30)	11.4 (7.60-17.0)

Valvulopathy	Serum Sodium	Sepsis	Right Atrial Dilation	Rheumatism	Respiratory Failure	Renal insufficiency	Prothrombin time
41 (9.79)	139 (136-141)	76 (18.1)	40 (9.55)	1 (0.24)	102 (24.3)	107 (25.5)	14.4 (13.2-18.4)
32 (23.2)	138 (135-140)	11 (7.97)	23 (16.7)	12 (8.70)	25 (18.1)	114 (82.6)	13.9 (13.0-16.4)
36 (76.6)	139 (135-142)	14 (29.8)	24 (51.1)	5 (10.6)	2 (4.26)	4 (8.51)	13.8 (12.8-19.6)
262 (22.0)	139 (136-141)	116 (9.72)	181 (15.2)	45 (3.77)	152 (12.7)	695 (58.3)	16.1 (13.4-24.5)
48 (64.0)	138 (136-141)	24 (32.0)	52 (69.3)	6 (8.00)	10 (13.3)	25 (33.3)	14.6 (13.1-17.5)
32 (43.8)	139 (136-141)	20 (27.4)	43 (58.9)	12 (16.4)	36 (49.3)	23 (31.5)	14.8 (13.1-22.1)
51 (27.0)	139 (136-141)	32 (16.9)	68 (36.0)	9 (4.76)	108 (57.1)	139 (73.5)	15.2 (13.4-19.6)
47 (8.19)	139 (135-142)	302 (52.6)	47 (8.19)	11 (1.92)	227 (39.6)	504 (87.8)	14.9 (13.3-19.6)
41 (16.5)	139 (137-142)	4 (1.61)	208 (83.9)	5 (2.02)	72 (29.0)	89 (35.9)	15.3 (13.3-21.2)
102 (53.4)	139 (136-141)	10 (5.24)	102 (53.4)	16 (8.38)	30 (15.7)	51 (26.7)	13.9 (12.8-18.3)
1478 (56.7)	139 (137-141)	16 (0.61)	602 (23.1)	152 (5.83)	71 (2.73)	243 (9.33)	14.0 (13.0-15.7)
41 (10.7)	139 (137-141)	23 (6.01)	320 (83.6)	13 (3.39)	40 (10.4)	21 (5.48)	14.1 (12.9-17.2)
20 (13.2)	137 (134-140)	121 (79.6)	136 (89.5)	6 (3.95)	139 (91.5)	135 (88.8)	15.8 (14.0-21.2)

