UNIVERSITAT DE BARCELONA

Application of Radiomics-based Machine Learning models on complex cardíac diseases

Cristian Izquierdo

UNIVERSITY OF BARCELONA

DOCTORAL THESIS

Application of Radiomics-based Machine Learning Models on Complex Cardiac Diseases

Author:
Cristian IZQUIERDO

Supervisor:
Karim LEKADIR
Polyxeni GKONTRA

UNIVERSITAT DE BARCELONA

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Mathematics and Computer Science Department

April 23, 2024

Declaration of Authorship

I, Cristian IZQUIERDO, declare that this thesis titled, "Application of Radiomics-based Machine Learning Models on Complex Cardiac Diseases" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- 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.

Signed: Cristian Izquierdo Morcillo

Date: 23/04/2024

""Some believe it is only great power that can hold evil in check, but that is not what I have found. I have found that it is the small everyday deeds of ordinary folk that keep the darkness at bay. Small acts of kindness and love.""

Gandalf. The Lord of The Rings.

UNIVERSITY OF BARCELONA

Abstract

Mathematics and Computer Science
Mathematics and Computer Science Department

Doctor of Philosophy

Application of Radiomics-based Machine Learning Models on Complex Cardiac Diseases

by Cristian IZQUIERDO

This doctoral thesis explores the integration of radiomics and machine learning (ML) in cardiology, focusing on the early detection and prognosis of complex cardiovascular diseases (CVDs). Radiomics transforms conventional medical images into rich, high-dimensional data representations that reveal intricate details of cardiac pathologies not visible to the naked eye. To facilitate systematic analysis and interpretation, we meticulously designed a pipeline to extract radiomic features and harness ML for classification and interpretation tasks. For each chapter of the thesis, we tailored and refined this radiomics analysis pipeline with ML, adapting it to suit varying scenarios, and ultimately transforming it into a survival analysis ML pipeline for the final chapter. The thesis begins with an overview of cardiovascular diseases, radiomics and ML, setting the foundation for their application in cardiac imaging.

The second chapter of this thesis demonstrates the utility of cardiovascular magnetic resonance (CMR) radiomics and ML in distinguishing left ventricular noncompaction cardiomyopathy (LVNC) from hypertrophic and dilated cardiomyopathies. The following chapter reveals the potential of combining CMR radiomics with electrocardiogram (ECG) data for improved atrial fibrillation (AF) detection, particularly enhancing accuracy among women. The fourth chapter assesses the ability of CMR radiomics and ML to predict major cardiovascular events like AF, heart failure (HF), myocardial infarction (MI), and stroke, utilizing UK Biobank data. The incorporation of radiomic features with vascular risk factors (VRFs) and CMR indices significantly boosts the predictive models' performance.

Lastly, the last chapter highlights the advantages of radiomics and ML in identifying genetic cardiomyopathy in excessively trabeculated patients, showing superior accuracy over traditional methods and predicting the risk and timing of Major Adverse Cardiac Events (MACE) more effectively. This thesis showcases the promise of radiomics and ML in advancing cardiac diagnostics and prognostics, offering a more precise, personalized approach to managing CVDs.

Acknowledgements

- To Karim Lekadir. Five years ago, I started my journey with my supervisor, and it has been filled with challenges. I am thankful our paths crossed. His guidance has transformed me. I am grateful for his patience, for challenging my research and personal beliefs, and for his support when I needed it most. Someday, the details of my thesis may fade from memory, but who I have become at the end of this journey will not. I truly mean it when I say that I owe all of this to him.

- I am grateful to my secondary supervisor, Polyxeni GKONTRA, whose unwavering support and assistance illuminated my path throughout this entire journey.

- My profound gratitude to Victor and Carlos. From the beginning of our journey, we have supported and motivated each other relentlessly. The number of times they have aided me in advancing my research is immeasurable. My appreciation for their help knows no bounds.

- To my family, I express my deepest gratitude. Your unwavering support and faith in me has been my foundation, providing the stability I needed throughout this journey. Your love and encouragement have been invaluable, reminding me that I am not alone in this pursuit. Thank you for standing by me every step of the way.

- To Cristina and Luke Skywalker, my Force. There is no better Jedi family than this.

Contents

List of Figures

List of Tables

List of Abbreviations

Heart-related abbreviations

CVD	**C**ardio **V**ascular **D**isease
CM	**C**ardio**M**yopathy
LVNC	**L**eft **V**entricle **N**on-**C**ompaction
HCM	**H**ypertrophic **C**ardio**M**yopathy
DCM	**D**ilated **C**ardio**M**yopathy
AF	**A**trial **F**ibrillation
HF	**H**eart **F**ailure
MI	**M**iocardial**I**nfarction
CMR	**C**ardiac **M**agnetic **R**esonance
LV	**L**eft**V**entricle
RV	**R**ight **V**entricle
MYO	**M**yocardium

Machine Learning abbreviations

AI	**A**rtificial **I**ntelligence
ML	**M**achine **L**earning
RF	**R**andom **F**orest
XGB	**X G B**oost
SVM	**S**upport **V**ector **M**achine
CV	**C**ross-**V**alidation

In honor of a truly good man, a beacon of optimism, filled with kindness and gentleness. In our final conversation, he uplifted my spirit. Wherever you are now, I hope these words find their way to you. May you have finally found the peace you deserved.

Chapter 1

Introduction

1.1 Introduction

The field of healthcare, despite being an area of continuous evolution and progress, still faces substantial hurdles in effectively addressing and treating life-threatening diseases Gawande, 2014; Chabner, 2016. In parallel, the realm of computer science is forging ahead, setting the stage for a future where artificial intelligence (AI) permeates every facet of our daily existence (Russell and Norvig, 2020).

Consider a future where AI is capable of combining thousands of medical records within seconds, unveiling patterns and trends that, in the past, would have been impossible to discern due to human limitations in data processing (Rajkomar, Dean, and Kohane, 2019; Obermeyer and Emanuel, 2016). This profound analytical power of AI could revolutionize diagnostic accuracy, patient care, and prognostic prediction, driving improvements across the entire healthcare continuum (Deo, 2015).

Moreover, envision a world where personalized medicine is no longer an ambitious ideal but the prevailing norm. The confluence of AI, genomics, and health informatics opens the doors for treatments tailored to an individual's unique genetic composition and health history, paving the way for more effective, precise, and less harmful therapeutic strategies.

Cardiovascular disease (CVD) is a poignant example of a global health challenge where these advancements could have significant impacts. It is currently the leading cause of mortality worldwide, contributing to an estimated 17.9 million deaths annually as of the latest World Health Organization (WHO) report in 2019 (Organization, 2019). This figure is expected to rise to over 23.6 million by 2030, reflecting the growing urgency of the problem (Roth et al., 2017). Genetic predispositions, coupled with modifiable risk factors like hypertension, high cholesterol, smoking, obesity, and physical inactivity, contribute to this burgeoning global health crisis (Yusuf et al., 2004; Mozaffarian et al., 2016).

The integration of AI into cardiovascular medicine could serve as a powerful tool in risk prediction, early detection, and management of CVD. Machine learning algorithms can analyze large volumes of data from electronic health records (EHRs), genetic databases, and wearable technology to identify risk factors and predict patient outcomes with remarkable precision (Krittanawong et al., 2017; Attia et al., 2019b). Moreover, advancements in genomics and pharmacogenomics offer the promise of personalized treatment strategies for CVD, considering an individual's genetic makeup in drug selection and dosing (Roden and Johnson, 2016; Luzum and Peterson, 2016).

As we move further into the 21st century, the application of AI and personalized medicine in clinical practice, particularly in combatting the devastating impacts of CVD (), presents an inspiring vision of the future of healthcare. Yet, it is important to recognize that realizing this vision will require ongoing research, multidisciplinary

collaboration, ethical considerations, and policy development to ensure these technologies are accessible, equitable, and beneficial for all(Organization, 2019).

The impact of AI on cardiology is profound, enhancing the early detection and diagnosis of heart diseases, optimizing treatment strategies, and paving the way for groundbreaking discoveries in heart health. AI's unparalleled ability to sift through and analyze vast datasets offers cardiologists critical insights into heart function and disorders, enabling more precise and timely interventions. This technological revolution in cardiac care is just beginning to unfold, promising to transform our approach to preventing, managing, and potentially curing cardiovascular diseases. Through dedicated research in this field, AI is set to significantly improve cardiac health outcomes and the quality of life for individuals worldwide, marking the dawn of a new era in cardiology.

However, cardiovascular imaging presents its own specific challenges. The use of AI presents a significant opportunity to revolutionize cardiovascular imaging, addressing longstanding challenges in the field to enhance diagnostic accuracy and patient outcomes. Traditional methods of cardiovascular imaging, while invaluable, often grapple with limitations such as variability in image interpretation, time-consuming analysis, and the potential for human error. AI, particularly machine learning and deep learning algorithms, can surmount these obstacles by enabling more precise, efficient, and reproducible image analysis. These technologies have the potential to automatically detect and quantify cardiovascular abnormalities, predict disease progression, and guide treatment decisions based on vast datasets that no human could feasibly analyze in a lifetime. Furthermore, AI-driven tools can streamline workflow efficiency, reducing the time from imaging to diagnosis and treatment, thus improving patient care and reducing healthcare costs. The integration of AI into cardiovascular imaging is not without its challenges, including the need for robust, high-quality datasets for algorithm training, concerns about algorithmic bias, and ensuring the interpretability of AI decisions to healthcare providers. Nonetheless, the promise of AI to enhance diagnostic accuracy, personalize patient care, and optimize clinical outcomes in cardiovascular medicine is both compelling and indispensable in the face of an ever-growing cardiovascular disease burden globally.

This dissertation delves into the promising frontier of integrating machine learning with medical imaging, to develop invaluable tools for improving detection and understanding of complex diseases. More precisely, the focal point of this exploration is the field of cardiology and pathological excessive trabeculations, with a specific interest in the integration of Radiomics and Machine Learning. Radiomics is a field that exploits the high-throughput extraction of large amounts of image features from radiographic images, offering a wealth of information beyond what the human eye can perceive (Gillies, Kinahan, and Hricak, 2016).

Throughout this thesis, we explore four unique applications of radiomics across various cardiovascular diseases (CVD). Each chapter delves into the specifics of data collection methods, the machine learning (ML) strategies applied, the benefits of utilizing radiomics within these contexts, and the potential for expanding these models' applications in future research. In summation, the fusion of radiomics and machine learning promises to markedly augment our comprehension of medical imaging data, its potential for predicting patient outcomes, and its capacity to identify disease patterns. Such advancements could usher in a new era of more effective, personalized treatment strategies for a broad spectrum of conditions.

1.2 Motivation

Complex CVDs, including atrial fibrillation, heart failure, stroke, left ventricular non-compaction (LVNC), and hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), among others, present with uncertain prognostic indicators, elusive etiological factors, and ambiguous morphological characteristics. These conditions, characterized by diverse manifestations such as excessive trabeculations in the left ventricle (specific to LVNC), irregular heart rhythms, or abnormal heart muscle growth depending on the disease, pose significant challenges in diagnosis and treatment. The implications of these varied pathological features and the associated risks remain largely unexplored and undefined. The inherent complexity of these diseases necessitates rigorous investigation to elucidate their pathophysiology, prognostic pathways, potential risk factors, and improve their diagnosis and prognosis. Currently, routine clinical variables are utilized in the prognosis and diagnosis of these conditions; however, their efficacy has been somewhat limited and tedious to acquire. This thesis underscores the principle that novel methodologies, especially those rooted in AI, will be key to diagnosis purposes.

In this thesis, we attempt to bridge the gap: The fusion of radiomics and machine learning (ML) emerges as a crucial advancement in the assessment of complex cardiovascular diseases, including atrial fibrillation, heart failure, stroke, DCM and HCM, and as well as rare diseases such as LVNC. This integration transforms traditional imaging data into a comprehensive array of quantifiable features, enabling the nuanced detection and analysis of patterns characteristic of these diseases, even in the most uncovered paths such as excessive trabeculation in LVNC. By leveraging these sophisticated technologies, clinicians can potentially overcome the limitations of subjective interpretation inherent in conventional imaging methods, offering a new frontier in the precise and early detection of complex cardiovascular conditions.

1.3 Objective

The main aim of this PhD thesis is to develop novel computational solutions based on machine learning to advance the diagnosis, prognosis of complex cardiovascular diseases, including a rare cardiovascular disease called LVNC, and forecasting of complications. Towards this aim, we leverage state-of-the art machine learning techniques and advanced image analysis to optimally exploit the information from cardiac magnetic resonance imaging (CMR), recognized as the reference modality for assessing the structure and function of the heart.

More precisely, to achieve the aim of this thesis, the following main objectives and relative sub-tasks had to be fulfilled:

- The first objective of this research is to investigate the potential of radiomics and ML in diagnosing patients with a range of different cardiomyopathy and complex cardiovascular diseases. To this end, we will develop and validate diverse radiomics-based machine learning models for the classification. Also will focus on the extraction of various radiomic features from cardiac imaging data from diseased patients, feature selection for model building, and cross-validation of the model's performance. We will focus in several CVDs, and we will look further into a specific condition denominated LVNC.

- The second objective of this study is to utilize radiomic signatures to predict disease progression in a specific cardiomyopathy such LVNC. Prognosis is a

critical aspect of patient management and has been challenging due to the heterogeneity of LVNC. By leveraging radiomics, we aim to derive quantitative features from cardiac imaging data that could potentially correlate with patient outcomes. We will conduct a longitudinal study, correlating the derived radiomic features with the clinical, biochemical, and histopathological data of LVNC patients. The performance of the proposed radiomics-based prognostic model will be evaluated in terms of its concordance index, calibration, and clinical usefulness.

- The third objective is to facilitate the adoption of radiomics into routine clinical practice through empirical demonstrations of its capabilities. Despite the advancements in radiomics, its translation into clinical practice has been slow, primarily due to the lack of understanding about its interpretability and applicability. We aim to bridge this gap by conducting comprehensive empirical studies that demonstrate the robustness, reproducibility, and clinical relevance of radiomics. This will involve sensitivity analysis, reproducibility studies, as well as evaluation of the model's performance across multiple centers and imaging platforms. The clinical utility of the radiomics model will be assessed by determining its added value to the conventional diagnostic and prognostic methods.

The culmination of these objectives represents a significant contribution to the field, promoting the enhanced utilization of radiomics in clinical practice, specially with the new scenario of LVNC patients for diagnosis and prognosis, and facilitating its wider adoption in the future.

It is worth noting that this thesis was undertaken as part of the european funded H2020 program euCanSHare. Networks program "Next generation training in cardiovascular research and innovation" ("euCanSHare", grant Agreement 608027). "euCanSHare" was an innovative program that ran from 2019 until 2023 with the aim of providing a tool for analysis of cardiovascular imaging, including automatic segmentation pipelines, radiomics extraction and advanced ML modelling.

1.4 Contributions

The outcomes of this dissertation have culminated in three (two accepted, one under submission) primary-author scholarly articles.

- Izquierdo C, Casas G, Martin-Isla C, Campello VM, Guala A, Gkontra P, Rodríguez-Palomares JF and Lekadir K (2021) Radiomics-Based Classification of Left Ventricular Non-compaction, Hypertrophic Cardiomyopathy, and Dilated Cardiomyopathy in Cardiovascular Magnetic Resonance. Front. Cardiovasc. Med. 8:764312. doi: 10.3389/fcvm.2021.764312

- Rauseo, E., Izquierdo Morcillo, C., Raisi-Estabragh, Z., Gkontra, P., Aung, N., Lekadir, K., Petersen, S. E. (2021). New imaging signatures of cardiac alterations in ischaemic heart disease and cerebrovascular disease using CMR radiomics. Frontiers in Cardiovascular Medicine, 8, 716577.

- Cristian Izquierdo, Guillem Casas, Carlos Martin-Isla, Victor M. Campello, Esmeralda Ruiz Pujadas , Polyxeni Gkontra, Alberto Morales-Galan , Jesus G Mirelis, Albert Teis, Coloma Tiron, José Manuel Garcia-Pinilla, Tomás Ripoll-Vera, Juan Jiménez-Jáimez, Eduardo Villacorta, Juan Ramon Gimeno-Blanes,

Esther Zorio, Roberto Barriales-Villa, José F. Rodríguez Palomares, Karim Lekadir, Andrea Guala. Radiomics analysis of cardiac magnetic resonance images for the detection of genetic and familial cases in excessive trabeculation of the left ventricle. On review process.

In addition to the primary authorships, significant contributions were made as part of this thesis as a co-author to several other publications related to this field of study, including the works presented in chapters 4 and 5:

- Pujadas, E. R., Raisi-Estabragh, Z., Szabo, L., McCracken, C., Morcillo, C. I., Campello, V. M., ... and Lekadir, K. (2023). Prediction of incident cardiovascular events using machine learning and CMR radiomics. European radiology, 33(5), 3488-3500.

- Pujadas, E.R., Raisi-Estabragh, Z., Szabo, L. et al. Atrial fibrillation prediction by combining ECG markers and CMR radiomics. Sci Rep 12, 18876 (2022). https://doi.org/10.1038/s41598-022-21663-

  Additionally, the author of this thesis have also participated in several other publications as co-author:

- Campello, V. M., Gkontra, P., Izquierdo, C., Martin-Isla, C., Sojoudi, A., Full, P. M., ... Lekadir, K. (2021). Multi-centre, multi-vendor and multi-disease cardiac segmentation: the MMs challenge. IEEE Transactions on Medical Imaging, 40(12), 3543-3554.

- Martin-Isla, C., Campello, V. M., Izquierdo, C., Raisi-Estabragh, Z., Baeßler, B., Petersen, S. E., Lekadir, K. (2020). Image-based cardiac diagnosis with machine learning: a review. Frontiers in cardiovascular medicine, 7, 1.

- Raisi-Estabragh, Z., Izquierdo, C., Campello, V. M., Martin-Isla, C., Jaggi, A., Harvey, N. C., ... Petersen, S. E. (2020). Cardiac magnetic resonance radiomics: basic principles and clinical perspectives. European Heart Journal-Cardiovascular Imaging, 21(4), 349-356.

- Campello, V. M., Martín-Isla, C., Izquierdo, C., Petersen, S. E., Ballester, M. A. G., Lekadir, K. (2020). Combining multi-sequence and synthetic images for improved segmentation of late gadolinium enhancement cardiac MRI. In Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges: 10th International Workshop, STACOM 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Revised Selected Papers 10 (pp. 290-299). Springer International Publishing.

- Martín-Isla, C., Campello, V. M., Izquierdo, C., Kushibar, K., Sendra-Balcells, C., Gkontra, P., ... Lekadir, K. (2023). Deep learning segmentation of the right ventricle in cardiac mri: The mms challenge. IEEE Journal of Biomedical and Health Informatics, 27(7), 3302-3313.

- Campello, V. M., Martín-Isla, C., Izquierdo, C., Guala, A., Palomares, J. F. R., Viladés, D., ... Lekadir, K. (2022). Minimising multi-centre radiomics variability through image normalisation: a pilot study. Scientific reports, 12(1), 12532.

Furthermore, the results of this work have contributed in the development of a web-based platform specifically designed for cardiovascular research analysis, including radiomics analysis and machine learning tools. The platform is accessible

at at https://vre.eucanshare.bsc.es/vre/home/and was developed as part of the H2020 euCanShare project.

1.5 Document organization

The current document is organized into 6 distinct chapters. The current chapter (Chapter 1) introduces the motivation for the development of this PhD thesis, its principal objectives and contributions.

Chapter 2 is dedicated to providing both the theoretical and technical groundwork that underpins the entirety of this thesis. This includes an in-depth exploration of the heart's primary characteristics, its pathophysiology, and a detailed analysis of several CVDs and cardiomyopathy diseases, whose prognosis or diagnosis is being addressed in this thesis. Subsequently, the chapter delves into the specifics of radiomics, expounding upon its definition, the methodologies employed for its extraction, and the subsequent utilization of this data in a clinical and research context. Moreover, we further investigate the principles of ML modeling, including a comprehensive description of the algorithms leveraged in this thesis. The aim is to elucidate how these computational tools can be used to decode the complexities of radiomics data and the heart's structure and function. To conclude this initial chapter, we enumerate state-of-the-art publications in the current literature within the domain of cardiology, and further honing in on its use in studying and managing CVDs. In each chapther, literature review was provided for each specific topic.

Chapter 3 focuses on the application of Machine Learning (ML) classification algorithms and radiomics in differentiating of LVNC, hypertrophic cardiomyopathy and dilated cardiomyopathy, using CMR. This chapter is partitioned into four subsections, each illustrating a unique classification target that leverages the interplay of radiomics and ML. We demonstrate that radiomics-based ML models outperform traditional CMR indices used in the clinic, including manually estimated trabecular indices, for the task at hand. Identifying the correct cardiomyopathy is a critical task for clinicians, as it lays the groundwork for accurate diagnosis and effective treatment. This step is vital in enhancing patient outcomes and increasing survival rates. Proper classification of the cardiomyopathy not only streamlines the diagnostic process but also ensures that patients receive the most appropriate and timely care, ultimately leading to improved quality of life and health prospects.