Hourly Rhythm Control CCBs	Hourly Rhythm Control PCBs	Hourly Rhythm Control BBs	Heart Rate	Arrhythmia History	Male Sex	Age
6.98 (4.81-11.6)	8.62 (6.22-13.1)	1.49 (0.92-2.88)	87.0 (75.0-103)	32 (7.64)	402 (95.9)	76.2 (65.3-82.1)
12.2 (4.16-45.7)	10.2 (7.08-23.5)	6.54 (3.39-16.8)	85.0 (75.0-96.0)	14 (10.1)	59 (42.8)	76.8 (70.2-83.0)
5.71 (0.00-23.5)	12.9 (7.31-40.6)	3.28 (0.00-7.58)	87.0 (77.0-100)	15 (31.9)	1 (2.13)	73.3 (60.2-85.5)
4.20 (2.99-5.88)	6.64 (4.77-9.82)	2.49 (1.85-3.03)	85.0 (72.0-100)	678 (56.8)	650 (54.5)	78.87 (69.89-86.0)
8.26 (0.64-17.9)	11.6 (7.46-39.6)	1.15 (0.24-3.33)	84.1 (72.0-105)	27 (36.0)	70 (93.3)	79.1 (67.2-83.7)
5.59 (2.65-20.3)	7.71 (3.12-25.7)	1.27 (0.11-4.71)	90.0 (78.0-105)	12 (16.4)	14 (19.2)	78.18 (70.6-85.0)
2.21 (0.46-8.13)	3.00 (1.33-8.78)	1.31 (0.67-2.91)	88.0 (74.0-101)	32 (16.9)	107 (56.6)	81.0 (74.3-87.1)
5.05 (3.61-7.49)	8.77 (5.80-14.7)	2.37 (1.69-4.28)	91.0 (77.0-109)	81 (14.1)	316 (55.1)	76.57 (65.3-84.1)
3.75 (2.08-8.06)	6.47 (4.10-14.2)	1.80 (0.89-2.62)	89.0 (74.8-103)	43 (17.3)	45 (18.2)	79.7 (70.5-87.5)
6.09 (2.88-11.4)	14.4 (10.0-24.5)	1.89 (0.90-3.27)	85.0 (73.0-99.0)	40 (20.9)	43 (22.5)	80.03 (72.1-84.9)
6.06 (4.09-9.95)	9.62 (8.65-11.2)	3.61 (2.57-4.62)	83.0 (77.0-90.0)	257 (9.87)	1622 (62.3)	73.4 (65.4-80.4)
4.07 (2.22-7.73)	11.2 (7.33-26.5)	1.96 (1.16-3.25)	86.0 (74.0-104)	110 (28.7)	215 (56.1)	74.45 (63.8-83.0)
3.36 (2.01-6.08)	6.59 (2.78-14.1)	2.90 (1.81-4.95)	95.0 (79.8-114)	21 (13.8)	108 (71.1)	74.9 (65.8-83.9)

Daily Mortality Rate BBs	Hourly Rate Control MgS	Hourly Rate Control CCBs	Hourly Rate Control PCBs	Hourly Rate Control BBs	Hourly Rhythm Control MgS
1.80 (1.15-2.78)	11.6 (3.70-55.2)	39.3 (21.6-59.7)	14.3 (6.69-25.6)	11.5 (5.98-31.6)	2.53 (1.24-6.91)
2.72 (0.11-5.32)	NaN	16.6 (8.53-133)	6.09 (3.24-27.4)	11.8 (6.86-100)	0.00 (0.00-0.00)
0.68 (0.00-4.18)	0.34 (0.00-33.3)	0.96 (0.00-76.3)	9.08 (0.00-58.7)	11.7 (5.00-100)	3.00 (0.00-13.7)
0.94 (0.71-1.66)	7.20 (2.94-17.9)	22.3 (15.7-35.3)	11.8 (6.28-26.0)	12.6 (7.32-19.1)	2.02 (1.36-3.24)
0.56 (0.00-2.26)	100 (100-100)	17.1 (9.20-236)	15.7 (0.00-182.0)	5.79 (2.49-18.1)	0.00 (0.00-0.00)
1.24 (0.33-5.66)	100 (100-100)	26.7 (13.8-119)	47.2 (22.1-91.9)	30.6 (24.5-145)	1.43 (0.00-6.65)
1.34 (0.57-2.23)	4.49 (0.00-45.7)	26.4 (12.1-40.2)	25.1 (5.56-52.7)	5.30 (1.75-19.8)	0.039 (0.00-2.3)
1.60 (0.79-2.65)	19.1 (12.0-47.8)	14.9 (7.48-22.5)	12.4 (6.07-31.7)	8.50 (5.02-21.7)	2.70 (1.46-5.28)
0.74 (0.33-1.43)	26.0 (20.0-75.3)	12.1 (6.85-37.4)	11.3 (3.15-33.2)	10.1 (6.76-24.8)	0.85 (0.18-3.13)
0.76 (0.34-1.79)	10.9 (7.14-100)	9.27 (1.34-36.6)	9.58 (4.42-31.7)	15.3 (6.61-25.8)	0.96 (0.094-5.3)
0.33 (0.090-0.56)	15.0 (11.3-39.3)	15.3 (9.58-31.3)	16.4 (12.9-18.9)	8.85 (6.03-15.5)	3.13 (2.24-5.84)
0.65 (0.37-1.4)	15.2 (5.85-100)	15.4 (8.24-22.3)	13.6 (7.52-27.4)	11.3 (5.22-31.7)	2.03 (1.00-3.99)
2.31 (0.97-3.8)	11.1 (3.02-49.5)	14.4 (4.27-33.1)	5.75 (1.32-19.4)	3.24 (0.32-8.93)	2.17 (0.40-5.91)