Chapter 4 studies the relationship between CMR radiomics in combination with ECG for atrial fibrillation (AF) prediction. This study explores the use of machine learning models to improve the detection of AF, the most common cardiac arrhythmia associated with serious health risks such as stroke and death. Traditionally, AF diagnosis relies on electrocardiograms (ECG), which may miss paroxysmal AF due to their time-limited nature. By combining image-derived radiomics phenotypes with ECG features, the research develops a more effective model for identifying AF, particularly focusing on the differences in heart remodeling between sexes. The integrated radiomics-ECG model outperformed the ECG-alone method, especially in women, where it significantly improved both accuracy and sensitivity in detecting AF. This novel approach offers deeper insights into AF's electro-anatomic remodeling and suggests a more efficient strategy for early AF detection, highlighting the importance of considering sex-specific differences in cardiac health assessments.

Chapter 5 is focused in a range of more wide-spread complex cardiac diseases. This study assesses the potential of cardiovascular magnetic resonance (CMR) radiomics, combined with machine learning techniques, to predict the occurrence of

major cardiovascular conditions, including atrial fibrillation (AF), heart failure (HF), myocardial infarction (MI), and stroke. Utilizing data from the UK Biobank, the research involved participants who experienced these cardiovascular diseases (CVDs) during follow-up. The approach included analyzing CMR images to extract radiomics features from specific regions of the heart and integrating these with vascular risk factors (VRFs) and CMR indices to create predictive models.

The findings revealed that the combined model incorporating VRFs, CMR indices, and radiomics features (VRF+CMR+Rad) showed superior performance in predicting AF, with notable accuracy and area under the curve (AUC) metrics. For HF, the inclusion of CMR metrics significantly enhanced the model's effectiveness. The study also demonstrated that adding radiomics features to VRFs alone yielded comparably strong predictive capabilities for HF, indicating the substantial value of radiomics in forecasting incident CVDs. Although improvements in predicting MI and stroke were more modest, the results overall underscore the incremental predictive value of radiomics features when combined with traditional risk factors and imaging indices in identifying future cardiovascular events.

Chapter 6 highlights the efficacy of integrating radiomics and machine learning (ML) in two different scenarios. The first consist in (i) identifying genetic cardiomyopathy in patients with excessive trabeculation (LVNC). Our research demonstrates that radiomics surpasses conventional clinical indices in performance. Traditional methods not only involve laborious processes but often necessitate the administration of contrast agents. Our proposed methodology not only enhances accuracy in detection but also streamlines the diagnostic process through automation. Furthermore, it eliminates the need for contrast agents, contributing to cost efficiency and reducing potential risks associated with their use. This approach marks a significant step forward in cardiac diagnostics, offering a more efficient, accurate, and safer alternative. The second scenario delves into the prognosis analysis of patients with excessive trabeculation (LVNC)(ii). In many cases, the extent of trabeculation does not provide clear insights into the patient's prognosis or the disease's progression. Our method utilizes radiomics combined with Survival Analysis Machine Learning models to enhance the prediction of the timeline and likelihood of Major Adverse Cardiac Events (MACE). This approach is crucial as it aids in anticipating the disease course, allowing for early intervention and potentially averting fatal outcomes. By accurately predicting the prognosis, clinicians can implement preemptive measures, significantly improving patient management and reducing the risk of severe cardiac incidents. This approach is geared towards providing a robust understanding of the prognosis of LVNC when these temporal factors are taken into consideration. By incorporating the temporal aspect into the analysis, this chapter aims to shed light on the prognostic trajectory of patients with LVNC over time, and elucidate how different factors may influence this trajectory. This will help to underscore the complexity and dynamic nature of LVNC prognosis, and provide a foundation for improving patient management strategies.

Chapter 2

Background and state-of-the-art

2.1 Clinical context

2.1.1 Demographical context

Cardiovascular diseases (CVD), a heterogeneous group of diseases affecting the heart muscle and its system, have emerged as a leading cause of global morbidity and mortality (see Fig. 2.1). Despite advances in cardiovascular medicine, the burden of these heart muscle disorders remains substantial, contributing significantly to the spectrum of heart failure cases worldwide. The insidious nature of CVD often leads to delayed diagnosis and management, underscoring a silent epidemic with a profound impact on public health. (Organization, 2021; Organization, 2019) The prevalence of CVD is challenging to ascertain due to varied diagnostic crite ria and the evolution of genetic testing; however, estimates suggest that CVD could affect as many as 1 in 500 individuals. The implications are grave, with heart failure resulting from dilated cardiomyopathy the most common form accounting for the largest number of heart transplants annually. Hypertrophic cardiomyopathy, on the other hand, although less prevalent, stands as a common cause of sudden cardiac death, especially among young athletes. The etiology of CVD is multifaceted, involving genetic predisposition, lifestyle factors, and possible environmental triggers. Genetic advancements have unveiled a myriad of mutations associated with these conditions, revealing a complex inter-play between genotype and phenotype. Yet, the translation of these findings into preventive strategies remains in its infancy. The management of CVD has evolved, focusing on mitigating risk factors, controlling symptoms, and halting disease progression. Therapeutic strategies encompass a spectrum from pharmacologic treatments to advanced interventions like implantable devices and heart transplantation. Despite such interventions, the five-year survival rate post diagnosis for certain CVD remains as low as 50%, reflecting the severity of these conditions. Global health systems face a considerable challenge in addressing the rise of CVD. Early detection and targeted therapies are imperative to alter the trajectory of this first-rate killer. Public health initiatives aimed at raising awareness and promoting cardiovascular health could pivot the direction towards better outcomes. As research delves deeper into the molecular underpinnings of these diseases, there is cautious optimism that personalized medicine may herald a new era in the management of CVD. The intertwining of robust epidemiological data and emerging biotechnological tools promises a future where the impact of these formidable diseases is significantly lessened.

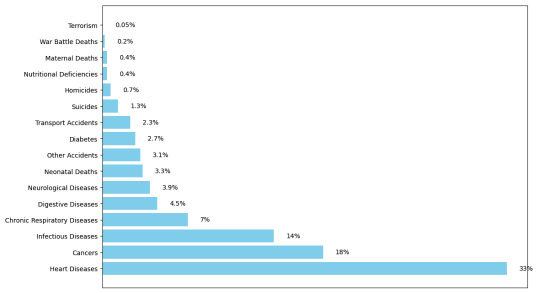

FIGURE 2.1: Causes of Death Globally in 2019. Data from IHME Global Burden of Disease and Global Terrorism Database and World Health Organization. The size of each bar represents the share of deaths due to a particular cause.

2.1.2 Heart structure and physiopathology

The heart is a vital muscular entity nestled within the thoracic cavity, tasked with the crucial job of propelling blood throughout the entire body (Buckberg et al., 2018; Torrent-Guasp et al., 2005). This dynamic organ ensures efficient delivery of oxygen and essential nutrients to all tissues while simultaneously aiding in the removal of metabolic waste. Comprised of four distinct sections - the left and right atria situated above and the larger left and right ventricles below - it functions as a superbly coordinated unit. The journey of blood commences in the right atrium, from where it travels to the right ventricle, making its way to the lungs for oxygen enrichment. This oxygen-laden blood returns to the heart's left atrium, gets pumped into the left ventricle, and is subsequently dispatched to nourish the body. Thus, the heart plays an indispensable role in upholding systemic circulation and overall wellness.

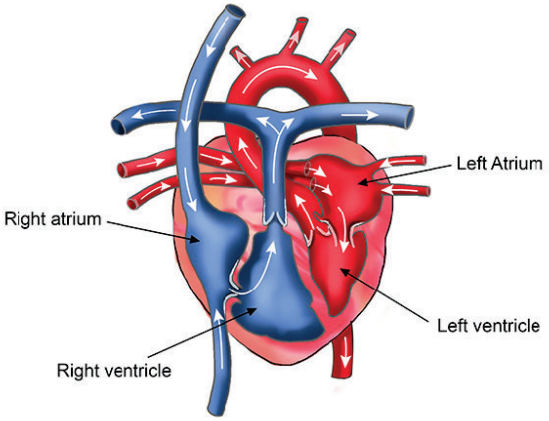

FIGURE 2.2: Cardiac cavities.

The heart is a complex organ consisting of multiple morphological regions, with the main divisions being the myocardium (MYO), the left ventricle (LV), and the right ventricle (RV). When acquiring images of the heart, cardiologists often visualize it from different perspectives depending on the specific pathology or trait they

are examining. In our study, we will focus on the Short-Axis (SAX) view (Figure 2.3), which encompasses the key features that are most relevant to our research objective.

FIGURE 2.3: Left image is the original MRI SAX view. Right image represents the cardiac cavities segmented in a Short Axis (SAX) MRI view. Green region represents myocardium (MYO), blue region represents Left Ventricle (LV) and yellow region represents Right Ventricle (RV).

2.1.3 Cardiovascular disease (CVD)

CVD refers to a class of diseases that involve the heart or blood vessels (Flora and Nayak, 2019; Lennon, Claussen, and Kuersteiner, 2018). CVD encompasses a wide range of conditions, including coronary artery disease, which affects the blood supply to the heart; arrhythmias, or disorders of heart rhythm; heart failure, where the heart is unable to pump blood efficiently; congenital heart disease, which is present from birth; and stroke, which arises from problems with blood flow to the brain, among others. The underlying mechanisms vary depending on the specific condition, but they often involve processes such as atherosclerosis (the buildup of plaque in the arteries), hypertension (high blood pressure), or genetic factors that affect heart function.

Risk factors for developing CVD include smoking, lack of exercise, obesity, high blood pressure, high cholesterol, diabetes, and family history of CVD (Flora and Nayak, 2019). Prevention and treatment strategies often focus on lifestyle modifications such as improving diet, increasing physical activity, and quitting smoking, in addition to medical interventions like medications to manage risk factors (e.g., blood pressure, cholesterol) and surgical procedures to correct or mitigate heart damage (Arbelo et al., 2023).

Early detection and management of risk factors are crucial for preventing or delaying the onset of CVD. Advances in medical research, including the use of technologies such as radiomics and machine learning (ML) in analyzing cardiac imaging, are enhancing our ability to diagnose, treat, and understand the complexities of cardiovascular diseases, potentially leading to improved patient outcomes and reduced mortality rates associated with these conditions. Depicted below can be found a brief description of the CVDs discussed in this thesis.

Atrial Fibrilation (AF)

Atrial fibrillation (AF) is a common cardiac arrhythmia characterized by rapid and irregular beating of the atrial chambers of the heart (Brundel et al., 2022). It represents a major public health (Chugh et al., 2014; Inohara et al., 2018) concern due to its association with increased morbidity and mortality, primarily from stroke and heart failure. The pathophysiology of AF involves a complex interplay of electrical, structural, and contractile remodeling of the atria, which facilitates the initiation and maintenance of the arrhythmia.

AF is the most prevalent sustained cardiac arrhythmia, affecting millions of individuals worldwide. Its prevalence increases with age, making it a significant burden in the aging population. Key risk factors for AF include hypertension, diabetes mellitus, obesity, sleep apnea, chronic kidney disease, and structural heart diseases such as valvular heart disease and heart failure (Geelhoed et al., 2020).

The underlying mechanism of AF involves multiple pathways that lead to aberrant atrial electrical activity and structural remodeling. Triggers for AF typically originate in the pulmonary veins, where ectopic beats can initiate the arrhythmia. The perpetuation of AF is supported by changes in atrial tissue that promote electrical re-entry and continuous activation of the atria. This electrical remodeling, coupled with structural changes such as fibrosis, contributes to the maintenance of AF.

AF can manifest in a variety of ways, ranging from asymptomatic episodes to significant palpitations, fatigue, dyspnea, and reduced exercise tolerance. The irregular heartbeat of AF can be detected on physical examination, with confirmation by electrocardiogram (ECG (Aizawa, Watanabe, and Okumura, 2017) showing absent P waves, irregular R-R intervals, and rapid atrial activity.

Management strategies for AF aim at preventing thromboembolic events, controlling heart rate, and restoring and maintaining sinus rhythm. Anticoagulation is a cornerstone in the management of AF to prevent stroke, with the choice of agent guided by risk stratification scores. Rate control, typically with beta-blockers, calcium channel blockers, or digoxin, is essential for symptomatic relief and to prevent tachycardia-induced cardiomyopathy. Rhythm control strategies, including pharmacologic agents and catheter-based ablation procedures, are considered for symptom management and in specific clinical scenarios.

Atrial fibrillation is a complex arrhythmia with significant implications for individual and public health. The push for innovation in imaging seeks to address the inherent challenges in AF management by providing clearer insights into the arrhythmia's onset, progression, and response to treatment. For instance, advancements in cardiac imaging are expected to improve the precision of diagnosing AF, assessing its severity, and monitoring its treatment response over time. This involves not just the refinement of existing techniques but also the introduction of entirely new imaging technologies (Bertelsen et al., 2020; Baalman et al., 2020) that can offer unprecedented views of the heart's electrical and structural characteristics (Bumgarner et al., 2018).

Furthermore, the development of novel imaging approaches is crucial for tailoring therapies to individual patient needs, thereby enhancing outcomes for those suffering from this prevalent arrhythmia. Through detailed visualization of the atrial structure and function, healthcare providers can better identify patients who would benefit most from specific interventions, whether they be pharmacological treatments, catheter ablation, or surgical procedures.(Dörr et al., 2019).

Stroke

Stroke represents a critical medical condition characterized by the abrupt loss of brain function due to a disturbance in the blood supply to the brain (Minnis and Quinn, 2024; Mozaffarian et al., 2016). This disturbance can result from either ischemia, due to blockage (as seen in ischemic stroke), or hemorrhage, as a result of bleeding (in hemorrhagic stroke). Stroke is a leading cause of disability and mortality globally, underscoring the importance of rapid diagnosis and treatment.

The incidence of stroke varies worldwide but is a significant health issue, particularly in older populations (Aparicio, Benjamin, Callaway, et al., 2021). Risk factors for stroke can be categorized into modifiable and non-modifiable factors. Modifiable risk factors include hypertension, diabetes mellitus, smoking, obesity, physical inactivity, and atrial fibrillation (Bikkina et al., 1995). Non-modifiable risk factors encompass age, gender, ethnicity, and genetic predisposition.

The pathophysiology of stroke depends on its type. Ischemic stroke, which accounts for approximately 80% of all strokes, occurs when a blood clot obstructs a blood vessel supplying the brain, leading to a deficiency in blood flow (ischemia) and resulting in cell death. Hemorrhagic stroke, on the other hand, results from the rupture of a blood vessel within or surrounding the brain, causing bleeding into or around the brain and consequent damage to brain tissue (Gosmanova, Mikkelsen, Molnar, et al., 2016).

Preventive strategies targeting modifiable risk factors are critical in reducing the risk of first or recurrent stroke (Flueckiger et al., 2018; Kim, Shim, Park, et al., 2016). These include lifestyle modifications, pharmacological treatment for hypertension, diabetes, and dyslipidemia, and, in some cases, anticoagulation for individuals with atrial fibrillation.

Stroke is a medical emergency that requires prompt recognition and treatment to reduce the risk of mortality and long-term disability. Efforts in public health aimed at stroke prevention are crucial, but equally important is the integration of imaging techniques for early detection and prognosis assessment. Alongside advancements in acute stroke management and rehabilitation, leveraging imaging technologies becomes imperative to mitigate the impact of strokes on both individuals and society. The ongoing research and development of novel therapeutic approaches are further enhanced by the integration of imaging, promising improved outcomes for stroke patients.

Myocardial infarction

Myocardial infarction (MI), commonly known as a heart attack, is a severe medical condition that occurs when blood flow to a part of the heart is abruptly blocked, leading to the death of heart muscle tissue. This condition not only represents a leading cause of morbidity and mortality worldwide but also poses significant challenges in diagnosis and management. The role of imaging in the assessment of myocardial infarction is pivotal, offering crucial insights into diagnosis, the extent of cardiac damage, and guiding therapeutic decisions (Thygesen et al., 2007).

Myocardial infarction is a major health concern globally, with its prevalence influenced by factors such as lifestyle, diet, and genetic predisposition. Risk factors for MI include hypertension, hyperlipidemia, smoking, diabetes, obesity, sedentary lifestyle, and a family history of coronary artery disease (Yusuf et al., 2004).

The pathogenesis of MI primarily involves the formation of atherosclerotic plaques in the coronary arteries, which can rupture and lead to thrombus formation. This

thrombus can occlude the artery, drastically reducing or completely stopping blood flow to a part of the heart muscle, resulting in ischemia and necrosis of the myocardial tissue.

Symptoms of MI can vary but often include chest pain or discomfort, which may radiate to the shoulders, arms, back, neck, or jaw. Other symptoms might include shortness of breath, nausea, vomiting, light-headedness, and cold sweats. However, not all MIs present with classic symptoms, making imaging and diagnostic tests critical for accurate diagnosis.

Imaging plays a fundamental role in the diagnosis, management, and prognostication of myocardial infarction. Various imaging modalities are utilized, each offering unique benefits.

CMR provides detailed images of the heart's structure and function, including the extent of myocardial damage, edema, and areas of microvascular obstruction. It is particularly useful in assessing myocardial viability and predicting recovery after MI. Also Cardiac CT angiography is a non-invasive alternative to coronary angiography for visualizing coronary artery disease. It can also assess for complications of MI, such as ventricular aneurysms. Technologies such as radiomics strive to improve the visualization of scar tissue and detect subtle changes in MI-affected regions without relying on contrast agents. These tools aim to capture nuances that may go unnoticed by conventional imaging techniques, potentially offering a more comprehensive understanding of stroke progression and its effects on brain tissue.

Imaging is indispensable in the comprehensive assessment of myocardial infarction, from initial diagnosis to detailed evaluation of cardiac anatomy, function, and post-infarction complications. (Baeßler, Mannil, Maintz, et al., 2018; Larroza et al., 2018) The choice of imaging modality depends on the clinical scenario, available resources, and specific information required. As imaging technologies advance, their role in enhancing the accuracy of MI diagnosis, guiding therapeutic interventions, and improving patient outcomes continues to evolve, underscoring the importance of these techniques in contemporary cardiac care.

Heart failure

Heart failure is a clinical syndrome characterized by the heart's inability to pump sufficient blood to meet the body's metabolic demands (Ekundayo et al., 2013). This condition can result from any structural or functional cardiac disorder that impairs the ventricles' capacity to fill with or eject blood. Heart failure is a significant global health issue, affecting millions of people worldwide, and is associated with high morbidity and mortality rates.

The pathophysiology of heart failure involves a complex interplay of hemodynamic, neurohormonal, and molecular changes. Initially, the heart tries to compensate for reduced pumping capacity through mechanisms such as ventricular dilation (to increase stroke volume), hypertrophy (to augment contractile force), and activation of the sympathetic nervous system and the renin-angiotensin-aldosterone system (to maintain arterial pressure and salt-water balance). However, over time, these compensatory mechanisms become maladaptive, leading to worsening heart failure (Geelhoed et al., 2020; Sahle et al., 2017).

Heart failure can be caused by a wide range of conditions that damage the heart muscle, including coronary artery disease, hypertension, valvular heart disease, cardiomyopathies, and myocarditis. Risk factors for developing heart failure include advanced age, diabetes, obesity, smoking, and a family history of cardiovascular disease.

Heart failure is classified based on the left ventricle's ejection fraction (EF), which measures the percentage of blood leaving the heart each time it contracts:

- Heart failure with reduced ejection fraction (HFrEF): EF less than 40%.
- Heart failure with reduced ejection fraction (HFrEF): EF less than 40%.
- Heart failure with preserved ejection fraction (HFpEF): EF 50% or higher, with symptoms of heart failure.
- Heart failure with mid-range ejection fraction (HFmrEF): EF between 40% and 49%.

Symptoms of heart failure may vary depending on the type and severity of the condition but commonly include dyspnea (shortness of breath), orthopnea (difficulty breathing while lying flat), paroxysmal nocturnal dyspnea, fatigue, reduced exercise capacity, and peripheral edema. Physical examination may reveal signs such as pulmonary rales, jugular venous distension, and a third heart sound (S3).

The diagnosis of heart failure is based on medical history, physical examination, and diagnostic tests, including electrocardiography, chest X-ray, and laboratory tests. Echocardiography is essential for assessing ventricular function, structure, and the presence of underlying heart disease. Additional tests such as CMR, stress testing, and cardiac catheterization may be used to identify the etiology of heart failure and guide treatment.

Management of heart failure aims to relieve symptoms, improve quality of life, and reduce hospitalization and mortality. Treatment includes lifestyle modifications (e.g., diet, exercise), pharmacological therapy (e.g., ACE inhibitors, beta-blockers, diuretics, aldosterone antagonists), and device therapy (e.g., implantable cardioverter-defibrillators, cardiac resynchronization therapy). In advanced cases, surgical options such as valve repair or replacement and heart transplantation may be considered.