Daily Mortality Rate PCBs	Daily Mortality Rate CCBs	Daily Mortality Rate MgS
1.45 (0.54-2.53)	1.79 (1.02-4.24)	0.68 (0.21-2.06)
0.58 (0.15-1.74)	1.08 (0.00-18.4)	1.16 (0.00-18.1)
2.69 (0.70-6.32)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
1.01 (0.56-1.71)	1.30 (0.80-2.11)	0.50 (0.25-1.21)
0.57 (0.00-5.32)	0.00 (0.00-0.00)	0.00 (0.00-2.28)
8.02 (3.36-24.0)	1.08 (0.00-3.53)	0.19 (0.00-4.49)
1.74 (0.49-3.91)	1.67 (0.62-4.41)	3.49 (1.32-9.16)
3.35 (2.37-5.58)	2.78 (1.75-4.27)	3.23 (1.54-6.17)
1.09 (0.21-3.54)	1.42 (0.329-3.18)	1.17 (0.00-4.16)
0.48 (0.00-1.82)	0.57 (0.078-1.56)	1.92 (0.075-3.34)
0.29 (0.15-0.42)	0.40 (0.058-0.97)	0.15 (0.071-0.941)
0.00 (0.00-0.00)	0.71 (0.26-1.66)	1.22 (0.40-2.83)
4.58 (1.23-8.45)	3.47 (1.81-6.85)	2.74 (1.11-5.08)

Table 7.6: Phenotype characteristics, continued.

	Clusters												
Variable	14	15	16	17	18	19	20	21	22	23	24	25	26
Anemia	50 (22.0)	10 (7.30)	89 (10.9)	3 (10.3)	52 (11.9)	14 (13.2)	41 (6.78)	13 (31.7)	6 (5.22)	28 (8.56)	12 (19.7)	4 (14.3)	20 (10.8)
Serum Calcium	8.50 (7.80-8.90)	8.70 (8.30-9.20)	8.60 (8.20-9.10)	8.20 (7.90-8.90)	8.72 (8.30-9.10)	8.60 (8.20-9.10)	8.70 (8.20-9.20)	9.10 (8.70-9.50)	8.40 (7.80-8.90)	8.66 (8.20-9.20)	8.50 (8.00-9.10)	8.50 (8.17-9.03)	8.60 (8.10-8.90)
COPD	98 (43.2)	16 (11.7)	68 (8.34)	2 (6.90)	56 (12.8)	5 (4.72)	64 (10.6)	25 (61.0)	24 (20.9)	30 (9.17)	10 (16.4)	3 (10.7)	17 (9.19)
Collagen disease	5 (2.20)	5 (3.65)	4 (0.49)	0 (0.00)	2 (0.46)	2 (1.89)	7 (1.16)	0 (0.00)	1 (0.87)	3 (0.92)	2 (3.28)	0 (0.00)	4 (2.16)
Cor pulmonale	13 (5.73)	3 (2.19)	11 (1.35)	1 (3.45)	51 (11.6)	9 (8.49)	39 (6.45)	7 (17.1)	5 (4.35)	15 (4.59)	3 (4.92)	0 (0.00)	11 (5.95)
Coronary artery atherosclerosis	37 (16.3)	19 (13.9)	50 (6.13)	5 (17.2)	378 (86.3)	8 (7.55)	99 (16.4)	3 (7.32)	1 (0.87)	107 (32.7)	16 (26.2)	13 (46.4)	20 (10.8)
Creatinine	1.40 (1.00-2.50)	0.86 (0.70-1.10)	0.90 (0.70-1.10)	1.20 (0.70-1.70)	1.30 (1.00-1.60)	1.00 (0.90-1.60)	1.50 (1.10-2.30)	1.50 (1.10-2.10)	0.90 (0.70-1.10)	1.00 (0.80-1.30)	1.40 (0.90-1.70)	2.25 (1.20-3.45)	1.00 (0.80-1.20)
Diabetes	130 (57.3)	13 (9.49)	105 (12.9)	3 (10.3)	207 (47.3)	4 (3.77)	154 (25.5)	31 (75.6)	4 (3.48)	51 (15.6)	9 (14.8)	1 (3.57)	32 (17.3)