Dilated cardiomyopathy

Dilated cardiomyopathy (DCM) is a condition characterized by the dilation and impaired contraction of the left or both ventricles of the heart (Schultheiss et al., 2019). It is a leading cause of heart failure and may result in arrhythmias, embolic events, or sudden cardiac death. Unlike hypertrophic cardiomyopathy, which involves thickening of the heart muscle, DCM involves thinning and weakening of the ventricular walls, leading to ineffective blood pumping (Weintraub, Semsarian, and Macdonald, 2017a).

DCM affects individuals of all ages, from infants to the elderly, with a higher prevalence in middle-aged adults. It can arise due to a variety of causes, including genetic mutations, viral infections of the heart, exposure to toxins (including alcohol and certain drugs), and autoimmune diseases (Hershberger, Hedges, and Morales, 2013). In many cases, however, the exact cause remains idiopathic, meaning it is unknown.

The fundamental pathology in DCM is a progressive dilation of the ventricles with concurrent systolic dysfunction. This dilation impairs the heart's ability to pump blood efficiently, leading to compensatory mechanisms such as ventricular hypertrophy and increased sympathetic nervous system activity. Over time, these compensations become maladaptive, exacerbating heart failure symptoms.

Symptoms of DCM often start insidiously and can include fatigue, weakness, shortness of breath, and edema. As the condition progresses, it can lead to significant morbidity from heart failure, arrhythmic events, and thromboembolic complications due to stasis of blood in the dilated ventricles (Schultheiss et al., 2019).

The diagnosis of DCM involves a thorough clinical evaluation, including a detailed medical history, physical examination, electrocardiogram (ECG), and chest X-ray. However, echocardiography is the cornerstone of diagnosis, providing detailed information on ventricular dimensions, systolic function, and the presence of any structural heart disease.

CMR plays a pivotal role in the assessment of DCM. It offers superior spatial resolution and tissue characterization, enabling precise measurement of ventricular volumes, wall thickness, and systolic function. CMR can also identify myocardial fibrosis, which is a predictor of adverse outcomes in DCM. This ability to visualize structural changes at a microscopic level and to differentiate DCM from other cardiomyopathies makes CMR an invaluable tool in the management and prognostication of patients with DCM (Weintraub, Semsarian, and Macdonald, 2017b).

Management of DCM is primarily aimed at treating heart failure symptoms and preventing complications such as arrhythmias and thromboembolic events. This may include the use of medications such as ACE inhibitors, beta-blockers, and diuretics, along with lifestyle modifications. In advanced cases, device therapy (e.g., implantable cardioverter-defibrillators or cardiac resynchronization therapy) or heart transplantation may be considered.

Dilated cardiomyopathy is a complex condition that significantly impacts patients' quality of life and survival. Advances in genetic research and imaging techniques, particularly the use of CMR, have greatly enhanced our understanding and management of DCM. Ongoing research into the underlying mechanisms and potential targeted tools such radiomics continues to offer hope for improved outcomes in this challenging condition.

Hypertrophic cardiomyopathy

Hypertrophic cardiomyopathy (HCM) is a complex cardiovascular disorder characterized by the abnormal thickening of the heart muscle, particularly the ventricular septum. This thickening can impede normal blood flow out of the heart and affect the heart's ability to pump effectively. HCM is a significant cause of sudden cardiac death, especially in young athletes, and can lead to a variety of symptoms ranging from benign to life-threatening (Maron, 1997).

HCM is among the most common genetic heart disorders (Gruner et al., 2013), affecting approximately 1 in 500 individuals globally. It can affect both men and women of any age. HCM is primarily inherited in an autosomal dominant pattern, meaning a mutation in just one copy of the responsible gene can cause the disorder. Mutations in several genes related to cardiac muscle proteins are known to contribute to the development of HCM, including the MYH7 and MYBPC3 genes.

The hallmark of HCM is myocardial hypertrophy that is not solely explained by abnormal loading conditions. This hypertrophy is most often asymmetric, with the septum between the ventricles being more commonly and severely affected than the ventricular free walls. The cellular architecture of the heart muscle in HCM is disorganized, a condition known as myocardial disarray, which contributes to the risk of arrhythmias.

This abnormal thickening and structural disarray can lead to several complications, including obstruction of blood flow from the left ventricle (known as obstructive HCM), diastolic dysfunction (difficulty with ventricular filling), mitral valve abnormalities, and an increased risk of atrial and ventricular arrhythmias.

The presentation of HCM is highly variable. Some individuals remain asymptomatic throughout their lives, while others develop symptoms such as shortness of breath, chest pain, palpitations, or episodes of lightheadedness and fainting. The risk of sudden cardiac death, although low overall, is a concern, particularly in younger individuals engaged in competitive sports (Stewart, Lavie, Shah, et al., 2018).

Diagnosis of HCM involves a comprehensive evaluation, including the patient's medical history, family history, physical examination, and several diagnostic tests. Echocardiography (ultrasound imaging of the heart) is the primary tool for diagnosing HCM, allowing visualization of the heart's structure and function, including the measurement of myocardial thickness and assessment of blood flow (Cardim et al., 2015).

Electrocardiography (ECG) can show abnormal heart rhythms and other changes indicative of HCM. Genetic testing may also be offered to identify mutations in genes associated with HCM and to guide family screening and management.

On the other hand, imaging is also important in the diagnosis of HCM via CMR. Clinicians often uses specific criteria, such as:

1. Myocardial Thickness: Wall thickness of 15 mm or more in adults (or significantly above the expected range for age and body size) without another cardiac or systemic cause.

2. Pattern of Hypertrophy: Identification of the pattern of hypertrophy (e.g., septal, apical) can aid in diagnosis and management.

3. Tissue Characterization: Presence of myocardial fibrosis, especially in areas of hypertrophy, may support the diagnosis and has prognostic implications. The thesis specifically focuses in this criteria section.

Management strategies for HCM are tailored to the individual's symptoms, risk factors for sudden cardiac death, and whether there is outflow tract obstruction. Treatment may include lifestyle modifications, medications (such as beta-blockers or calcium channel blockers to manage symptoms and improve heart function), and, in some cases, invasive procedures. Septal myectomy or alcohol septal ablation can be considered for patients with severe symptoms due to left ventricular outflow tract obstruction. Implantable cardioverter-defibrillators (ICDs) may be recommended for those at high risk of sudden cardiac death.

Hypertrophic cardiomyopathy is a genetically diverse and clinically variable condition that requires careful management to mitigate its complications, including sudden cardiac death. Advances in genetic understanding and imaging techniques have significantly improved the diagnosis and management of HCM, enhancing the quality of life and prognosis for those affected. Despite the advancements in cardiac imaging, HCM is still difficult to differentiate from other specific conditions, as we will see in the future chapters.

Left Ventricle Non-Compaction (LVNC)

Left Ventricular Non-compaction (LVNC) is an uncommon cardiac condition, distinguished by a unique, sponge-like morphology of the left ventricle - one of the heart's

primary chambers responsible for pumping blood. This anomalous presentation is attributed to the failure of the normal myocardial compaction process during fetal development, leading to this distinctive structural appearance (Petersen et al., 2005; Petersen, Matthews, Francis, et al., 2016).

LVNC can compromise the heart's ability to efficiently pump blood, leading to a range of symptoms that may include dyspnea, chest discomfort, and arrhythmias. The condition's etiology can be either genetic, stemming from inherited factors, or acquired, resulting from specific pathologies or environmental influences. It is also noteworthy that LVNC frequently co-occurs with other cardiac conditions, such as various forms of cardiomyopathy.

Management of LVNC generally comprises a multifaceted approach. Medications are often prescribed to alleviate symptoms and improve heart function. Lifestyle modifications may be recommended to mitigate the exacerbation of symptoms and prevent further cardiac complications. In more severe cases, or when concurrent cardiac conditions are present, surgical interventions such as ventricular assist devices or transplantation may be considered.

The most challenging aspects in managing LVNC lie in (i) distinguishing it from other cardiomyopathies, and (ii) predicting its prognosis, encompassing both its evolution and the prediction of those likely to experience cardiac events. The complexity and heterogeneity of LVNC present substantial difficulties in accurately identifying it and predicting its course. Certain patients may remain asymptomatic and stable over time, while others may exhibit a progressive course leading to heart failure, arrhythmias, or thromboembolic events. This variability in clinical progression makes prognostication a particularly arduous task. The ability to accurately differentiate it from other complex CVDs and prognosticate adverse cardiac events is crucial in managing LVNC as it enables timely intervention, optimized treatment strategies, and can significantly improve patient outcomes.

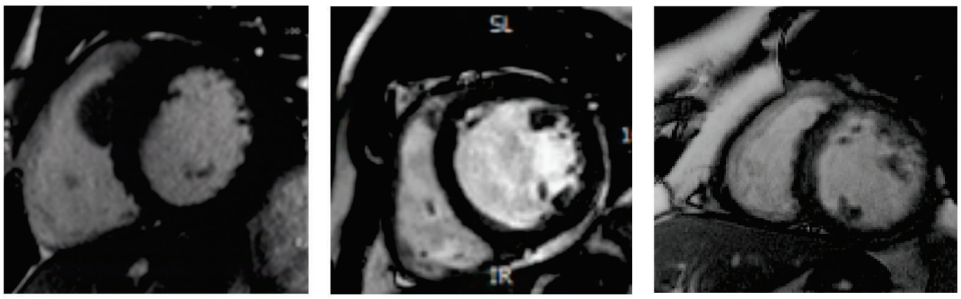

FIGURE 2.4: ExampleS of CMRs depicting three complex cardiac diseases challenging to distinguish: dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), and left ventricular non-compaction (LVNC). The leftmost image showcases DCM, which is characterized by a noticeable thickening of the heart muscle. The middle image depicts HCM, where the myocardium appears thinner in comparison. Lastly, the right image exemplifies LVNC, identifiable by its unique sponge-like structure. The distinctive trabeculae in this case are found in the region of the left ventricle.

2.2 Methodological context

2.2.1 Clinical measurements

Cardiologists utilize CMR (Pennell, 2004) as a sophisticated tool to extract vital clinical indices critical for predicting, diagnosing, planning treatment, and managing cardiovascular diseases. The unparalleled detail provided by MRI in visualizing cardiac anatomy and function is due to its superior spatial resolution and the ability to differentiate between various tissue types without exposing patients to ionizing radiation. This imaging technique is instrumental in quantifying a broad array of cardiac metrics essential for evaluating the health and functionality of the heart.

The procedure initiates with the capture of precise images using specialized MRI sequences designed specifically for cardiac assessment. These include cine MRI for motion assessment, myocardial tagging for evaluating strain, and late gadolinium enhancement to identify fibrosis or scarring. Each sequence furnishes distinct and valuable data contributing to a thorough cardiac examination.

Following image capture, sophisticated software aids cardiologists and radiologists in analyzing the images. This phase entails meticulous measurement of structural and functional cardiac indices. Among the primary metrics measured are the volumes of the ventricles, the mass of the myocardium, the fraction of blood ejected from the heart during each beat (ejection fraction), and the degree of myocardial strain. These measurements are derived through either manual, semi-automated, or automated processes of image segmentation, where the contours of the heart's chambers and the myocardium are precisely delineated.

One of the most important measures utilized by cardiologist is Left Ventricle Ejection Fraction (LVEF)(Maceira et al., 2006). Assessing the left ventricle ejection fraction (LVEF) is a fundamental procedure in cardiology that involves calculating the percentage of blood the left ventricle ejects with each heartbeat relative to its total volume during diastole. This metric is crucial for gauging heart function, particularly in diagnosing and managing heart failure and other cardiac conditions. The measurement process typically employs imaging modalities such as echocardiography, CMR, or ventriculography.

Echocardiography, often the first choice due to its accessibility and non-invasive nature, uses ultrasound waves to create images of the heart in motion, enabling the visualization of the left ventricle as it contracts and relaxes. CMR, known for its high spatial resolution and accuracy, provides detailed images of the heart's structure and function, allowing for precise calculation of LVEF. In some cases, ventriculography, an invasive technique using contrast dyes and X-rays, is utilized, especially when other methods are inconclusive (Grothues et al., 2002).

To measure LVEF, healthcare professionals first determine the volumes of the left ventricle at the end of diastole (when it is fullest) and at the end of systole (when it is least full). The LVEF is then calculated by subtracting the systolic volume from the diastolic volume, dividing this difference by the diastolic volume, and multiplying by 100 to express the result as a percentage.

LVEF is a vital measure because it directly reflects the heart's ability to pump blood effectively throughout the body (Maceira et al., 2006). A normal LVEF ranges from 55% to 70%, indicating efficient heart function. Values below this range suggest systolic heart failure or cardiomyopathy, where the heart's pumping ability is compromised. Conversely, an LVEF higher than normal can indicate diastolic dysfunction, where the heart is stiff and cannot fill properly. Monitoring LVEF over time can help guide treatment decisions, including medication adjustments, the need for

implantable devices, or surgical interventions, making it an indispensable tool in the management of patients with heart disease.

Trabeculations in the heart are mesh-like, irregular muscular columns that line the inner surface of the ventricles, primarily the left ventricle. These structures are formed during normal embryological development of the heart and consist of networks of muscle fibers (see Figure 2.4. They play a role in the functioning of the heart by helping to manage blood flow and minimize stress on the ventricular walls during the contraction cycle. In certain conditions, such as left ventricular non-compaction cardiomyopathy, the trabeculations are excessively prominent, which can be detected through imaging techniques like CMR and may affect heart function. In evaluating ventricular structures, especially trabeculations, the use of the Petersen and Jacquier coefficients is crucial. These metrics, derived from advanced CMR techniques, allow for a precise differentiation between trabeculated and compact myocardial masses. The Petersen coefficient (Petersen et al., 2005) provides a ratio of trabeculated to total myocardial mass, helping to assess the extent of trabeculation. Conversely, the Jacquier coefficient (Jacquier et al., 2010) evaluates the trabecular mass relative to the total ventricular mass. These coefficients are instrumental for diagnosing conditions like left ventricular non-compaction cardiomyopathy (LVNC), where excessive trabeculations are characteristic.

Identifying the degree of trabeculation involves high-resolution imaging and detailed segmentation techniques to distinguish pathological from normal physiological trabeculations. This differentiation is essential due to the variability in trabeculae among individuals and the complexity of the heart's structures. The precise measurement facilitated by these coefficients aids in the standardization of diagnoses and helps to ensure accuracy in identifying pathological conditions.

Additionally, late gadolinium enhancement (LGE) is a technique utilized in CMR to identify scar tissue or MI within the heart muscle (Kim et al., 2009). This method involves the use of a gadolinium-based contrast medium, which highlights differences in the tissue composition of the heart muscle by accumulating in regions with increased extracellular space, often a hallmark of fibrosis or scarring.

In the process of LGE, the contrast agent makes the affected myocardial areas appear bright on MRI images, contrasting sharply with the darker appearance of healthy myocardial tissue (see Figure. 2.5). This distinction is particularly useful for delineating the presence and extent of scar tissue, offering critical insights into the condition of the heart following injury or disease.

LGE MRI has become an indispensable tool in the evaluation of cardiac health, further used in the final chapter of this thesis, especially for detecting and quantifying myocardial damage due to ischemic heart disease or other cardiomyopathies (Stirrat and White, 2013). It plays a pivotal role in assessing myocardial viability, informing treatment strategies, and predicting patient outcomes, such as susceptibility to arrhythmias or the potential for heart failure progression.

Accuracy in these measurements is paramount, as inaccuracies can lead to misdiagnosis or inappropriate management strategies. Adherence to standardized measurement protocols is crucial for ensuring the reliability and consistency of these indices.

2.2.2 Challenges in cardiovascular imaging

Imaging techniques are in continuous development but still face many challenges (Fujikura, 2022). One of the primary difficulties arises from the inherent limitations

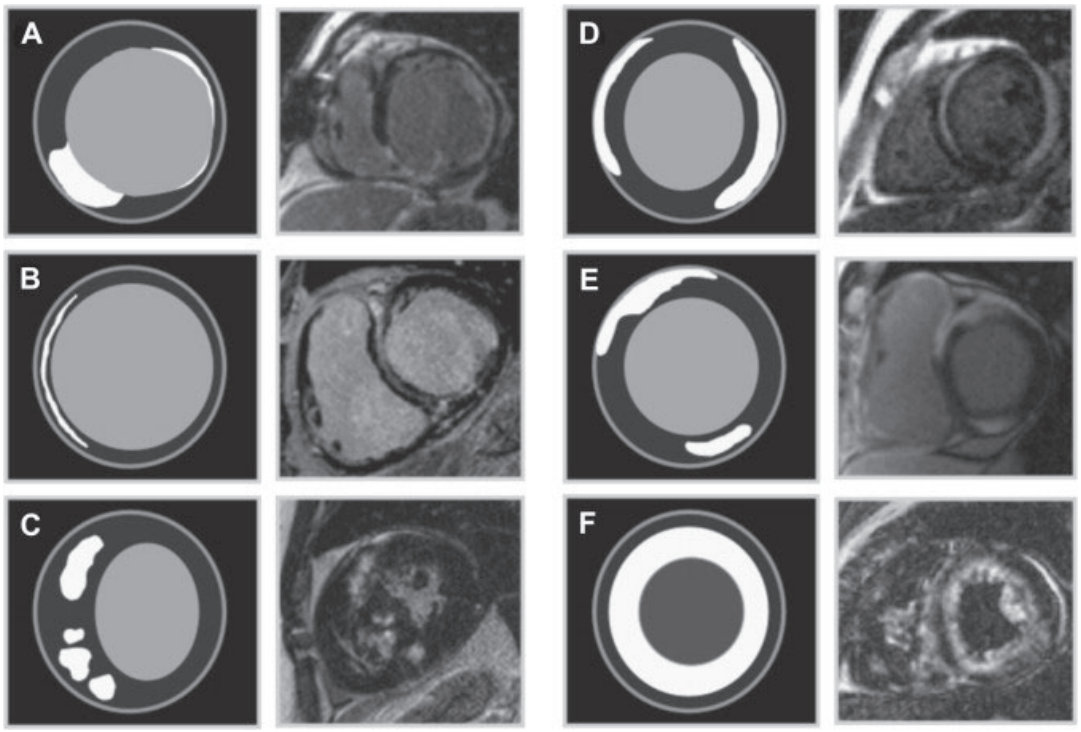

FIGURE 2.5: Schematic of typical hyperenhancement (HE) patterns identified by late gadolinium enhancement imaging in patients with ischemic and nonischemic cardiomyopathy. Source of the image: Stirrat J, White JA. The prognostic role of late gadolinium enhancement magnetic resonance imaging in patients with cardiomyopathy. Can J Cardiol. 2013 Mar;29(3):329-36. doi: 10.1016/j.cjca.2012.11.033. PMID: 23439019.(Stirrat and White, 2013)

of imaging modalities. While technologies like CMR and echocardiography are invaluable for visualizing the heart's structure, their ability to delineate fine trabeculations can be influenced by factors such as image resolution, the specific imaging sequence or protocol used, and the observer's experience. CMR, despite its superior spatial resolution, requires careful optimization of sequences to adequately capture the contrast between trabeculated and compact myocardium. Similarly, echocardiography's effectiveness can be hampered by acoustic shadowing and the dependency on the patient's anatomy and window of imaging (Dwivedi et al., 2013).

Moreover, the criteria for diagnosing conditions related to trabeculation, such as LVNC, are subject to ongoing debate and refinement. The lack of universally accepted standards for measuring and interpreting trabeculations adds another layer of complexity. Researchers and clinicians employ various quantitative indices, like the trabeculation-to-compact layer ratio, but these metrics can vary significantly based on the methodology and thresholds used (Casas, Rodríguez-Palomares, and Ferreira-González, 2022).

Inter-individual variability further complicates the assessment of left ventricle trabeculations. There is a wide range of normal trabeculation among healthy individuals, influenced by factors such as age, gender, and ethnicity. This variability necessitates a tailored approach to interpreting trabeculation levels, considering the patient's unique background and the clinical context.

The dynamic nature of the heart, with trabeculations that may alter in appearance and function over time or in response to physiological and pathological conditions, requires a dynamic and flexible assessment strategy. This necessitates ongoing research and development of more refined imaging techniques and analytical tools, as well as a multidisciplinary approach involving cardiologists, radiologists, and imaging specialists to accurately assess and interpret the significance of left ventricle trabeculations.

Regarding the use of late gadolinium enhancement, while is a powerful imaging technique in CMR for identifying myocardial scarring and fibrosis, it does have several drawbacks:

- Nephrogenic Systemic Fibrosis Risk: Gadolinium-based contrast agents can pose a risk of nephrogenic systemic fibrosis (NSF) in patients with severe renal impairment. NSF is a rare but serious condition characterized by fibrosis of the skin, joints, and internal organs.

- Allergic Reactions: Although rare, some patients may experience allergic reactions to the gadolinium contrast agent, ranging from mild symptoms to more severe anaphylactic reactions.

- Limited Accessibility and Cost: LGE MRI requires specialized equipment and expertise, which may not be available in all medical facilities. Additionally, the cost of the procedure and the contrast agent can be high, limiting accessibility for some patients.

- Contraindications: Patients with certain types of metal implants, pacemakers, or defibrillators may not be eligible for MRI procedures, including those involving gadolinium contrast, due to safety concerns.

- Temporal Resolution: While LGE provides excellent spatial resolution of myocardial scarring, it may not offer the best temporal resolution for dynamic cardiac function assessments compared to other imaging modalities.

- Quantification Challenges: Although LGE is excellent for qualitative assessment of scar tissue, quantifying the exact amount of myocardial fibrosis can be challenging and may require advanced imaging and analysis techniques.
- Kidney Function Monitoring: Patients undergoing LGE MRI may require monitoring of their kidney function before and after the administration of gadolinium, adding an additional step to the imaging process.

Extracting clinical indices from CMR images represents a sophisticated and essential task that integrates advanced imaging technology with medical proficiency. However, this process is not without its challenges (Pennell, 2004). Our research aims to address these obstacles by exploring how emerging technologies can be leveraged to surmount these difficulties, such automatize the extraction process, standarize the protocols, recognize and visualize unseen new patterns or remove the use of contrast agents.

2.2.3 Radiomics analysis

Recent advancements in MRI technology and image processing, including the integration of machine learning and artificial intelligence, are refining the precision with which these clinical indices are measured (Najjar, 2023; Pinto-Coelho, 2023). These innovations are streamlining the measurement process, reducing variability between observers, and paving the way for more accurate, reliable, and predictive cardiac assessments. In this scenario is where radiomics play an important role. Radiomics is a technique that involves using mathematical and geometrical algorithms to extract large amounts of data from images (Gillies, Kinahan, and Hricak, 2016). This approach is used to obtain numerical values from complex image characteristics, which can be difficult to be interpreted directly. Radiomics can provide both geometric or morphological information and more complex data information such as variations in gray textures or histogram-based characteristics.

In the context of medical imaging, the application of radiomics has demonstrated its reliability and effectiveness in various domains, mainly used in oncology for over a decade (Ding et al., 2021). Radiomics has proven to be effective in capturing essential textural and shape details that play a crucial role in comprehending various types of cancer, as well as predicting prognosis and assessing risk.(Wu et al., 2018). In recent times, researchers have put forth the suggestion that the unique capabilities of radiomics make it potentially applicable in the field of cardiology (Raisi-Estabragh et al., 2020a). Numerous publications have highlighted the efficacy of radiomics in predicting and classifying cardiac diseases (Cetin et al., 2018; Cetin et al., 2020; Rauseo et al., 2021). While the medical community has made significant efforts to establish standardized protocols and define quantitative metrics for different diseases, there remain certain procedures that still rely on subjective evaluation by clinicians. To address this challenge, radiomics aims to provide a quantitative approach by extracting numerical features from medical images. These features can then be used to establish value ranges and evaluation scales specific to each disease, facilitating easier extrapolation of findings to other clinical studies and promoting the adoption of a unified criterion for disease prognosis and diagnosis. To perform a radiomics analysis, it is essential to follow a series of crucial steps, depicted below:

Region of interest (ROI) segmentation

Medical images are reservoirs of substantial pixel-based information, where only a fraction of the image is the target, while surrounding organs or tissues may not be relevant for the study. In radiomics analysis, it is critical to accurately delineate the organ - the area of interest within the overall image. In this work, we concentrate exclusively on the heart, emphasizing the importance of precise segmentation. Failure to accurately segment the ROI can lead to data extraction from extraneous regions outside the heart (i.e. lungs), diluting the focus of our analysis.

This dissertation focused on CMR of SAX view, as it provides extensive morphological and functional information regarding the studied cardiac disease. Within this perspective, the key structures requiring detailed delineation include the left ventricle (LV), the right ventricle (RV), and the myocardium (MYO)(Fig.2.7).

In order to aid the healthcare professionals in this intricate process of segmentation, numerous software tools are available. These range from freely accessible platforms such as ITK-Snap (Fig.2.6) to proprietary software like CVi42 *CVI42 Version x.x* Year. By leveraging these digital tools, clinicians can accomplish accurate segmentation of the heart's intricate structures, thereby enabling an efficient and precise radiomics analysis.

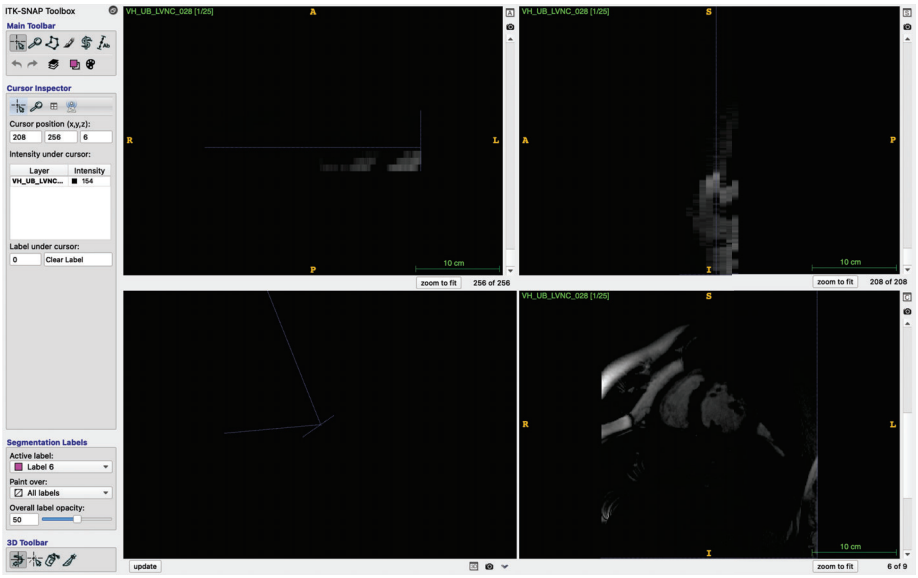

FIGURE 2.6: ITK-Snap is a free open source tool that allows clinicians/researchers to delineate the contours of the ROI.

Radiomics extraction

The researchers in Griethuysen et al., 2017a have introduced PyRadiomics, an open-source software code that has gained significant popularity in the field. PyRadiomics facilitates the calculation and extraction of radiomics features. It offers a user-friendly object-oriented code, requiring only the original image and the corresponding mask or contours to enable efficient extraction of radiomics. The extracted radiomics are conveniently presented in a tabular data format, allowing for ease of analysis and interpretation.

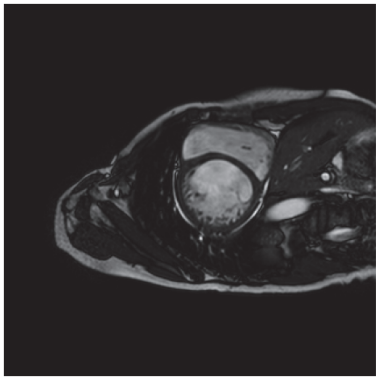

FIGURE 2.7:
Original image

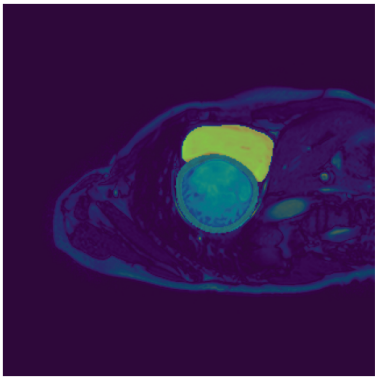

FIGURE 2.8:
ROI selection

FIGURE 2.9: Process of delineation. Clinicians must select the ROI before radiomics extraction

In our particular scenario, we acknowledge the unique characteristics of each region of the heart and recognize the significance of extracting and analyzing radiomics separately based on the specific type of cardiomyopathy under investigation. This approach proves crucial in identifying structural alterations at a localized level and determining the most relevant radiomics for detecting these changes. For instance, in cases where the injury is confined to a particular region, such as a myocardial scar resulting from a myocardial infarction, it becomes imperative to extract radiomics individually from the left ventricle (LV), right ventricle (RV), and myocardium (MYO). This enables a comprehensive evaluation of the distinct radiomic features associated with each region, facilitating accurate detection and characterization of pathological conditions.

Radiomics can be broadly categorized into three distinct types based on the mathematical attributes they describe. These include Shape Radiomics, First-Order Radiomics, and Textural Radiomics.

Shape Radiomics characteristics assess and gauge the morphological attributes within delineated contours, disregarding the distribution of gray-level intensity. This category of features is intuitively comprehensible as they correspond closely to volumes and surfaces that are routinely measured with CMR indices in clinical practice. Examples of these features range from basic metrics like volume, elongation, or surface area, to more complex metrics such as sphericity.

First-order radiomics features are derived from the statistical properties of voxel intensities within a ROI, as defined by a segmentation mask. These statistics articulate the pixel/voxel intensity distribution in the form of fundamental metrics, independent of spatial interrelationships. Examples of first-order radiomics features include simpler metrics such as the median or mean, as well as more complex mathematical metrics like entropy, energy, or kurtosis.

In contrast, texture radiomics features identify nuanced alterations in the pixel's gray-scale distribution by detecting patterns and changes in adjacent gray-scale values through advanced matrix computations. These texture features can be classified into five categories based on the matrix used for their derivation: Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), Neighbouring Gray Tone Difference Matrix (NGTDM), and Gray Level Dependence Matrix (GLDM).

It is anticipated that shape features will encapsulate cardiac morphological traits typically associated with each specific cardiac condition. Meanwhile, first-Order and texture features are anticipated to significantly contribute to discerning gray-scale alterations within the Left Ventricle (LV) or Left Ventricle Myocardium (LVMYO) tissue. This is especially pertinent for distinguishing trabeculations in patients with LVNC, a characteristic attribute of this condition.

Challenges in radiomics implementation

The utilization of radiomics for cardiovascular imaging, while holding significant promise, encounters various obstacles and limitations that impact its efficacy and broader implementation (Jin, 2023; Najjar, 2023; "Radiomics in Cardiovascular Imaging" 2023). These challenges are spread across technical, clinical, and logistical areas, including:

- Variability and Lack of Standardization in Imaging Data: The diversity in imaging protocols, equipment from various manufacturers, and settings used across different healthcare facilities introduces inconsistency in radiomic feature extraction. Achieving uniformity in imaging protocols and feature extraction techniques remains a significant challenge.
- Complex Data Sets and the Risk of Overfitting: The extraction of numerous features from cardiovascular images results in complex datasets. Addressing these with advanced statistical and machine learning methods is crucial to prevent overfitting and ensure the model's applicability to novel data.
- Issues with Reproducibility and Independent Validation: The reproducibility of radiomic findings can be hampered by differences in image acquisition techniques, feature extraction algorithms, and analysis procedures. Conducting independent validation to verify these studies is essential yet often complicated by the need for extensive, annotated data collections (Raisi-Estabragh et al., 2020c).
- Incorporation into Clinical Practice: The integration of radiomics into existing clinical workflows poses both logistical and technical hurdles. It requires training for healthcare professionals to interpret radiomic assessments and seamless integration with current medical IT infrastructures.
- Model Interpretability: The complexity and lack of transparency in many high-performing machine learning models challenge their acceptance by clinicians. This opacity can hinder trust and limit the practical use of these models in clinical environments.

In the field of cardiovascular imaging, the effectiveness of radiomics is particularly contingent upon the precision of image segmentation techniques. These procedures are critical for delineating the relevant anatomical structures within the heart, ensuring that the subsequent radiomic analysis accurately reflects the underlying cardiac pathology. The accuracy of these segmentation steps directly impacts the quality and reliability of the extracted radiomic features, underscoring their importance in the overall process of leveraging radiomics for diagnostic and prognostic purposes in cardiovascular medicine. In our research, accurately identifying and

outlining the presence of excessive trabeculations within the left ventricle or myocardium is of paramount importance. This process is crucial for assessing the structural abnormalities of the heart, particularly in conditions such as Left Ventricular Non-compaction (LVNC).

2.2.4 Machine Learning

Machine learning, a subset of artificial intelligence, encompasses the training of algorithms using data to make predictions or detect patterns. It has found extensive employment in diverse fields such as image and speech recognition, natural language processing, and predictive analytics (Russell and Norvig, 2016).

In machine learning, algorithms undergo training using sizable datasets, enabling them to generalize and make accurate predictions or identify patterns in new, unseen data. This process involves supervised learning, unsupervised learning, or reinforcement learning methods.

Supervised learning employs labeled datasets where each example is associated with a known output, allowing the algorithm to learn from the provided ground truth (Russell and Norvig, 2016). Conversely, unsupervised learning operates on unlabeled datasets, requiring algorithms to autonomously uncover patterns and structure within the data (Bishop, 2006). Reinforcement learning involves an agent interacting with an environment, receiving feedback in the form of rewards or penalties as it learns to achieve a specific goal (Sutton and Barto, 2018).

Within the field of cardiology, machine learning has proven its utility in various applications. For instance, it has been used in image analysis for cardiac image segmentation, aiding in the precise delineation of anatomical structures (Litjens et al., 2017). Moreover, machine learning techniques have shown promise in the detection and classification of cardiac arrhythmias, contributing to more accurate diagnoses and personalized treatment strategies (Attia et al., 2019a).

By harnessing the power of machine learning, advancements in cardiology are facilitated, leading to improved patient outcomes, refined diagnostics, and enhanced decision-making processes.

Throughout the development of this thesis, various machine learning algorithms were employed for classification, ranging from tree-based decision models to more sophisticated methods such as SVM and XGBoost.

Classification

Support Vector Machine (SVM) is a powerful supervised machine learning algorithm predominantly used for classification tasks, although it can be employed for regression as well. Originally introduced by Vapnik and Chervonenkis in the 1960s, its modern and more popular incarnation was developed in the 1990s, when it was formulated in terms of an optimization problem that aimed to maximize the margin between classes(Cortes and Vapnik, 1995).

The core idea behind SVM is to find an optimal hyperplane that separates data into distinct classes. In two-dimensional space, this hyperplane is simply a line. However, in higher dimensions, it takes the form of a plane or a set of planes. The "optimal" hyperplane is the one that achieves the maximum margin from both classes, and the data points that lie closest to this hyperplane (and effectively define its position) are known as 'support vectors' – hence the name.

Another algorithm frequently used in ML is Random Forest (RF)(Breiman, 2001). A Random Forest Classifier is a machine learning model formed from a collection of

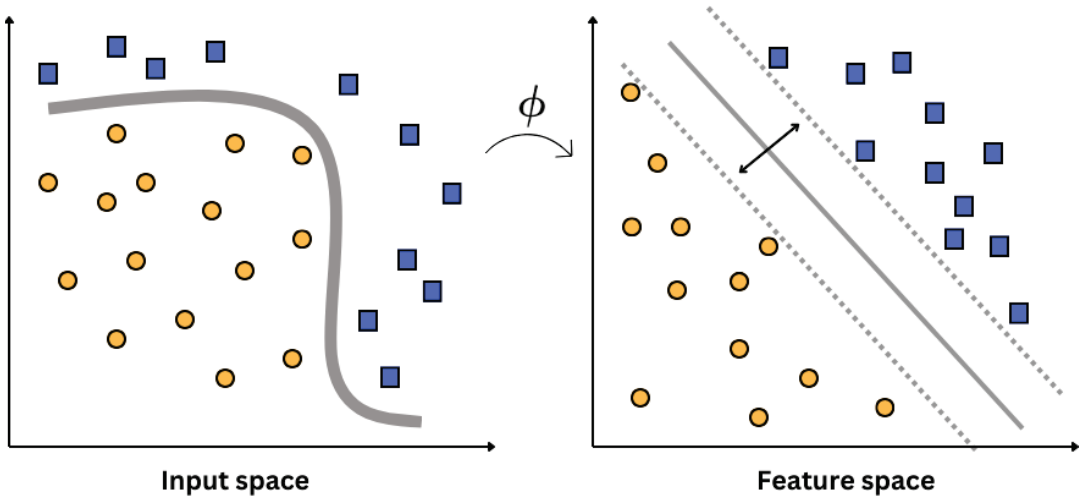

FIGURE 2.10: Support Vector Machine schematic.

multiple decision trees. To create diversity and reduce bias, each tree is trained on random subsets of data and features. When making predictions, every tree provides its own class 'vote', with the majority determining the final classification. This ensemble method not only increases prediction accuracy but also makes the model more resilient against overfitting. Furthermore, the Random Forest has the capability to rank feature importance, offering valuable insights into which variables most influence predictions.

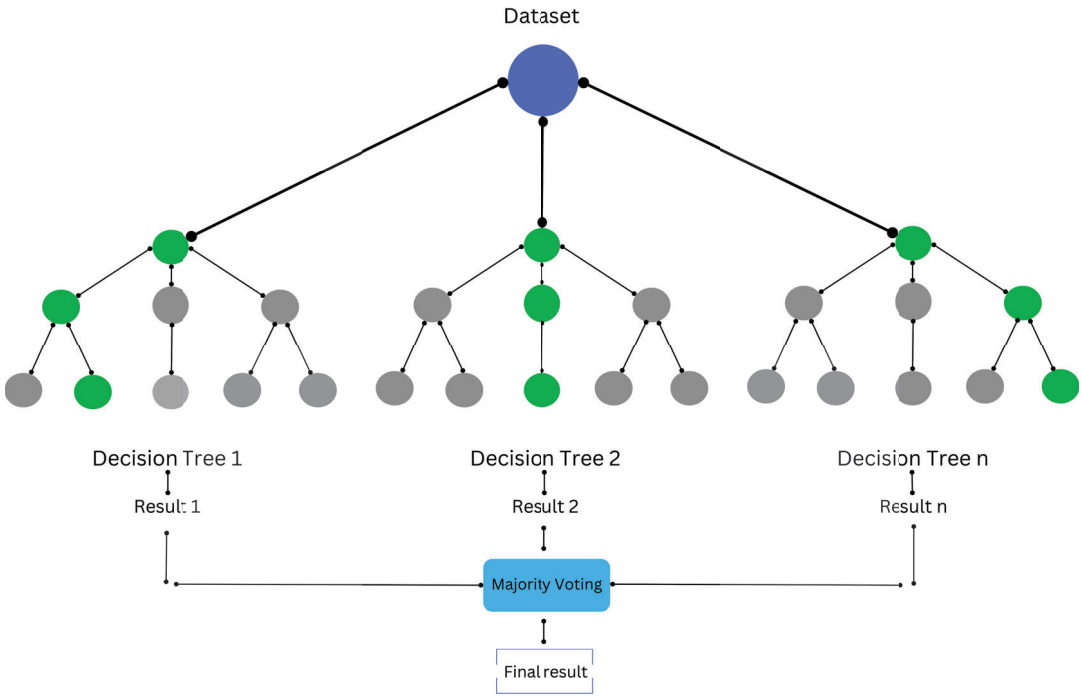

FIGURE 2.11: Decision tree schematic

XGBoost (eXtreme Gradient Boosting, Chen and Guestrin, 2016) is a powerful machine learning algorithm that has gained significant attention in the field of classification. It belongs to the family of gradient boosting algorithms, which are ensemble learning methods that combine the predictions of multiple weak models to create a strong predictive model.

In the context of classification, XGBoost excels at constructing predictive models that can accurately classify data into different classes or categories. It operates by sequentially building a series of decision trees, where each subsequent tree is trained to correct the mistakes made by the ensemble of previously trained trees. This iterative process allows XGBoost to progressively improve the model's performance and capture complex relationships within the data.

One of the key advantages of XGBoost is its ability to handle diverse types of data and effectively manage high-dimensional feature spaces. It employs a unique regularization technique known as "gradient-based regularization," which helps prevent overfitting and enhances the model's generalization capability. Additionally, XGBoost incorporates a robust optimization framework that optimizes a specific objective function, such as log-loss or cross-entropy, to find the best possible model parameters.

The algorithm incorporates several innovative features, including parallel tree construction, which accelerates the training process, and hardware optimization, which further enhances efficiency. XGBoost also supports various advanced techniques, such as handling missing values, handling imbalanced datasets, and providing feature importance analysis.

In summary, XGBoost is a state-of-the-art machine learning algorithm specifically designed for classification tasks. Its ability to handle diverse data types, manage high-dimensional feature spaces, and optimize objective functions makes it a valuable tool for developing accurate and robust classification models.

Survival analysis

Survival analysis (Flynn, 2012; Stel et al., 2011), also known as time-to-event analysis, is a statistical method used to examine the time until a specific event of interest occurs. In the context of our scenario, it pertains to analyzing the time until a cardiac event takes place.

In survival analysis, the event of interest is referred to as the "failure" event or "censoring" event, while the time until the event happens is called the "survival time." The data utilized in survival analysis typically comprises two types of observations:

- Failure observations: These include instances where the exact failure time is known.
- Censoring observations: These consist of cases where the failure time is not known, but it is established that the subject remained at risk for the event until a specific censoring time.

The primary objective of survival analysis is to estimate the underlying probability distribution of the survival time and draw inferences about the population based on the available sample data. One widely employed model in survival analysis is the Cox proportional hazards model (Abd ElHafeez et al., 2021; Deo, Deo, and Sundaram, 2021). This model estimates the hazard rate of the event as a function of one or more co-variates.