Left Atrial Dilation	Hypertension	Hemoglobin	Heart Failure	Erythrocyte Count	Erythrocyte distribution width	Thyroid Disorder
65 (28.6)	47 (20.7)	10.5 (9.30-12.1)	113 (49.8)	3.57 (3.1-4.09)	15.7 (14.5-17.3)	43 (18.9)
38 (27.7)	116 (84.7)	12.3 (11.2-13.6)	13 (9.49)	3.99 (3.61-4.34)	14.2 (13.4-15.3)	124 (90.5)
81 (9.94)	318 (39.0)	12.2 (10.8-13.6)	158 (19.4)	3.99 (3.59-4.48)	14.1 (13.4-15.3)	41 (5.03)
28 (96.6)	27 (93.1)	11.3 (10.9-12.1)	1 (3.45)	3.72 (3.4-3.95)	14.2 (13.1-16.6)	9 (31.0)
344 (78.5)	179 (40.9)	12.4 (11.1-13.5)	315 (71.9)	4.04 (3.51-4.47)	14.2 (13.6-15.2)	31 (7.08)
99 (93.4)	61 (57.6)	12.0 (10.2-13.3)	85 (80.2)	3.90 (3.46-4.5)	14.5 (13.7-16.2)	45 (42.5)
514 (85.0)	117 (19.3)	11.6 (10.2-13.1)	462 (76.4)	3.79 (3.31-4.33)	15.2 (14.2-16.9)	50 (8.26)
40 (97.6)	24 (58.5)	11.6 (10.0-13.4)	39 (95.1)	3.99 (3.45-4.36)	15.7 (14.7-16.7)	3 (7.32)
14 (12.2)	8 (6.96)	12.2 (10.5-13.6)	14 (12.2)	3.99 (3.51-4.43)	14.3 (13.6-15.6)	5 (4.35)
12 (3.67)	273 (83.5)	12.1 (10.8-13.3)	234 (71.6)	3.99 (3.54-4.42)	14.6 (13.7-15.8)	46 (14.1)
12 (19.7)	2 (3.28)	11.7 (10.3-12.9)	19 (31.2)	3.65 (3.03-4.14)	14.4 (13.7-16.0)	2 (3.28)
26 (92.9)	8 (28.6)	11.0 (9.97-12.1)	21 (75.0)	3.65 (3.22-4.08)	14.7 (13.6-15.5)	3 (10.7)
7 (3.78)	107 (57.8)	12.7 (11.0-14.3)	40 (21.6)	4.15 (3.55-4.57)	14.3 (13.7-15.6)	10 (5.41)

Serum Potassium	Postoperative Condition	Platelet Count	OSA	Myocardial infarction	Serum Magnesium	Leukocyte Count
4.30 (3.80-4.90)	4 (1.76)	233 (158-308)	31 (13.7)	5 (2.20)	2.00 (1.70-2.20)	10.9 (7.40-15.8)
4.10 (3.80-4.40)	5 (3.65)	239 (196-298)	5 (3.65)	1 (0.73)	1.90 (1.80-2.00)	10.6 (7.00-14.1)
4.10 (3.80-4.50)	33 (4.05)	242 (183-308)	11 (1.35)	25 (3.07)	1.89 (1.70-2.10)	10.6 (7.70-13.9)
4.30 (4.00-5.10)	1 (3.45)	207 (126-268)	0 (0.00)	3 (10.3)	2.00 (1.80-2.20)	11.6 (6.70-14.6)
4.20 (3.90-4.70)	351 (80.1)	206 (163-256)	31 (7.08)	127 (29.0)	2.10 (1.90-2.30)	8.90 (6.80-12.6)
4.20 (3.60-4.60)	1 (0.94)	235 (174-319)	4 (3.77)	0 (0.00)	2.00 (1.80-2.20)	10.4 (7.50-14.3)
4.30 (3.90-4.90)	23 (3.80)	209 (157-267)	23 (3.80)	27 (4.46)	2.00 (1.80-2.30)	10.0 (7.00-14.5)
4.60 (4.10-5.40)	0 (0.00)	240 (189-276)	1 (2.44)	3 (7.32)	2.10 (1.90-2.30)	9.60 (7.80-13.4)
4.00 (3.70-4.60)	0 (0.00)	227 (151-302)	1 (0.87)	6 (5.22)	1.90 (1.60-2.10)	11.3 (8.00-15.4)
4.10 (3.80-4.60)	18 (5.50)	233 (176-298)	2 (0.61)	10 (3.06)	1.90 (1.70-2.10)	10.7 (7.70-13.8)
4.30 (3.80-4.70)	44 (72.1)	207 (151-263)	2 (3.28)	5 (8.2)	2.20 (1.90-2.40)	8.90 (7.60-12.8)
4.60 (3.90-4.92)	1 (3.57)	219 (181-289)	0 (0.00)	24 (85.7)	2.00 (1.70-2.32)	11.7 (10.25-16.3)
4.10 (3.80-4.50)	7 (3.78)	222 (167-279)	9 (4.86)	7 (3.78)	1.90 (1.70-2.10)	9.50 (6.80-13.0)

Valvulopathy	Serum Sodium	Sepsis	Right Atrial Dilation	Rheumatism	Respiratory Failure	Renal insufficiency	Prothrombin time
30 (13.22)	138 (135-141)	84 (37.0)	17 (7.49)	2 (0.88)	58 (25.6)	168 (74.0)	15.9 (13.8-25.3)
15 (10.95)	138 (136-141)	6 (4.38)	29 (21.2)	5 (3.65)	23 (16.8)	12 (8.76)	14.0 (12.9-21.4)
58 (7.12)	139 (137-141)	41 (5.03)	38 (4.66)	9 (1.10)	16 (1.96)	35 (4.29)	13.9 (12.7-17.2)
14 (48.3)	138 (134-142)	12 (41.4)	10 (34.5)	0 (0.00)	5 (17.2)	19 (65.5)	15.8 (13.4-23.5)
172 (39.3)	139 (136-141)	23 (5.25)	279 (63.7)	24 (5.48)	48 (11.0)	229 (52.3)	14.2 (13.0-16.3)
7 (6.60)	139 (136-142)	6 (5.66)	64 (60.4)	1 (0.94)	9 (8.49)	40 (37.7)	14.2 (12.8-21.6)
204 (33.7)	138 (135-141)	118 (19.5)	463 (76.53)	37 (6.12)	68 (11.2)	409 (67.6)	15.7 (13.7-21.6)
9 (22.0)	137 (134-140)	9 (22.0)	35 (85.4)	0 (0.00)	7 (17.1)	31 (75.6)	14.7 (13.3-22.3)
3 (2.61)	140 (137-142)	18 (15.7)	2 (1.74)	1 (0.87)	112 (97.4)	10 (8.70)	14.2 (13.2-17.3)
39 (11.9)	139 (137-142)	25 (7.65)	6 (1.83)	25 (7.65)	60 (18.4)	66 (20.2)	14.5 (13.0-19.5)
56 (91.8)	139 (137-141)	0 (0.00)	0 (0.00)	2 (3.28)	5 (8.20)	50 (82.0)	14.1 (13.4-16.1)
1 (3.57)	139 (136-141)	12 (42.9)	18 (64.3)	1 (3.57)	5 (17.9)	24 (85.7)	15.0 (14.2-20.8)
26 (14.1)	139 (137-141)	21 (11.4)	6 (3.24)	3 (1.62)	20 (10.8)	18 (9.73)	15.2 (13.2-21.4)