Survival analysis explores various aspects of the survival time distribution, such as the event of interest, probability of survival over time, hazard rate, cumulative hazard, and survival function.

Machine learning (ML) can be employed in conjunction with survival analysis to identify patterns and predictors related to survival time. In our healthcare context, an ML model could be trained using a dataset containing patient information

encompassing demographics, medical history, imaging features, ECG indexes, etc. This ML model aims to predict the likelihood of survival for individual patients. By analyzing intricate relationships and interactions between variables, ML algorithms can enhance the accuracy of survival predictions compared to traditional statistical techniques.

Traditional survival analysis methods, like the Cox proportional hazards model (Benítez-Parejo, Rodríguez del Águila, and Pérez-Vicente, 2011), make specific assumptions about data (e.g., the risk associated with variables is constant over time). However, real-world data can be more complex.

Machine learning offers a flexible approach. Neural networks, for instance, can be adapted to survival analysis. DeepSurv(Katzman et al., 2018), a deep learning approach to survival analysis, allows for the non-linear relationships between covariates and provides a risk score that can be used for personalized treatment recommendations. The reduced explainability of complex models, particularly in the area of deep learning, brings about significant disadvantages. These downsides not only impede the practical deployment of such models but also raise ethical concerns. Given these considerations, the aforementioned limitations in explainability are precisely why (Katzman et al., 2018) will be excluded from our evaluations in this study of survival analysis.

Random Survival Forests (Pölsterl, 2020), an adaptation of Random Forests for survival data, is another tool that captures complex interactions between variables and offers importance measures for each predictor.

2.2.5 Feature selection in modelling

Feature selection (FS) is a process in machine learning where the most relevant variables (features) are identified and selected for use in model construction (Xu et al., 2023). FS plays a pivotal role in machine learning (ML) within clinical settings, primarily due to its impact on the accuracy, efficiency, and interpretability of predictive models. In healthcare, where decisions can have direct consequences on patient outcomes, the selection of relevant features from clinical data is crucial for several reasons. Its use implies:

- Improves Model Performance: By choosing the most relevant features, ML models can focus on the most significant predictors of outcomes, potentially increasing accuracy and reducing the risk of overfitting. This is especially important in clinical setups where the prediction accuracy can directly affect patient care and treatment plans.
- Enhances Interpretability: In clinical practice, understanding why a model makes a certain prediction is as important as the prediction itself. Feature selection helps in simplifying models, making it easier for healthcare professionals to interpret the results and trust the ML-based decisions.
- Reduces Training Time: By eliminating redundant or irrelevant features, the dimensionality of the dataset is reduced, which can significantly decrease the computational resources and time required to train models. This is particularly beneficial in clinical environments where rapid decision-making is often needed.
- Facilitates Generalization: Selecting a robust set of features can help in developing models that generalize well to new, unseen data, which is critical when applying models across different patient populations or clinical settings.

- Aids in Data Understanding: The process of feature selection can reveal important insights about the underlying structure of clinical data and the relationships between different variables. This can lead to a better understanding of the disease processes and patient characteristics that are most relevant to health outcomes.
- Supports Cost-effective Testing: In scenarios where collecting certain data points is expensive or invasive, feature selection can identify the most informative features that need to be collected, potentially reducing the costs and burdens on patients.

In this thesis, we employ two widely recognized feature selection (FS) techniques, reflecting their prevalent use in the field.

- Chi-Square Test (McHugh, 2013): A filter method that assesses the statistical significance of the relationship between each feature and the target variable, particularly for categorical data. It measures how expectations compare to actual observed data. Features with higher values from this test indicate a stronger association with the target, making them prime candidates for inclusion in the model. This approach is effective for preliminary reduction of features, helping to streamline the modeling process.
- Sequential Forward Selection (SFS) (Ververidis and Kotropoulos, 2005): As a wrapper approach, SFS begins with no features and incrementally adds them based on their contribution to model performance. It evaluates each candidate feature by its ability to improve the model when combined with features already selected. The selection process continues until no significant performance gain is observed. Although SFS can be resource-intensive, it meticulously constructs a feature set that is optimized for the model at hand.

These methods, the Chi-Square test and Sequential Forward Selection, are instrumental in simplifying the feature space and enhancing model performance. The former is particularly useful for quick, initial screening of categorical features, while the latter offers a comprehensive strategy to build an effective feature combination through an iterative process. Leveraging these techniques allows for the development of more accurate and efficient predictive models.

2.2.6 Challenges in Machine Learning modelling

Several papers have highlighted the challenges associated with integrating machine learning (ML) into medical practice, particularly emphasizing the ethical considerations and practical hurdles in critical care and emergency medicine (Vayena, Blasimme, and Cohen, 2018; Sendak et al., 2019; Kang and Yoon, 2023). These challenges can be summarized below:

- Data Quality and Availability: One of the primary hurdles in ML is obtaining high-quality, relevant data. Models are only as good as the data they're trained on, and issues like missing values, inaccuracies, and biases in the data can significantly skew outcomes. In current clinical practice, a significant issue arises from incomplete data due to patients discontinuing or not adhering to follow-up processes. Additionally, there is inconsistency in the quality of data acquisition, and much of the data from the past five decades has been collected through manual annotations and paperwork, difficulting the translation to computational environments.

- Model Complexity and Interpretability: As ML models become more complex, interpreting their decision-making processes becomes more challenging. This "black box" nature can hinder trust and acceptance, especially in critical fields like healthcare.
- Computational Resources: Advanced ML models, particularly deep learning, require substantial computational power for training and inference, which can be costly and environmentally impactful.
- Generalization vs. Overfitting: Creating models that generalize well to new, unseen data while avoiding overfitting to the training dataset is a delicate balance. Overfitting leads to models that perform well on training data but poorly in real-world applications. In clinical practice, it's imperative for models to be versatile and applicable across various settings, including different centers and countries, ensuring they deliver reliable and generalizable outcomes.
- Ethical Concerns and Bias: ML models can inadvertently perpetuate or amplify biases present in their training data, leading to unfair or unethical outcomes. Ensuring models are fair and unbiased is a significant challenge.
- Dynamic Environments: In many applications, the environment in which the model operates can change over time, a phenomenon known as concept drift. Models need to be adaptable and capable of learning from new data without forgetting previous knowledge.
- Labeling Costs: For supervised learning, obtaining a large amount of accurately labeled data can be expensive and time-consuming, limiting the speed and scope of model development.
- Integration with Existing Systems: Integrating ML models into existing IT systems and workflows can be challenging, requiring significant effort to ensure compatibility and performance. Clinicians should be trained in their use and understanding for a better clinical outcome.

2.3 State-of-the-Art

Radiomics, originally emerging within the domains of oncology and radiology imaging (Shur et al., 2021), have recently seen a surge in their application to cardiovascular imaging. Given the voluminous influx of new research on this topic, several works have pivoted towards offering extensive reviews to encapsulate the evolving landscape (Polidori et al., 2023; Rizzo et al., 2018).

Radiomics have been applied to a wide range of cardiovascular applications. For example, in Kolossváry et al., 2017; Koskinas et al., 2015; Schlett et al., 2013 radiomics are used to evaluate and characterize cardiovascular issues related to coronary plaques. In Baessler et al., 2019, they delved into the differential potential of radiomic features extracted from cardiac magnetic resonance (CMR) imaging, aiming to distinguish between ischemic cardiomyopathy and dilated cardiomyopathy. Their findings indicated high diagnostic accuracy from machine learning models built upon radiomic signatures. Another publication, A et al., 2023 team focused on left ventricular remodeling in patients with dilated cardiomyopathy, utilizing radiomic features from echocardiography. Their findings illustrated that specific radiomic metrics could potentially predict adverse remodeling. In Pu et al., 2022, they

embark on a radiomic analysis to identify fibrosis within hypertrophic cardiomyopathy using cardiac magnetic resonance (CMR) cine imaging. Fibrosis, a critical pathological feature of hypertrophic cardiomyopathy, can play a pivotal role in prognosis and management. The research potentially delves into the capabilities of radiomics to enhance the precision of fibrosis detection through CMR, offering insights that could refine diagnostic and prognostic processes. In Neisius et al., 2019, they delve into the potential of radiomic analysis in distinguishing between hypertensive heart disease and hypertrophic cardiomyopathy using myocardial native T1 imaging. Both conditions, although distinct, might present overlapping clinical and imaging characteristics. This study possibly highlights the utility of radiomic features to offer enhanced diagnostic precision, aiding in the differentiation between these two heart conditions. Another example can be depicted in Neisius et al., 2020. In this work, the authors investigate the potential of texture signatures derived from native myocardial T1 imaging as innovative markers for identifying patients with hypertrophic cardiomyopathy, especially those without evident myocardial scarring. Hypertrophic cardiomyopathy, a prevalent heart condition, can often present without clear scar tissue, making diagnosis challenging. This study seeks to understand if specific texture signatures from native myocardial T1 imaging can offer enhanced diagnostic capabilities and help in the early and precise identification of such patients. Another similar example as the previous one can be found in Schofield et al., 2019. The primary objective of their study was to discern whether such texture analysis using radiomics can differentiate between various causes or aetiologies of left ventricular hypertrophy (LVH), a condition characterized by the thickening of the heart's left ventricular wall. Differentiating the underlying causes of LVH can be vital for appropriate treatment and prognosis. The study underscores the potential of advanced imaging techniques, such as texture analysis of CMR cine images, in enhancing diagnostic precision and providing more insights into the heterogeneity of LVH presentations. Throughout the various chapters outlined below, we will explore the current state-of-the-art and findings related to the specific research inquiries addressed in each of the chapters.

Chapter 3

Radiomics-Based Classification of Left Ventricular Non-compaction, Hypertrophic Cardiomyopathy, and Dilated Cardiomyopathy in Cardiovascular Magnetic Resonance

3.1 Introduction

Cardiomyopathies (CMs) are defined as primary myocardial disorders in the absence of other conditions that may affect the structural or functional properties of the heart's muscle (Elliott et al., 2007). CMs are divided into distinct morphologic phenotypes (Elliott et al., 2007), including hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM) as two of the most prevalent CMs. HCM is characterized by an increase in left ventricular (LV) wall thickness ("2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy" 2014), DCM by LV (or biventricular) systolic dysfunction and dilatation (Pinto et al., 2016), while both disorders are unexplained by loading conditions. Left ventricular noncompaction (LVNC) is a recently defined and poorly understood condition, characterized by prominent LV trabeculae, a thin compacted myocardial layer, and deep intertrabeculae recesses (Jenni, Oechslin, and Loo, 2007).

Cardiac magnetic resonance (CMR) is current the gold standard imaging modality for the clinical assessment of CMs, as well as to identify and differentiate the different phenotypes. CMR is widely used in the diagnosis of HCM (Cardim et al., 2015), DCM (Donal et al., 2019) and LVNC (Petersen et al., 2005; Jacquier et al., 2010; Captur et al., 2013). However, some LVNC features can overlap with those of other CMs and LVNC patients might present with morphological findings of HCM and/or DCM (Oechslin and Jenni, 2018). Furthermore, hypertrabeculation may also occur in the healthy population, which makes it challenging to differentiate physiologic from pathological hyper-trabeculation forms (Oechslin and Klaassen, 2019) by CMR. The difficulties to differentially and timely diagnose LVNC in clinical practice has motivated the development of new imaging indices, in particular, the Petersen (Petersen et al., 2005) and Jacquier (Jacquier et al., 2010) coefficients, which estimate the level of hypertrabeculation in the LV myocardium. However, these coefficients, while they improve LVNC diagnosis (Captur et al., 2013), are challenging and tedious to estimate in practice, as they require expert and accurate identification and

delineation of the trabeculae on the CMR images. This is a time-consuming task that is furthermore subject to inter-observer variability given the inherent complexity of the trabeculae.

Radiomics is an emerging image analysis technique for deeper phenotyping of cardiovascular health and disease in CMR (Raisi-Estabragh et al., 2020b). It enables the examination of a large pool of advanced imaging features that describe a wide range of complex, as well as subtle traits of the cardiac tissues at different scales and locations. Compared to existing cardiac indices such as those listed above, radiomics features encode multivariate information by capturing and combining heterogeneous morphological (e.g. sphericity, compactness) and appearance (e.g. entropy, coarseness) properties of the tissues. Hence, in the last years, several works have shown its potential for identifying new imaging signatures that can be leveraged for enhanced cardiac disease understanding ((Cetin et al., 2020; Amano et al., 2018)) and quantification (Aerts et al., 2014; Neisius et al., 2019; Cheng et al., 2018). In addition to providing comprehensive indicators of cardiac health and disease, CMR radiomics features are easier to calculate as they only require the segmentation of the myocardial boundaries, and even this segmentation process can be automatized (Campello et al., 2021).

This chapter presents the first to develop and evaluate a radiomics model for automatically differentiating LVNC, DCM, and HCM phenotypes in CMR. Based on a clinical dataset comprising different CM subgroups as well as healthy subjects from routine clinical practice, a machine learning pipeline is implemented to combine multiple radiomics features into a novel discriminative model of LVNC, HCM, and DCM. Subsequently, the obtained radiomics model is evaluated in great detail and its performance compared to the one obtained based on CMR indices, including the existing, manually estimated trabecular indices of LVNC. The results described in chapter show the promise of the proposed radiomics approach for achieving state-of-the-art LVNC, HCM, and DCM differential diagnosis more efficiently, while removing the need for the delineation of the LV trabeculae.

3.2 Data and Methodology

3.2.1 Dataset

The study cohort consists of 118 subjects, including 37 DCM, 25 HCM, and 35 LVNC patients, as well as 21 healthy control (NOR) subjects. HCM and DCM populations were available from the 2020 M&Ms MICCAI Challenge dataset (Campello et al., 2021). All patients for this study were assessed at the Hospital Universitari Vall d'Hebron (HUVH) following standard CMR protocols. Table 3.1 summarizes the clinical diagnostic criteria for each disease. In short, HCM, DCM, and LVNC diagnoses were established by expert cardiologists based on currently accepted imaging criteria (Cardim et al., 2015; Donal et al., 2019; Petersen et al., 2005; Jacquier et al., 2010) combined with other clinical data, such as electrocardiography, family history, and genetics. The mean age of the cohort was 49.4 $\pm$17.97 and 76 subjects (65% of the cohort) were men (see Table 3.2 for more detailed information).

3.2.2 CMR clinical indices

All patients underwent a standard CMR protocol. In brief, all scans were performed with a 1.5 Tesla scanner (Avanto, Siemens Healthcare, Erlangen, Germany), with typical cine parameters as follows: TR/TE (repetition time/echo time) =3.2/1.5 ms,

TABLE 3.1: Cohort size for each specific disease/control and clinical criteria for inclusion.

Disease	Cohort size	Clinical inclusion criteria
DCM	37	Depressed LVEF with increased LV volumes. Usually normal LV mass, wall thickness and asymmetry.
HCM	25	Increased LV mass with wall thickness $> 15mm$ and/or asymmetry > 1.3. Ussually preserved LVEF.
LVNC	35	Jacquier ratio $> 20\%$ and Petersen ratio > 2.3. LVEF, LV volumes, LV mass and wall thickness can be normal or not.
NOR	21	Normal conventional CMR values. No history of relevant cardiovascular disease or systemic diseases.

voxel size 1.4 × 1.4 × 8 mm, and a slice gap of 2.0 mm. The temporal resolution was interpolated to 25 phases per cardiac cycle (28-37 ms). The protocol includes a complete cine short-axis ventricular stack with the base to apex coverage acquired using balanced steady-state free procession (bSSFP) with one breath-hold per image slice. Short axis cine images were obtained and analyzed. Semi-automatic contouring of LV endocardial and epicardial end-diastolic (ED) and end-systolic (ES) borders was performed with Circle 42 (CVi 42) software (Calgary, Canada). A total of 9 existing CMR indices were quantified, including LV ejection fraction (LVEF), end-diastolic and end-systolic LV volumes (EDLVV, ESLVV), LV mass, inter-ventricular septum (IVS), posterior wall (PW) thickness, asymmetry (IVS/PW), and Petersen and Jacquier coefficients (Petersen et al., 2005; Jacquier et al., 2010). Right ventricle contours were not provided for the study; thus features were not considered. However, these diseases are predominantly related to the LV cavity, therefore missing information from the right ventricle was not considered relevant. Fractal dimensions (Captur et al., 2013) were not included in the experiments due to its limited applicability in daily clinical routine and its lack of prognostic correlation (Ivanov et al., 2017). Additionally, Petersen and Jacquier coefficients (Petersen et al., 2005; Jacquier et al., 2010; Captur et al., 2013) were considered more validated for this experiments. All CMR analyses were performed by an expert cardiologist with several years of experience in the field.

3.2.3 Radiomics extraction

From the region of interest provided by the LV endocardial and epicardial contours, radiomics were extracted from end-diastole (ED) and end-systole (ES) phases, following a pre-established pipeline from the open-source Python (Van Rossum and Drake, 2009, version 3.7.9) PyRadiomics library (Griethuysen et al., 2017a, version 3.0). A set of 420 radiomics features were extracted from the LV cavity (LV) and LV myocardium (LVMYO) within the original filter, including different types: 52 shape, 72 first-order, and 296 texture features (see Suppl. Material for a full list of radiomics extracted). We perform a radiomics features characterization by considering both

TABLE 3.2: Cohort characteristics. The table presents the average and standard deviation for the characteristics of the entire cohort, including age, sex, percentage of men and BMI. On the right are presented the characteristics of each disease group.

Characteristics	Full cohort	DCM	HCM	LVNC	NOR
Sample size	118	37	25	35	21
Age (years)	49.39 ±17.97	50.24 ±15.12	60.88 ±17.84	44.57 ±17.46	40.85 ±16.99
Sex (M/F)	76 / 42	26 / 11	18 / 7	19 / 16	13 / 8
Male percentage (%)	64.41 %	70.27 %	72.00 %	54.28 %	61.90 %
BMI (Kg/m^2)	26.37 ±3.65	27.12 ±3.89	26.32 ±2.59	25.24 ±3.45	25.67 ±4.38
EDLVV (ml)	169.02 ±63.88	225.40 ±67.72	126.54 ±34.58	159.07 ±49.61	136.87 ±24.47
ESLVV (ml)	97.03 ±66.49	158.19 ±75.45	51.34 ±21.01	87.36 ±45.56	57.65 ±12.52
LV Mass (g)	122.29 ±46.06	153.04 ±51.70	139.05 ±46.65	95.53 ±26.37	102.52 ±18.70
LVEF (%)	47.18 ±16.03	32.47 ±12.72	60.23 ±9.79	47.26 ±14.14	58.07 ±4.09
Petersen coefficient	1.77 ±0.90	1.61 ±0.53	0.80 ±0.39	2.86 ±0.48	1.45 ±0.54
Jacquier (%)	17.47 ±10.08	16.70 ±4.80	12.32 ±2.72	24.30 ±15.66	14.10 ±3.19
ISV (mm)	11.15 ±4.86	9.40 ±1.70	19.12 ±4.41	8.48 ±1.61	9.09 ±1.57
PW (mm)	7.26 ±1.42	7.48 ±1.46	8.25 ±1.49	6.51 ±1.10	6.95 ±0.92
Assymetry (IVS/PW)	1.51 ±0.57	1.27 ±0.18	2.34 ±0.75	1.30 ±0.18	1.30 ±0.16

TABLE 3.3: Mann Whitney U statistical test for demographics and clinical characteristics group comparison. P-values provided for existing CMR indices involved in classification. *IVS*: Inter-ventricular septum. *PW*: Posterior wall.

Demographics	DCM-HCM	DCM-LVNC	DCM-NOR	HCM-LVNC	HCM-NOR	LVNC-NOR
Age (years)	0.011	0.08	0.05	0.0011	< 0.001	0.26
BMI (Kg/m^2)	0.354	0.144	0.208	0.378	0.415	0.432
CMR indices						
EDLVV (ml)	< 0.0001	< 0.0001	< 0.0001	< 0.0036	0.08	0.031
ESLVV (ml)	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.028	< 0.0014
LV Mass (g)	0.02	< 0.0001	< 0.0001	0.0011	0.013	0.1214
LVEF (%)	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.099	0.0005
Petersen coefficient	< 0.0001	< 0.0001	0.094	< 0.0001	< 0.0001	< 0.0001
Jacquier (%)	< 0.0001	0.014	0.019	< 0.0001	0.040	0.0005
IVS (mm)	< 0.0001	0.017	0.304	< 0.0001	< 0.0001	0.09
PW (mm)	0.014	0.0007	0.058	< 0.0001	0.0006	0.046
Asymetry	< 0.0001	0.11	0.12	< 0.0001	< 0.0001	0.4934

the ED and ES phases to be able to identify disease-specific changes over a heartbeat cycle of the radiomics features.