Age	Male Sex	Arrhythmia History	Heart Rate	Hourly Rhythm Control BBs	Hourly Rhythm Control PCBs	Hourly Rhythm Control CCBs
75.6 (69.1-83.3)	104 (45.8)	104 (45.8)	87 (75.0-103)	1.28 (0.56-3.31)	8.76 (4.64-17.1)	3.40 (0.98-5.33)
81.8 (73.4-87.3)	15 (11.0)	45 (32.9)	85.0 (71.0-99.0)	3.81 (1.92-10.8)	6.56 (2.02-45.5)	6.60 (3.73-18.6)
76.5 (65.2-84.7)	264 (32.4)	18 (2.21)	87.0 (74.0-103)	3.41 (2.44-5.05)	8.74 (6.45-16.4)	4.09 (3.32-6.84)
80.1 (65.8-84.5)	1 (3.45)	3 (10.3)	91.0 (78.0-108)	1.90 (0.00-5.60)	13.7 (9.57-42.6)	0.12 (0.00-7.15)
74.0 (66.7-79.8)	399 (91.1)	84 (19.2)	84.0 (75.0-93.0)	3.38 (2.20-4.52)	8.08 (6.17-13.8)	2.53 (0.96-5.87)
84.4 (77.4-88.6)	14 (13.2)	67 (63.2)	89.5 (76.0-110)	3.93 (1.46-6.16)	3.64 (0.00-40.0)	3.91 (2.07-6.50)
78.3 (70.0-84.7)	542 (89.6)	83 (13.7)	86.3 (74.0-104)	1.58 (1.13-2.24)	5.22 (3.21-8.77)	3.77 (2.98-5.93)
80.3 (73.6-85.5)	10 (24.4)	0 (0.00)	86.0 (70.0-104)	2.88 (0.92-6.58)	6.08 (6.08-6.08)	1.95 (0.11-5.65)
73.6 (63.1-81.8)	71 (61.7)	11 (9.57)	91.0 (79.0-108)	2.11 (0.092-6.77)	9.22 (5.84-30.9)	5.86 (3.74-25.2)
81.7 (74.4-87.4)	110 (33.6)	13 (3.98)	83.0 (70.0-96.0)	2.14 (1.08-3.35)	7.00 (3.25-37.1)	4.58 (2.41-7.7)
75.4 (67.2-79.8)	22 (36.1)	7 (11.5)	86.0 (75.0-90.0)	0.00 (0.00-0.00)	15.0 (9.61-35.6)	NaN
84.9 (81.2-87.7)	17 (60.7)	5 (17.9)	76.0 (68.5-103)	5.62 (1.45-48.0)	0.00 (0.00-7.60)	0.00 (0.00-0.00)
77.0 (65.3-84.2)	143 (77.3)	157 (84.9)	86.0 (71.0-100)	1.63 (0.62-3.82)	10.2 (4.12-46.6)	4.34 (1.66-8.48)]