3.2.4 Machine learning scheme

From the previous section, a total of 420 radiomics features were extracted and were potential candidates for inclusion in the targeted radiomics model for disease classification. However, not all of these features will have predictive power, and hence feature selection will be first applied to select the most optimal features for the classification task. We separated this feature selection process into 2 steps. First, we identify those that are highly correlated, for each feature, and remove them from the radiomics set as they carry a similar predictive signal. For this purpose, we estimated Pearson correlation between all features, and those above 0.9 were considered redundant. It is well-known that radiomics are highly redundant thus, with this first step we only aimed to remove the most correlated ones, and be further reduced with a more sophisticated feature selector. The procedure resulted in a reduction from 420 radiomics features to only 120. This reduced subset is introduced in the Pipeline function from Python Sci-kit learn package (Pedregosa et al., 2011, version 0.24.2) that has three different steps, including the additional feature selection method mentioned above:

1. **Normalization:** Data normalization is required before introducing the data into the machine learning models because different scales in the variables measured represent that features have different contributions to the model fitting, and this may introduce bias. We applied the MinMaxScaling function from the Sci-kit Python library (Pedregosa et al., 2011). The StandardScaler was also considered, although no significant difference was found with a min-max scaling.

2. **Feature selection:** The number of features remaining was still large and had to be reduced before reaching the model building section. For this purpose, the SelectKBest function from Python's Sci-kit learn library (Pedregosa et al., 2011) was performed. The algorithm works by selecting the best features based on univariate statistical tests. It selects the features according to two different parameters: highest *score function* and number of features (k). These parameters had to be defined beforehand. The score function parameter selected was *f_classif*, which computes the ANOVA f-value for the sample and provides the associated p-value for each feature based on the correlation with the class label. The *k* parameter is defined as the number of features selected by the feature selector. Prior the analysis, we do not know the optimal number of features to be selected, therefore the *k* value before the analysis had to be defined within the range of parameters from 5 to 120 (i.e. number of possible features it may select), to be later introduced in the hyper-parameter optimization Grid Search CV. The number of features selected is tested iteratively, selecting the best k radiomics according to the score function and tested further with the model.

3. **Model building:** Three different machine learning algorithms were trained and tested: One Vs Rest Support Vector Machine (SVM), Multi-class Random Forest (RF), and Multi-class Logistic regression (LR) for classification (Pedregosa et al., 2011). According to a recent review (Martin-Isla et al., 2020), both SVM and RF models were the most used techniques for conventional ML

for image-based diagnosis. Additionally, prior knowledge from a recent publication (Rauseo et al., 2021) proved also the reliability of SVM and RF when dealing with radiomics. Finally, we decided to include Logistic Regression as one of the most-used techniques in statistical analysis for the purpose of increasing the comparison benchmark.

For evaluation, the experiment is validated in a nested CV scheme Cawley and Talbot, 2010. We performed a 10-fold outer loop and a 3-fold inner loop (see Figure 3.1 for a more graphical description). This represents that for each fold in the outer loop, 90% of the data is kept for training and validation, while the remaining 10% will be held for testing. The same procedure was performed in the inner loop for each fold. The remaining train and validation data were split into 66% for training and 33% for validation. All the splits in our scheme were performed with Stratified K-fold sci-kit learn function (Pedregosa et al., 2011) to keep classes balanced. The normalization and feature selection steps were performed in each fold of the nested CV scheme to avoid data leakage. This means that no knowledge of the held test set was introduced into the training stage, which could corrupt the learning process and its posterior generalization.

Models' performance are dependent on the hyper-parameters selected Probst, Boulesteix, and Bischl, 2019. For this purpose, a Grid Search CV (Pedregosa et al., 2011) was applied in the inner loop, thus ensuring the optimal hyper-parameters were selected. Grid Search CV (Pedregosa et al., 2011) is an optimizer algorithm that calculates the model's performance for each combination of hyper-parameters and keeps the one that achieved the highest prediction metric, to be later tested on the testing held data. (see Table 3.4 in Suppl. Material for the full list of hyper-parameters used).

Paired t-test on both distributions of testing AUC performances was performed to analyze the statistical significance for each general machine learning model across CMR indices and radiomics, as well as for each differential diagnosis (i.e. identifying a single disease class from the whole cohort) and prove they were comparable. Additionally, Receiver Operating Characteristic (ROC) curves were calculated to provide a better representation of the true and false-positive rates for each differential diagnosis. Due to the architecture of our Nested CV scheme, we obtain 10 different tests AUC, one per fold (10Fold outer loop, see Figure 3.1, left side). This means that each of the 10 models resulted from the Grid Search CV in the inner loop might be different (i.e. the combination of hyper-parameters and the number of features selected may vary depending on different characteristics of the training and validation set). Thus, to present the most relevant features on average in a single list, we analyzed the features selected by each of the 10 models and selected for representation those features that remained constant across the 10 folds and were finally sorted by feature importance score. With this, we create a highly approximated list of the most relevant radiomics features for each classification.

Since it is possible that different iterations selected a different number of features (i.e. for example, fold 1 could select 30 features and fold 2, 40 features), we analyzed the validation AUC values for each number of features (k=10,20,30,40...) across all the combination of hyper-parameters to see the effect of increasing the number of features on AUC (see Figure 3.3). Finally, to provide a more clinical perspective, we analyzed the implications of feature type (shape, first-order, or texture), region (LVMYO, LV cavity), and phase (ED, ES) for each differential diagnosis, and linked them with the existing clinical knowledge.

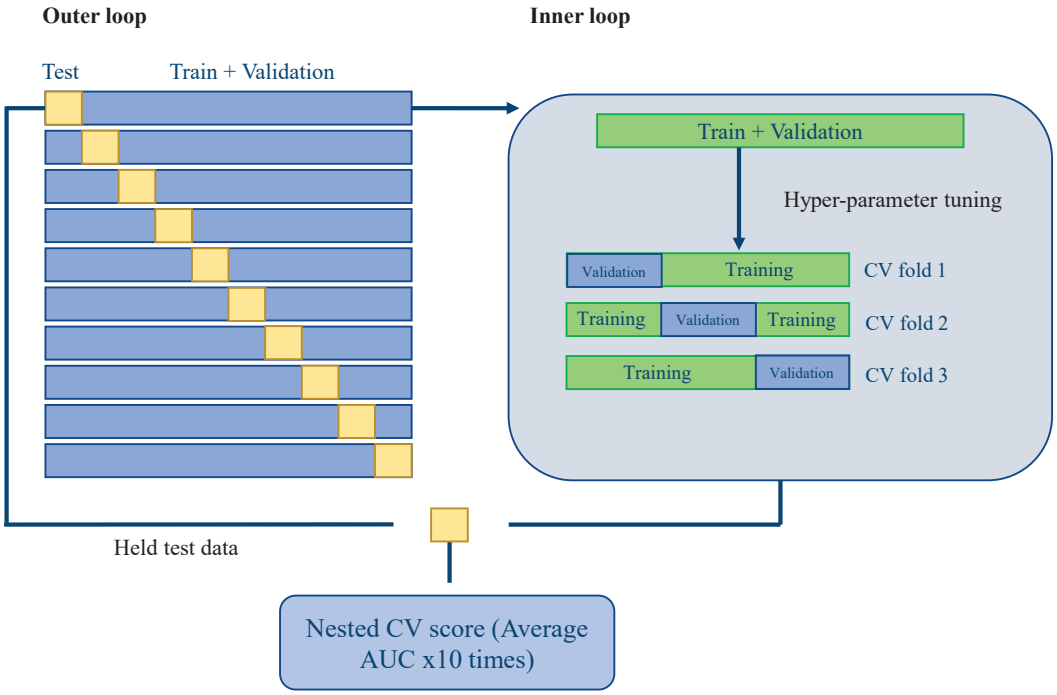

FIGURE 3.1: Machine learning validation: 10-fold outer loop Nested CV and a 3-fold CV inner loop scheme.

3.3 Results

3.3.1 Differentiation of LVNC, HCM, DCM and Normal subjects

In the first experiment, we evaluated and compared the performance of the different machine learning models (namely RF, OVR-SVM and LR). As it can be seen in Table 3.5, the highest AUC values were obtained by the RF technique, for all models. Hence, the RF technique is used as the baseline models in the remainder of the experiments.

Subsequently, we performed a comparison between the AUC scores obtained by the existing CMR indices (i.e. standard model) and radiomics models for classification. As it can be seen from Table 3.5, the radiomics model had a comparable performance to the standard model and there was no statistically significant differences between the two models ($p\text{-}value > 0.05$). However, the radiomics model was obtained without the expert delineation of the trabeculae. In more detail, the ROC curves for the classification models are presented in Figure 3.2.

3.3.2 Radiomics signatures analysis

In this section, we provide more details on the contributions of the different radiomics features to the classification models. We analyzed incrementally the number of features selected for all the hyper-parameter combinations and we showed in Figure 3.3 that the AUC values did not increase significantly after integrating 30-40 radiomics features in the model.

To further illustrate the predictive power of the radiomics features, Table 3.7 presents across subsections the 10 best performing radiomics for each differential diagnosis, sorted by their weighted feature importance (in percentage). Moreover, Table 3.8 shows the 10 best radiomics features involved for the general RF model.

TABLE 3.4: Hyper-parameters grid for optimization.

Model	Hyper-parameter: [range]
One-vs-Rest SVM	C: [0.01,1,10,100,1000] gamma: [1,0.1,0.001,0.0001] kernel = ['rbf']
Random Forest	bootstrap: ['True','False'] min_samples_leaf: [3, 4, 5, 6] n_estimators: [100, 200, 300, 500, 1000] min_samples_split: [8, 10, 12] multi_class='ovr'
Logistic Regression	penalty: [l1] Solver: ['liblinear'] multiclass = ['ovr'] max_iter = [100,150]
*Select KBest	k: [5,10,15,20,25,30,40,50,60,70,80,90,100,110,120]

C: regularization parameter. RBF: Radial Basis Function kernel. OVR: One-vs-Rest.

TABLE 3.5: Summary table of the testing performance of the selected models. The table provides AUC values for CMR existing indices and radiomics, along with the p-value.

	CMR indices		**Radiomics**		**p-value**
Models	**Mean**	**STD**	**Mean**	**STD**	
One Vs Rest SVM	0.972	0.03	0.942	0.03	> 0.05
Random Forest	**0.978**	**0.03**	**0.964**	**0.01**	> 0.05
Logistic Regression (multinomial)	0.970	0.03	0.956	0.03	> 0.05

TABLE 3.6: Generic and differential diagnosis AUC testing metrics for Random Forest model. P-values are presented in the table for statistical significance analysis and prove they are comparable.

	CMR indices		**Radiomics**		**p-value**
General model	**Mean**	**STD**	**Mean**	**STD**	
Random Forest	0.978	0.03	0.964	0.01	> 0.05
Differential diagnosis models					
DCM-vs-Rest	**0.97**	**0.02**	0.93	0.03	> 0.05
HCM-vs-Rest	**1.00**	**0.00**	0.99	0.03	> 0.05
NOR-vs-Rest	0.95	0.02	**0.97**	**0.02**	> 0.05
LVNC-vs-Rest	**0.96**	**0.04**	0.92	0.03	> 0.05

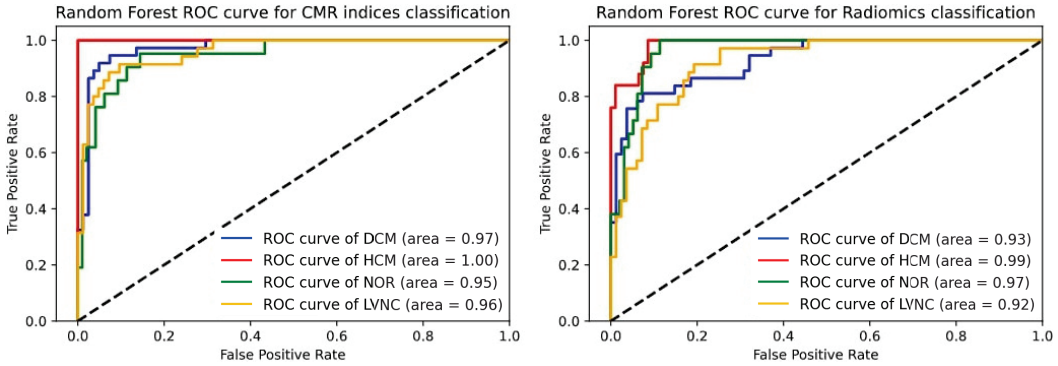

FIGURE 3.2: Roc curve comparison for differential diagnosis. The left subfigure represents the CMR indices ROC curve. The right subfigure represents the radiomics ROC curve.

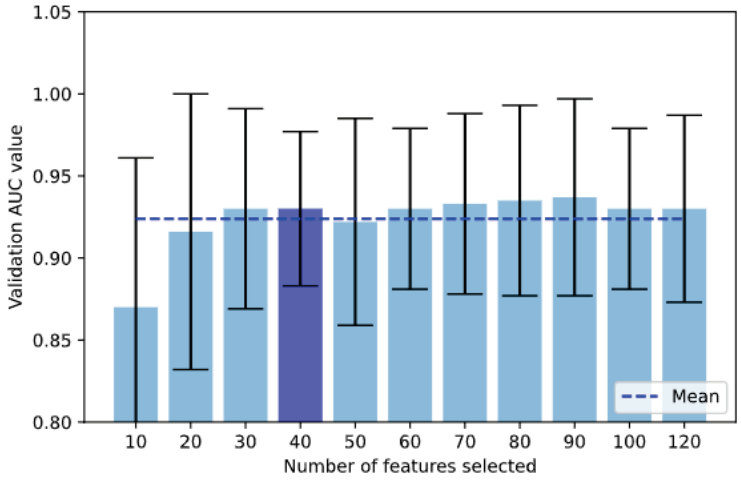

FIGURE 3.3: In our grid search scheme, we analyzed incrementally the number of features selected for all the hyper-parameter combinations and we show that validation AUC was not increasing significantly once reached 30-40 features, with the lowest STD at 40 features. AUC when selecting under 10 features reduced drastically.

By looking at Figure 3.4, we can observe that shape features play an important role when classifying DCM against the rest of the diseases. Alternatively, texture features have a higher impact in the classification of HCM and healthy subjects, and even a higher impact in the differential diagnoses of LVNC.

Left side of figure 3.5 shows the distribution of the selected radiomics features across the cardiac structures (i.e. LVMYO vs. LV cavity). For the identification of DCM, most of the radiomics features (65%) pertained to the LV cavity, while the remaining 35% belonged to the LVMYO. Conversely, features extracted from the LV cavity and the LVMYO participated equally to the identification of HCM and healthy subjects. Finally, the largest difference in terms of region importance can be found for LVNC, where almost 96% of the features belonged to LVMYO, with only a 4% to LV cavity (see Figure 3.5, left image.). This last finding is in line with the existing clinical knowledge. Normally, papillary muscles are considered inside the LV cavity and not as a myocardial mass, according to the guidelines and clinical consensus among cardiologists. But in patients with LVNC conditions, trabeculae and papillary muscles are quantified (within Jacquier coefficient (Jacquier et al., 2010) outside

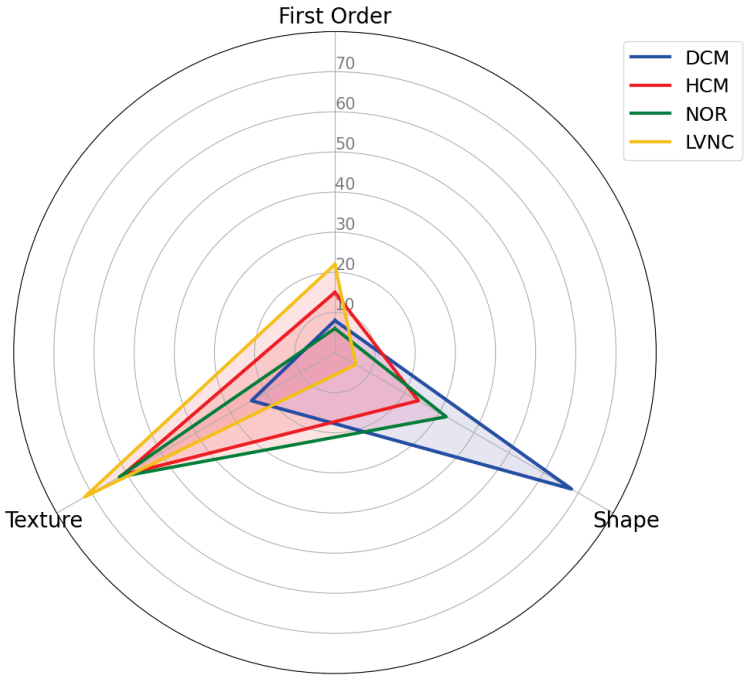

FIGURE 3.4: Differential diagnosis overlapping radar plot, comparing the distribution of the selected radiomics across types of radiomics.

the LV cavity and included in the myocardial mass (LVMYO). For this reason, contrary to what the name itself suggests, the assessment between LVNC and the rest of cardiomyopathies is determined by changes or differences in the LVMYO.

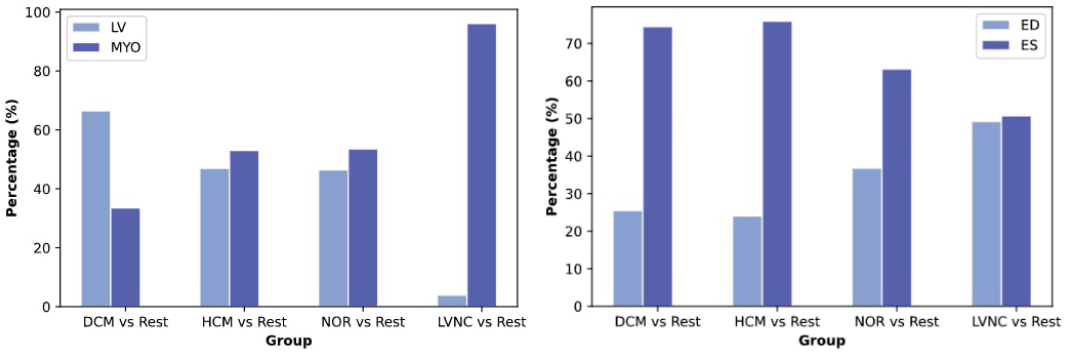

FIGURE 3.5: Distribution of the selected radiomics features across the cardiac structures (left image, i.e. LV myocardium (LVMYO) vs. LV blood pool (LV)) and cardiac cycle phase (right image). For this analysis, all the radiomics selected by the RF model were used.

Regarding the contribution of cardiac cycle phases, ES radiomics were more important than ED features for the classification of DCM, HCM, and healthy subjects. However, ES and ED phases play a similar role when assessing LVNC (See Figure 3.5, right image).

Finally, we compared the time needed to obtain the diagnoses using the standard as well as the proposed radiomics models. With the existing CMR indexes, our clinical experts spent approximately 9-12 minutes to delineate the trabeculae and

then derive an LVNC diagnosis, on average. In contrast, for the proposed radiomics-based approach, the time to assess one subject was reduced to 10-20 seconds depending on the image characteristics (i.e. volume, slice images or bin width Griethuysen et al., 2017a).

3.4 Discussion

3.4.1 Summary of findings

Machine learning-based radiomics models showed excellent performance for differentiating between hypertrophic cardiomyopathy, dilated cardiomyopathy, left ventricular non-compaction, as well as healthy subjects. According to the results presented in Table 3.5 and Figure 3.2, radiomics and existing CMR indices resulted in similar performances.

The 10-most significant radiomics features for the general RF model comprise a combination of radiomics types, regions, and heart cycle phases (see Table 3.8). Looking in detail, myocardium sphericity seems to play an important role in the overall classification, occupying the first and third spots in terms of predictive power for the ES and ED phases, respectively. This can be explained by the remodeling of the heart that affects the global and regional structure of the left ventricle.

Moreover, texture features were found to add substantial information, positioning in the second top position of the ranking, which underlines the widely accepted diagnostic importance of myocardial tissue characteristics such as myocardial fibrosis (See Table 3.8). Specifically, Long Run High Gray Level Emphasis is a texture feature that explains longer contrasted gray level strings/regions, which can be related to the existence and prominence of myocardial trabeculations, typically associated with LVNC.

Regarding the radiomics features for each differential diagnosis, differences were found among the various diseases. Concretely, we can observe how the contribution of the radiomics features to the prediction are distributed, by type (Fig. 3.4), region and phase (Fig.3.5). While shape features seem to play the most important role in DCM classification, texture features are more important when classifying HCM and LVNC subjects. This finding is in line with clinical knowledge: DCM is defined primarily by a dilation of the left ventricle, while myocardial fibrosis has a pivotal role in HCM diagnosis and LVNC is defined by the presence of myocardial trabeculations.

3.4.2 Limitations and future work

The findings presented in this chapter must be considered in light of the study limitations, and future work may take different directions. Firstly, while the radiomics performance for automated diagnosis is promising, these results were obtained based on a single-centre small-size clinical cohort. To confirm these promising results, future studies should be extended towards multi-centre studies. Furthermore, this work relied on semi-automated, manually controlled, delineations of the LV endo and epi-cardial contours on the short-axis images, before the extraction of the existing indices and radiomics features. However, automatic segmentation of the ventricular boundaries has been extensively investigated using deep learning (Campello et al., 2021) and these models could be extended to segment pathological cases in particular for LVNC. Finally, this chapter focused on the diagnosis of LVNC and related CMs. In clinical practice, subsequently, prediction of LVNC related events before

they occur would enable early and personalized prevention. Our plan is to extend the proposed radiomics models to enable patient-specific risk prediction and prognosis estimation after LVNC diagnosis.

3.5 Conclusions of Chapter 3

CMR radiomics constitutes a promising approach to differentially diagnose overlapping and complex conditions such as HCM, DCM, and LVNC. The classification performance of radiomics models are in-line with the one obtained by using existing CMR indexes but the diagnoses can be reached fully automatically without the need for expert delineation of the trabeculae as in previous works.

TABLE 3.7: Top-10 best performing radiomics for each differential diagnosis, divided across subsection.

DCM vs Rest				
Radiomics feature	Type	Region	Phase	Weight (%)
Minor Axis	Shape	LVMYO	ES	7.0
Volume	Shape	LV	ES	6.7
Least Axis	Shape	LVMYO	ES	5.8
Max2D diameter Slice	Shape	LV	ES	5.5
Least Axis	Shape	LV	ED	5.4
Least Axis	Shape	LVMYO	ED	5.1
Long Run High Gray Level Emphasis	GLRLM	LV	ES	5.1
Minor Axis	Shape	LVMYO	ED	5.0
Volume	Shape	LV	ED	5.0
Sphericity	Shape	LVMYO	ES	5.0
HCM vs Rest				
Radiomics feature	Type	Region	Phase	Weight (%)
Surface Area to Volume Ratio	Shape	LVMYO	ES	7.4
Sphericity	Shape	LVMYO	ES	7.2
Large Dependence High Gray Level E.	GLDM	LV	ES	6.9
Long Run High Gray Level E.	GLRLM	LV	ES	6.7
Sphericity	Shape	LVMYO	ED	6.2
Skewness	First Order	LV	ES	5.3
Gray Level Non Uniformity	GLSZM	LV	ES	5.1
Autocorrelation	GLCM	LV	ES	5.0
Energy	First Order	LV	ES	4.9
Surface Area to Volume Ratio	Shape	LV	ES	4.8
NOR vs Rest				
Radiomics feature	Type	Region	Phase	Weight (%)
Gray Level Non Uniformity	GLRLM	LV	ES	6.5
Run Length Non-Uniformity	GLRLM	LVMYO	ES	6.0
Sphericity	Shape	LVMYO	ES	5.9
Low Gray Level Run E.	GLRLM	LVMYO	ED	5.7
Dependence Variance	GLDM	LVMYO	ES	5.7
Max2D diameter Slice	Shape	LVMYO	ES	5.2
Gray Level Non Uniformity	GLSZM	LVMYO	ES	5.2
Long Run High Gray Level Emphasis	GLRLM	LVMYO	ED	5.2
Max2D diameter Slice	Shape	LV	ES	5.0
Minor Axis	Shape	LV	ED	4.9
LVNC vs Rest				
Radiomics feature	Type	Region	Phase	Weight (%)
Large Dependence Low Gray Level E.	GLDM	LVMYO	ES	6.6
Surface Area to Volume Ratio	Shape	LVMYO	ED	6.5
Inverse Variance	GLCM	LVMYO	ES	6.1
Long Run Emphasis	GLRLM	LVMYO	ES	5.9
Large Area Low Gray Level E.	GLSZM	LVMYO	ES	5.7
Large Dependence Low Gray Level E.	GLDM	LVMYO	ED	5.6
Large Area Emphasis	GLSZM	LVMYO	ED	5.2
Zone Percentage	GLSZM	LVMYO	ED	5.1
Long Run Emphasis	GLRLM	LVMYO	ED	5.1
Percentile 90	First Order	LVMYO	ES	4.8

TABLE 3.8: Top-10 best performing radiomics for RF model.

General RF model

Radiomics feature	Type	Region	Phase	Weight (%)
Sphericity	Shape	LVMYO	ES	4.9
Long Run High Gray Level Emphasis	GLRLM	LV	ES	4.6
Sphericity	Shape	LVMYO	ED	3.6
Surface Area to Volume Ratio	Shape	LVMYO	ES	3.6
Gray Level Non Uniformity	GLSZM	LV	ES	3.4
Minimum	First Order	LVMYO	ED	3.4
Least Axis	Shape	LV	ED	2.5
Large Area Low Gray Level Emphasis	GLSZM	LVMYO	ES	3.0
Volume	Shape	LV	ED	2.9
Coarseness	NGTDM	LV	ES	2.2

Chapter 4

Atrial Fibrillation Prediction by Combining ECG Markers and CMR Radiomics

4.1 Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia. It is characterized by an irregular heart rhythm and often abnormally rapid heart rate. The most common complications of AF are increased risk of stroke, heart failure, and death (Chugh et al., 2014). These complications may be mitigated by early AF detection and initiation of appropriate treatments, such as anticoagulation and rate control therapies. Cardiac structure and electrical activity are two important, inter-linked aspects of cardiac health and disease. The cardiac conduction system is complex and depends on the global and local structure of the cardiac chambers. The occurrence of AF is linked to distinct electro-anatomic cardiovascular remodeling (Pellman and Sheikh, 2015). Electrical recordings of the heart such as the 12-lead electrocardiograms (ECG) provide indications of cardiovascular health. The ECG is a dynamic physiological signal that represents the electrical activity of the heart. It is widely used to identify patterns or abnormalities in cardiac rhythms and waveforms. ECG recordings are the main clinical tool for AF diagnosis (Hagiwara et al., 2018). The best indicators are the absence of the p-wave degenerating into small magnitude fibrillatory waves and the irregularity of R-R intervals indicating irregular conduction of atrial impulses through the atrioventricular (AV) node to the ventricles. The study of the QRS complex, a combination of the Q wave, R wave and S wave that represents ventricular depolarization, might also add some information by analyzing the height of the amplitude or the size of the interval. But the latter indicator might have normal values even when the AF is present (Fuster et al., 2006). Furthermore, AF frequently occurs intermittently with the characteristic AF-defining features only apparent when an individual experiences a paroxysm of AF. Whilst paroxysmal AF is more challenging to diagnose, it confers the same adverse risks as individuals continuously in an AF rhythm (Lip, 2001). Emerging deep learning approaches have shown promise in quantifying complex patterns in cardiac electrical activity (Somani et al., 2021; Liu et al., 2020). However, there is room for improvement. For patients with undiagnosed AF, ischemic stroke may be the first clinical manifestation of the condition. AF is detected for first time in approximately one-fourth of patients presenting with ischemic stroke(Sposato et al., 2022; Howlett et al., 2015). Early detection of AF may enable early intervention and prevention of ischemic stroke. There are numerous conference challenges, particularly organized by Physionet, which aim to address

early detection using machine learning techniques. In spite of the successful results, the existing works in the literature do not stratify by sex. This is an important consideration given significant sex differential patterns in AF highlighted in clinical papers. The estimated prevalence of AF is lower in women, whilst this may reflect genuine lower burden of AF in women it may also indicate under-diagnosis in this population (Ko et al., 2016). Indeed, women with AF experience higher mortality and ischemic stroke and are less often prescribed anticoagulation treatment (Kassim et al., 2017). Cardiovascular Magnetic Resonance Imaging (CMR) plays an important role in the diagnosis of complex cardiac diseases. Recently, the concept of radiomics has attracted significant attention in the cardiac imaging community (Raisi-Estabragh et al., 2020b) due to its ability to quantify and analyse large pools of advanced imaging phenotypes, which are descriptive of complex shape, size, intensity or textural patterns. Preliminary results have shown the promise of CMR radiomics for AF discrimination (Cetin et al., 2018). CMR radiomics extracts a large number of quantitative features using data characterization algorithms. These techniques are very promising for deeper image phenotyping of cardiac structure and tissue (Raisi-Estabragh et al., 2021b). The combination of imaging phenotypes and ECG features for AF detection has not been explored in the existing literature. Yet, such an approach may enable integration of complementary signals and hence improve AF detection by considering both anatomical and electrical alterations. In this chapter, we aim to evaluate the feasibility of combining cardiac imaging with ECG features for AF detection considering sex-differential patterns. Integrated risk prediction models were built combining CMR radiomics and ECG parameters, separately for men and women. Morphological, temporal and non-linear features were extracted from the ECG waveforms. The study was performed using the UK Biobank resource, a large-scale health database publicly available under request. To our knowledge, it is the first time, that the combination of ECG and imaging are explored. The inclusion of radiomics allows a more precise information of the AF event and quantifies the complexity of cardiac structure and remodeling providing a complementary information additionally to the ECG test.

4.2 Related work

The related works can be divided into three categories according to the computation of the features: ECG features extracted from the waveforms with machine learning techniques, ECG features extracted from deep learning methods and hybrid frameworks that combines traditional ECG features with the ones extracted using deep learning algorithms. Classical approaches were mainly based on morphological features of the ECG signal in time domain such as heartbeat, analysis of intervals and amplitudes of QRS, QT, PR and R-R (Athif, Yasawardene, and Daluwatte, 2018; Zong, Mukkamala, and Mark, n.d.). Those studies, with satisfactory results, may be sensitive to the ECG noise. To alleviate this issue, the morphological features were computed in other domains such as in the frequency or time–frequency domain. Some examples of these features are power spectral density of the R-R intervals and frequency bands (e.g., ultra low , very low and low). Non-linear features were also considered as the model of the heart cannot be reduced to a linear function as it also involves a nonlinear contribution (Acharya et al., 2006; Rizwan et al., 2021). Some works combining ECG features in different domains are the following: Yin et al., 2019 proposed a multi-domain ECG feature extraction method20. The multi-domain features were composed of nonlinear and frequency domain features, which were

used as input features to train and test an SVM classifier model. Zabihi et al., 2017 also proposed a multi-domain ECG feature extraction which included time-domain features, time-frequency, phase-space based on non-linear features and meta-level information. Random forest classifier was applied for feature selection as well as for classification. As the technology evolves, more data can be processed, and deep learning techniques emerge. Many works have applied deep learning for feature extraction using convolutional neural networks (CNN)(Zubair, Kim, and Yoon, 2016; Kiranyaz, Ince, and Gabbouj, 2016; Hsieh et al., 2020), long and short memory networks (LSTM) (Schwab et al., 2017; Faust et al., 2018) as well as their variants (Xia et al., 2018; Andersen, Peimankar, and Puthusserypady, 2019; Fan et al., 2018). Other studies have applied deep learning to obtain new features and fused them with traditional features. Some examples are the following. Smoleń, 2017 created an initial model using Recurrent Neural Network (RNN) classifier, that was fed by lengths of intervals between following R peaks. The computed probabilities for each class were combined with hand-designed features and used as an input for Gradient Boosting Machine (GBM) classifier. The features selected were categorized into 5 categories: statistical features, QRS morphology features, RR-interval features, noise features, and frequency-based features. The performance of those methods was very promising. Most of them used publicly available databases within conference challenges. However, the main issue of deep learning techniques is that a large number of samples are required in order to 'learn' and generally, hospitals do not have enough cases to use this type of methods.

Additionally, the AF might not be registered on standard of care 12-lead ECG during hospital visits, opening a new line of research into AF detection using portable devices such as smartwatches (Dörr et al., 2019; Bumgarner et al., 2018; Perez et al., 2019). Those devices might not be as sophisticated and exact as clinical ECG devices. But it has the advantage that it can deal with the early detection of the AF as long periods of signal are recorded. Screening is suggested as one strategy to increase AF detection rates and start anticoagulation in an earlier stage in high-risk individuals. Screening by opportunistic pulse palpation or ECG rhythm strip is already recommended by the European Society of Cardiology (ESC) in all patients older than 65 years contacting health services and by the National Institute for Health and Care Excellence (NICE) where patients have a symptom suggestive of AF (Welton et al., 2017).

Another line of research is using risk factors, biomarkers, ECGs or a combination of risk factors and imaging features in order to predict incident cases of AF. A strong causal relationship between natriuretic peptides NT-proBNP, BNP and MR-proANP, and incidence of AF was ruled out by Geelhoed et al., 2020 Vascular risk factors including diabetes, hypertension as well as daily lifestyle variables such as smoking and obesity(Chyou et al., 2015; Wong et al., 2020) have also been studied in relation to incident atrial fibrillation as well as the inclusion of CMR Imaging (Pujadas et al., 2020), in recent studies. ECG features have also been analyzed to study the possibility to develop AF (Aizawa, Watanabe, and Okumura, 2017). In spite of the promising results, the studies are in an initial stage of research and have not been integrated in clinical routine.

4.3 Methods

4.3.1 Population and Setting

UK Biobank is a large-scale health database containing over a half million of participants aged between 40 and 69 years old and recruited across UK between 2006 and 2010. It is a powerful research resource including biomarkers, medical records, risk factors, clinical tests and physical measurements to study the most common and life-threatening diseases. The database is regularly updated with additional data, making it a potential source for research purposes. AF was detected through the Hospital Episode Statistics (HES) system, a database containing clinical details of all the admissions of the NHS hospitals in England, to provide a continuous follow-up of the participants.

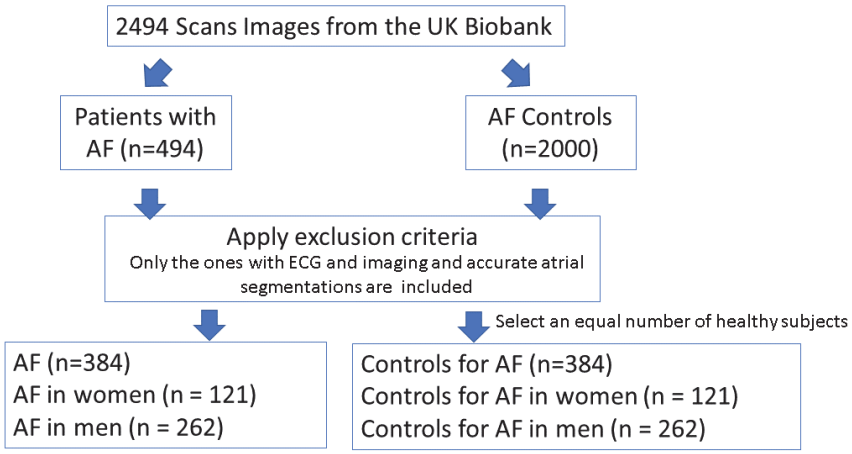

FIGURE 4.1: The process to select the data from the UK Biobank.

4.3.2 Study design and data

From the 495 prevalent AF cases in the UK Biobank cohort, we selected all the patients with AF who underwent both ECG and the CMR scan and the corresponding segmentation of the Left Ventricle (LV) and Right Ventricle (RV) cavities as well as the left and right atria were available (n=383). To analyze the differences between sexes, we separated the data into female (n=121) and male (n=262) participants. Of these, 45 women and 49 men were in sinus rhythm at the time of their ECG recording. The healthy controls were defined as participants who were not diagnosed with AF and had a normal sinus rhythm on their ECG. For the healthy controls, we considered the first 2000 UK Biobank participants for computational purposes with ECG and CMR imaging. To avoid unbalanced models, the same number of healthy controls were randomly selected for each sex (Figure 4.1).

4.3.3 Feature extraction

The features of ECG were extracted in temporal, morphological and non-linear domains. Radiomics features were computed from the LV and RV segmentations in

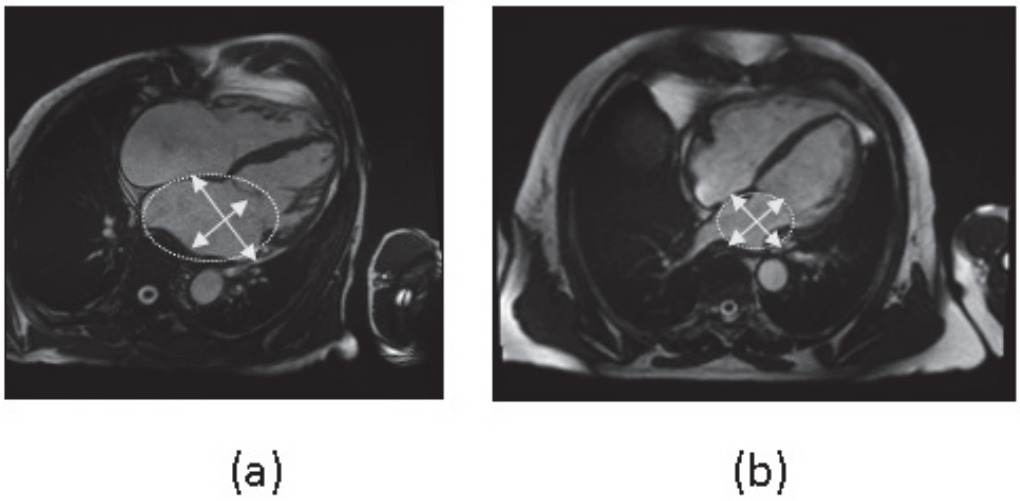

FIGURE 4.2: The figure shows four-chamber cine CMR images in end-diastole from two female UK Biobank participants. Our models selected the most important radiomics features from the left atrial region of interest. The arrows show the axis, and the circular shape indicates the sphericity. The first image (a) shows an AF patient with a larger axis and pronounced oval sphericity. The second image (b) illustrates a healthy subject with normal atrial dimensions, with more circular sphericity and a smaller axis than an AF patient.

end-systole (ES) and end-diastole (ED) phases from short- and long-axis cine CMR images. The radiomics of the atrias were computed from the long-axis images. In this section, we will explain in detail how the features were extracted.

4.3.4 Radiomics Feature extraction

Radiomics features were extracted from the CMR images and the corresponding contours from three segmented ROIs: LV and RV cavities and LV myocardium in ED and ES in short axis. The segmentation of the ROIs was performed manually by expert cardiologists by defining the contours with points with a different label for each ROI using CMR and CT Software (CVI42). The segmentations for each patient were exported in a single xml file containing the contour points for the RV, LV and MYO segmentations. In order to convert each contour into a binary mask, we developed an in-house software that transforms CVI42 contours into readable format contours (*euCanSHare* n.d.). We also obtained the atrial segmentation using an automatic segmentation model based on a traditional U-Net architecture. First, a manual segmentation was performed by clinical experts in 764 datasets from Barts Heart Centre, UK.Then, data augmentation techniques were used for generalizability of the model such as small rotations, random contrast adjustments and random intensity histogram shifting. The Adam optimizer was used with a learning rate of 0.0001 and 0.9 and 0.999 for first and second moments, respectively. The model was then trained with a batch size of 16 256x256 images with 100 epochs. The loss function used was cross entropy. We computed the radiomics using the open-source python-based PyRadiomics library (version 2.2.0). To harmonize the images, the histogram matching technique was applied given a reference image. A binwidth of 25 was used to discretize the grey values of the image as it is the default parameter selected

by pyradiomics. We extracted the relevant information present in the image by using three classes of features: 1. First Order Features: are histogram-based features related to the distribution of the gray level values in the tissue, without focusing on their spatial relationships. 2. Shape Features: describe geometrical properties of the organ, such as volume, diameter, minor/major axis and sphericity. 3. Texture Features: are derived from images and allow quantification of spatial relationships among pixels. The shape radiomics of all the ROIs both for the short axis and long axis were all considered. However, the first-order and textural features were only considered from the LV myocardium as the other ROIs included parts such as the papillary muscles that can alter the intensity signals within the ventricular and atrial cavities. Shape features derived from the LV myocardium, LV, RV and LA, RA were selected for the analysis, while first-order and textural features derived only from the LV myocardium were used. A total of 420 atrial radiomics features were computed in long axis where each ROI contained the same number of features of each type (ROI shape n=24, ROI first-order n=36, ROI texture n=150). Additionally, 262 radiomics features both for short and for long axis were included from each CMR study (LV shape n=26, RV shape n=26, MYO shape n=26, LV myocardium first-order n=36, LV myocardium texture n=148).

4.3.5 ECG Feature extraction

We extracted the features of the ECG signals that are related according to literature with AF. We do not use the whole ECG signal as an input of the classification method to avoid overfitting. The ECG features for morphological, classical and non-linear features were computed using the open source code for ECG feature extraction in AF implemented in Matlab and mainly based on Physionet library (Andreotti et al., 2017). Firstly, classical ECG features were extracted based on morphological features in time domain including heartbeat intervals, analysis of QRS, QT, PR, R-R intervals and amplitude. For robustness, morphological features in frequency domain were also extracted including power spectral density of the R-R intervals and frequency bands (ultra-low, very low, low and high frequency and Ratio of low- to high-frequency power). Finally, non-linear features were also considered as the model of the heart is not only linear but also involves a nonlinear contribution. In this work, Poincaré Plot was used to extract non-linear features in ECG. Poincaré Plot is a 2D dimensional scatter plot where each point represents the RR interval as a function of the previous RR interval. The Poincaré analyzes quantitatively the shape of the plot which provides rich information of the behavior of the heart. For example, the plot for a patient with AF has a more circular shape than a healthy subject that is similar to a comet along the line of identity (Henriques et al., 2020).In order to determine the geometric appearance of the plot quantitively, some techniques such as ellipse fitting, correlation coefficient and histogram-based methods were implemented. Additionally, the Sample entropy was computed to measure the complexity of the time series (Shaffer and Ginsberg, 2017). We proposed a multi-domain ECG feature extraction method including classical, non-linear and frequency domain features with a total number of 116 features. The second lead was used to extract the features as both old devices and the wearable devices are using a single lead. According to literature, the second lead provides the most valuable information (S. Luz et al., 2016; Murat et al., 2021) including P, QRS and T waves. For that reason, it is the most used within the single-lead ECG works and the one with better results from 12-lead ECG recordings (Baalman et al., 2020).