Daily Mortality Rate BBs	Hourly Rate Control MgS	Hourly Rate Control CCBs	Hourly Rate Control PCBs	Hourly Rate Control BBs	Hourly Rhythm Control MgS
1.32 (0.54-3.24)	7.07 (0.00-100)	12.0 (7.37-32.4)	9.97 (6.49-25.2)	3.17 (1.46-17.0)	0.64 (0.027-3.89)
0.66 (0.00-1.78)	33.3 (33.3-33.3)	21.7 (1.06-86.9)	11.7 (0.00-169)	9.99 (5.78-30.8)	0.19 (0.00-5.78)
1.24 (0.64-2.17)	4.45 (0.00-41.2)	19.2 (13.2-33.6)	13.3 (6.67-24.1)	10.3 (4.94-16.7)	1.59 (0.62-3.55)
6.71 (0.00-14.7)	50.0 (50.0-50.0)	6.31 (0.00-137)	8.75 (1.92-51.8)	29.5 (25.0-100)	0.00 (0.00-0.00)
0.73 (0.26-1.31)	16.5 (8.58-100.0)	9.12 (4.46-25.8)	14.9 (9.96-26.1)	19.2 (9.97-32.0)	3.29 (1.01-7.45)
1.85 (0.76-3.45)	16.6 (9.20-55.5)	6.38 (2.86-21.3)	14.7 (0.00-265)	5.77 (0.00-13.5)	0.00 (0.00-2.41)
1.61 (0.92-2.02)	13.1 (7.12-31.1)	13.0 (8.56-28.8)	15.1 (7.30-34.8)	5.28 (3.80-11.2)	2.08 (1.27-3.53)
2.08 (0.14-11.0)	13.3 (13.3-13.3)	21.8 (14.8-157)	117 (100-130)	12.9 (4.23-56.5)	0.51 (0.00-12.5)
3.21 (0.93-10.5)	21.8 (7.48-76.9)	14.8 (7.77-35.4)	8.35 (0.00-24.8)	7.74 (1.19-50.0)	3.47 (1.20-23.3)
1.73 (1.03-3.06)	9.17 (0.00-20.6)	14.5 (6.39-40.1)	6.58 (0.00-77.0)	17.3 (6.38-24.9)	0.00 (0.00-0.73)
14.3 (0.00-14.3)	0.00 (0.00-0.00)	NaN	8.09 (1.73-27.3)	25.0 (25.0-25.0)	2.93 (0.00-5.00)
3.04 (0.00-26.7)	0.00 (0.00-0.00)	215 (146-600)	8.93 (7.55-8.96)	9.67 (2.50-300)	1.68 (0.00-33.3)
0.00 (0.00-0.42)	100 (100-100)	19.2 (6.33-71.7)	14.1 (1.98-29.0)	15.8 (6.33-152)	1.41 (0.093-4.91)

Daily Mortality Rate PCBs	2.31 (0.82-6.88)	3.48 (0.61-8.40)	0.85 (0.30-1.70)
Daily Mortality Rate CCBs	1.92 (0.62-5.64)	0.53 (0.00-2.62)	0.86 (0.44-1.22)
Daily Mortality Rate MgS	2.88 (1.26-8.51)	0.14 (0.00-5.64)	1.55 (0.96-4.22)

Daily Mortality Rate PCBs	4.81 (0.94-12.1)	0.96 (0.64-1.81)	0.14 (0.00-6.78)
Daily Mortality Rate CCBs	6.06 (0.00-150)	0.089 (0.00-0.58)	1.26 (0.25-2.86)
Daily Mortality Rate MgS	13.8 (0.00-33.3)	0.00 (0.00-0.00)	0.00 (0.00-1.99)

Daily Mortality Rate PCBs	2.44 (1.28-3.88)	0.00 (0.00-0.00)	0.87 (0.00-2.63)
Daily Mortality Rate CCBs	0.84 (0.55-1.77)	2.45 (0.00-11.7)	2.37 (0.48-5.65)
Daily Mortality Rate MgS	1.03 (0.61-1.89)	0.00 (0.00-0.00)	1.17 (0.14-6.46)

Daily Mortality Rate PCBs	2.64 (0.48-11.4)	1.40 (0.043-2.58)	8.46 (6.36-48.0)
Daily Mortality Rate CCBs	2.22 (1.00-4.04)	NaN	1.86 (0.00-14.2)
Daily Mortality Rate MgS	1.01 (0.067-15.3)	9.49 (5.56-25.0)	9.42 (7.69-14.3)

Daily Mortality Rate PCBs	0.00 (0.00-0.00)	
Daily Mortality Rate CCBs	1.12 (0.00-3.17)	
Daily Mortality Rate MgS	2.98 (1.03-12.0)	