4.3.6 Feature Selection

Chi-squared test is applied to the features, and selects metrics statistically significantly linked to the outcome. The Chi-squared test can be defined as given the data of two variables, we can get observed count O and expected count E. Chi-Square measures how expected count E and observed count O deviate from each other. The formulation is as follows:

A small p-value of the test statistic indicates that the corresponding features is dependent on the outcome, and it is an important feature. The statistical test returns each feature's importance score using the -log of the p-value. A large score value indicates that the corresponding feature is important. In our approach, the number of features selected was 30 as the model stabilizes after 30 features.

4.3.7 Statistical Analysis

The experiments were conducted using the Matlab 2021b software. The correlation between ECG and radiomics was performed using Pearson's correlation. We used the fscchi2 function to apply the Chi-Squared test to select the most relevant features. A hierarchical model was built by combining radiomics with ECG to show the added value of incorporating radiomics features into the model for women, men and for both sexes. For comparison, we built the ECG and Radiomics-based models alone. The models were trained with a Support Vector Machine (SVM) technique which has been widely used in cardiovascular risk predictions ("Discrete Cosine Transform and Support Vector Machines for Classification Cardiac Atrial Arrhythmia and Cardiac Normal" 2020; Martinez-Alanis et al., 2020) due to its numerous advantages such as computationally efficient and robustness for real-world applications as well as the ability to find non-linear relationships through the kernel trick. The models were tested following a nested cross validation also known as double cross-validation, in order to minimize a biased evaluation of the accuracy of the model. Nested cross validation is widely employed in the machine learning field and was mainly developed to work with small datasets. Compared to standard cross validation techniques, nested cross validation can help in the reduction of overfitting and alleviate the limitation of optimistic biases, especially in relatively small samples. Varma and Simon et al., showed that nested cross validation methods provide an almost unbiased estimate of the true error compared to standard k-fold cross-validation particularly when used for both hyperparameter tuning and evaluation (Varma and Simon, 2006; Iizuka et al., 2003; Raschka, 2018). The method is divided into two loops: the inner loop is responsible for the selection of the best parameters, and the outer loop estimates the generalization accuracy66. This procedure splits the data into training and test folds k times in an outer loop. For each training fold, the hyperparameter optimization process is performed in an inner loop and returns the best parameters that minimize the error following the same procedure of partitioning and rotating the training fold into training and validation sets. Using this scheme, the test folds are never used to build the model, decreasing the possibility of overfitting. Notice that we have ten models trained with different partitions of the data not a single partition, making this procedure robust and reliable. Additionally, all the data has been used for testing making the performance measurements more reliable. The hyperparameter optimization procedure was performed using greedy optimization which apply a brute force exhaustive search by trying each combination of each parameter. Five partitions are used for tuning the parameters of the SVM for each training fold in the inner loop (5-cross validation) and 10 cross-validation

for the outer loop with partitions of 90% for training and 10% of testing in each outer fold. The summary of this procedure is shown in Figure 4.3. We computed to assess the performance of the models, the receiver operating characteristics (ROC) curve and area under the curve (AUC), as well as F1-score, accuracy, sensitivity and specificity over the test set. Additionally, Welch´s t-test was computed for group-wise comparisons. Several healthy partitions are randomly selected to show that the model does not depend on the selected data using different random seeds and we computed the ROC curve for each different partition of the healthy cohort. To compare the models, a paired t-test on the distributions of AUC performances was performed to analyze the statistical significance in a nested cross validation framework (Izquierdo et al., 2021). Data Availability The datasets generated and/or analysed during the current study are available online from the , UK Biobank database.

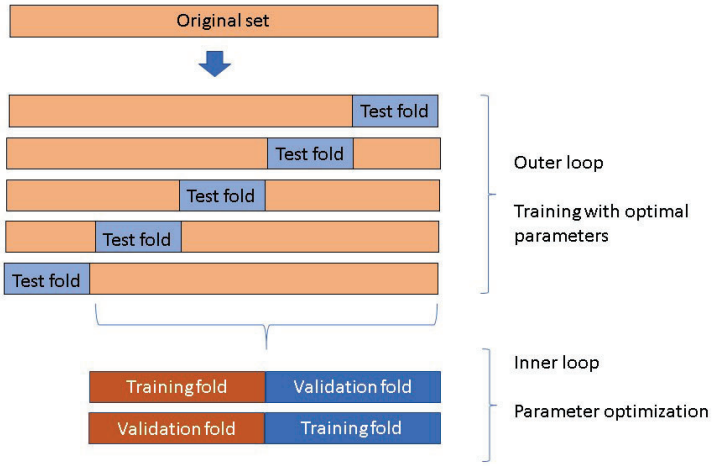

FIGURE 4.3: A nested cross validation scheme.

4.4 Results

4.4.1 Baseline characteristics

We studied 32,121 UK Biobank participants with an average age of 63 (±7.53) years. 51% of the participants were female. A total of 495 participants had prevalent AF. The AF cohort included a greater proportion of men (69.3%), slightly older individuals with greater comorbidity burden, and higher BMI. For most baseline metrics there was no statistically significant difference between men and women except in education level and alcohol intake. Specifically, men were more likely to participate in higher education (48% vs. 34%) than women and consume alcohol more than 1-2 time a week. The Table 4.1 summarizes the baseline characteristics.

4.4.2 Correlation between ECG and Radiomics features

The correlation between ECG and radiomics features was not very high, as illustrated in Figure 4.4. The morphological features were the ones with a certain correlation along all the radiomics features both in short and in long axis. Moreover, a

higher correlation is shown for radiomics features computed from long axis images (vs short axis) as these features include atrial radiomics, particularly in the heart rate variability in temporal and non-linear domain. Thus, ECG features seem to have a higher correlation with the features related to the atria than with the other regions of interest of the heart. However, the correlation found is not high between ECG and radiomics with the two providing additive and complementary information.

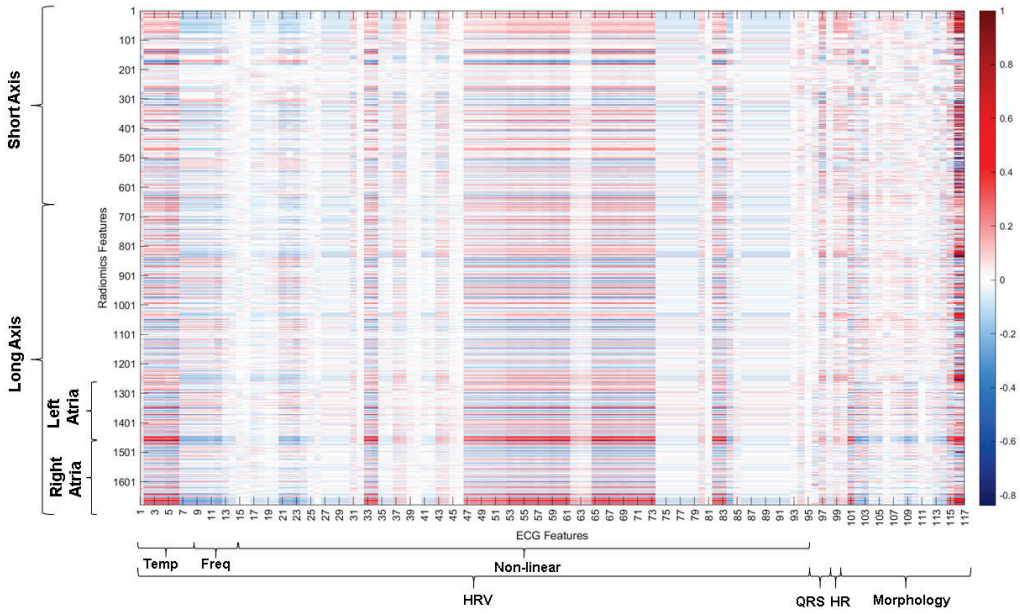

FIGURE 4.4: Correlation between ECG and radiomics features showing low correlation between radiomics features extracted from the short-axis images and a slightly higher correlation with the features from the long axis images including atrial metrics. Temp temporal, Freq frequency, HRV heart rate variability, HR heart rate.

4.4.3 Feature Selection for each Model

For the model that includes both sexes, the ECG features which are related to the heart rate (such as tachycardia) were the most predominant features. The volume and surface of the left atrial were also important features in the model. Most of the relevant radiomics features are first, shape and then texture. The region of interest (ROI) selected for all the features are left atrium (LA) and the phase end diastole (ED). In Table 4.5, all the selected features are shown for the general model.

For the separate model in women, the most predominant features are mean of RR and diameter of the LA. The shape and texture variables are also informative model features. The ranking of importance is lower than the other models. In the model built with only male participants, the sphericity and volume of the LA are selected as the most relevant features followed by the ECG features such as tachycardia and bradycardia. The shape features are the most selected and secondly the texture.

In Table 4.6 and Table 4.7, the features for female and male are described with the ranking score. In Table 4.8, the repeatability of the variables in women is also shown in the partitions of the nested-cross validation.

In the three models, ED phase is selected the most and most of the radiomics features are from the left atrial ROI. The most predominant features are mainly based on shape and secondly textural features. First-order features do not have a high

presence in the models. Figure 4.5 highlights visually the most relevant CMR markers, in the ED phase from the left atrial, for the women case but for men, it would be equivalent. The arrows indicate the axis, and the circular shape shows the sphericity. The AF patient has larger axis with a more oval sphericity than the healthy patient with a more circular shape of the left atrial.

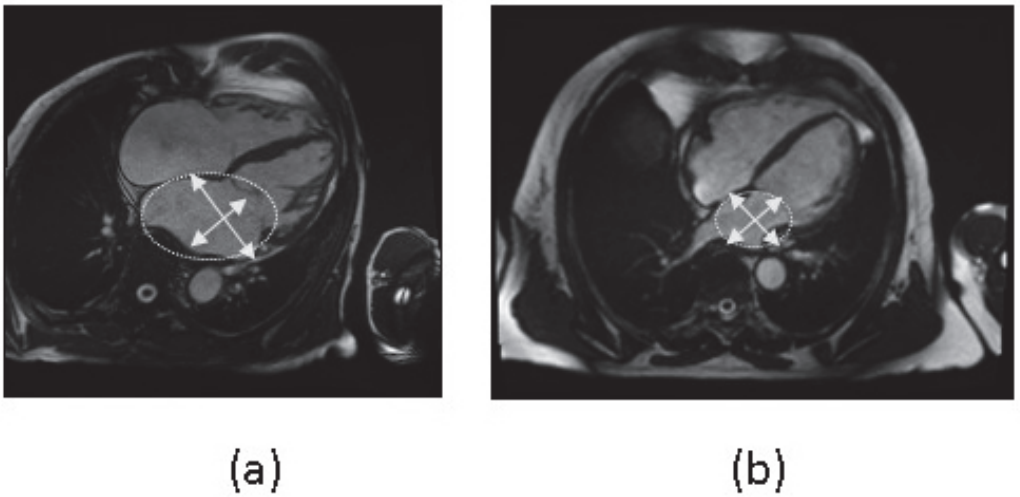

FIGURE 4.5: The figure shows four-chamber cine CMR images in end-diastole from two female UK Bi-obank participants. Our models selected the most important radiomics features from the left atrial region of interest. The arrows show the axis, and the circular shape indicates the sphericity. The first image (a) shows an AF patient with a larger axis and pronounced oval sphericity. The second image (b) illustrates a healthy subject with normal atrial dimensions, with more circular sphericity and a smaller axis than an AF patient.

4.4.4 Performance of Electro-Radiomics Models

Table 4.2 shows the performance of the models adjusted by sex for the whole sample and for men and women separately. In the model with both sexes, radiomics did not show an added value compared with ECG alone or with the combination of both. In sex-specific analyses, we found poorer performance of the ECG model in women than men (AUC: 0.77 vs 0.88, $p<0.05$). The addition of radiomics features improved the model accuracy for women to similar levels as for the ECG only model in men (AUC: 0.87 vs 0.88, $p>0.05$). The sensitivity also increases considerably in women by adding the radiomics (Sensitivity:0.68 vs 0.79) having a higher detection of AF cases. According to our experiments, the addition of radiomics features has greater incremental value for AF discrimination in women than for men, where the added value is not clear. This behavior is not observed if we do not separate the data between men and women. To show, that the added value of radiomics in women does not depend on the data selected, we repeated the experiments with another randomly selected healthy comparator, observing consistent results throughout (Figure 4.6). In order to test the robustness of the results with respect to covariates, we repeated the sex-specific experiments adjusting the models by: i) age and sex ($p>0.05$) and ii) age, sex and main comorbidities related to AF which are diabetes, high cholesterol and hypertension ($p>0.05$). The results in Table 4.3 follow the same pattern than the

model adjusted by sex. It shows the robustness and strength of the features selected related to AF of the models.

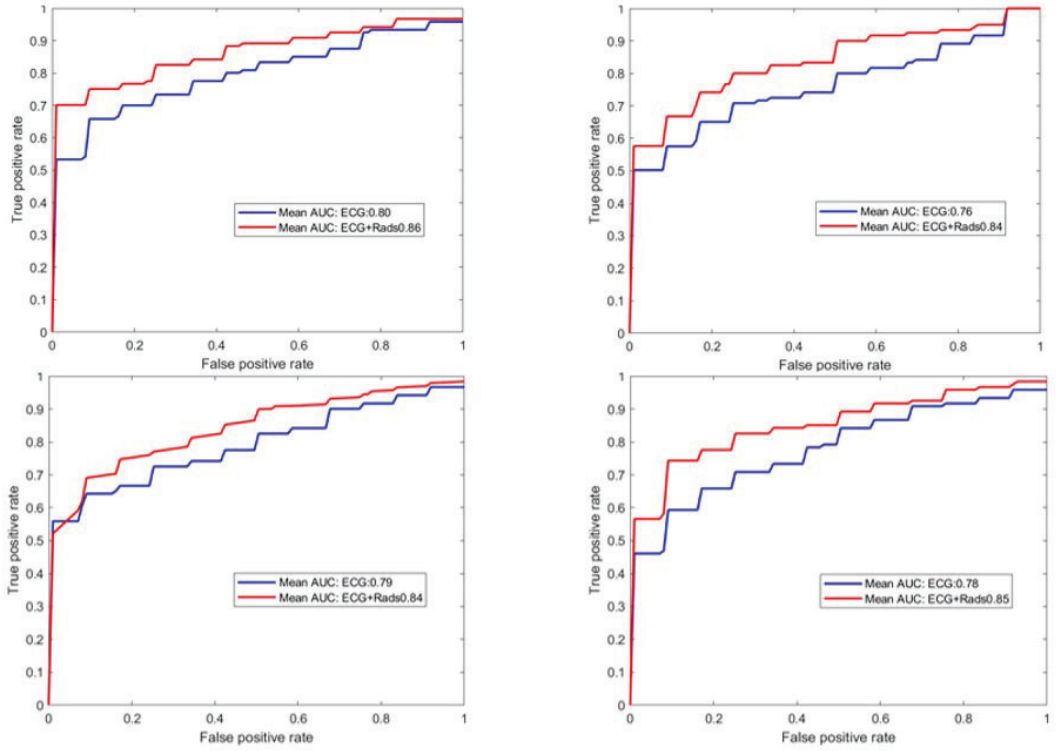

FIGURE 4.6: Different random partitions of the healthy cohort were randomly selected to show the added value of radiomics versus ECG alone in women. For all the cases, the improvement is clear and statistically significant ($p<0.05$).

Finally, we extended the statistical analysis of phenotyping prevalent AF by selecting only the cases with patients of AF with a sinus rhythm without being differentiated with a normal ECG of a healthy patient and randomly matched with the healthy cohort with N=45 and N = 49 for men and women, respectively. Again, the best added value of adding radiomics is for women reaching an 0.72 of AUC vs 0.54 ($p<0.05$). The sensitivity increased significantly compared with ECG alone (0.65 vs 0.72). The best general model combining women and men was ECG+ radiomics with an AUC of 0.61. For men the most predictive model was using radiomics with an 0.59 of AUC. Then, we observe that using radiomics in this scenario, the prediction improves for all cases, particularly in women. A summary of the results is shown in Table 4.4.

4.5 Discussion

In this study we demonstrate the feasibility and clinical utility of using an integrative electro-anatomic model for AF diagnosis. We demonstrate the usefulness of these models in understanding phenotypic alterations that occur in AF. Importantly, we identified different electro-anatomical remodeling patterns in male and female patients with AF. Our findings indicate the usefulness of a more integrative approach to disease in women, who may have more subtle phenotypic alterations than men,

particularly in the early disease stages. As ECG is the main clinical tool for AF diagnosis, we expected ECG to have better results than radiomics alone, as was shown in the results for the men and general models. However, we found lower performance of the ECG model for women than men in AF. This behavior is clearly seen when the models are split into female and male subjects. The combination of ECG with radiomics predictors was able to improve the model performance among female subjects. Radiomics showed less added value for men, however the most relevant features selected by the Chi-Squared test were radiomics-based features, particularly from the left atrial. Although, it did not improve the model's overall accuracy, this finding suggests that radiomics features may precede ECG changes in both men and women. The underlying mechanisms of the sex differences in AF are incompletely understood. The main driving factors reported in the literature are higher body mass index, larger atria and ventricle size among males (Kishi et al., 2015; McManus et al., 2010; Magnussen et al., 2017). Notably, atrial enlargement has been linked to higher risk of incident AF and AF recurrence (Zacà et al., 2007; Raisi-Estabragh et al., 2021a). Moreover a study by Vegte et al., 2021 demonstrated that genetically susceptibility to AF increases indexed left atrial volumes and decreases LA ejection fraction (Vegte et al., 2021). On the other hand, these factors might also impact the interpretation of the ECG signal. Our results suggest that women with AF have less overt ECG changes than men. Indeed, women have a higher heart rate at rest due to hormone effects, autonomic nervous system influences, and intrinsic properties of the sinus node. The P-wave is significantly shorter as well as the PR interval and the QRS duration. QT has also a more prolonged corrected interval in women (Boriani et al., 2017). As an example, prolonged QT interval possibly cause lower sensitivity for ECG in women with leading to false positive cases. Moreover, shorter P-waves with lower amplitude might make ECG recordings susceptible to noise and motion artifacts (Hossain et al., 2019). This means that the subtler radiomics feature changes are important for improving AF detection in women. Due to the more pronounced ECG changes among male participants our model can differentiate between cases and controls with high accuracy using these features alone. Notably, radiomics features appear dominant in the combined models even for men. This suggests that radiomics features are more sensitive at picking up AF-related alterations and these changes may complement the information derived from the ECG. We also performed an extension including only the patients with the diagnosed AF who were in sinus rhythm at time of their ECG. As expected, ECG was not able to distinguish between healthy and unhealthy participants. However, the inclusion of radiomics substantially improved the model performance, particularly in women. Importantly, increased atrial volume (Bertelsen et al., 2020) and atrial fibrosis (Sohns and Marrouche, 2020) might serve as a substrate for AF, and these alterations can be picked up by radiomics features. Although further information is needed to better describe the link between atrial radiomics features and biological precursors of AF. The utility of artificial intelligence-based methods has been already demonstrated in the detection of AF, importantly sex differences are rarely addressed in these studies. The Apple Heart Study assessed the ability of an irregular pulse notification algorithm to identify AF in 419,297 (42% female) individuals (Perez et al., 2019). Overall, 2161 (21% female) participants received a notification and 34% of cases were clinically confirmed from the total number of users detected by the smartwatch. In the study positive predictive value of an irregular pulse notification was 0.84 (95% CI, 0.76–0.92), supporting the ability of the algorithm to correctly identify atrial fibrillation, mainly among white male subjects. Notably, the datasets collected among smart device users rarely permit the assessment of sex differences, as man are more

likely to own these devices in the first place (Guo et al., 2019). AI applications are also used in the monitoring (Gopinathannair et al., 2020), risk stratification (Inohara et al., 2018) and management of AF patients. As a future work, we will extend this research to other cohorts to generalize the models and validate them to external data. With the inclusion of more data, we will also explore deep learning techniques combining all leads with the features that we identified in this work to improve the model accuracy. Moreover, we will also differentiate between certain types of atrial fibrillation to find phenotypes in each category instead of atrial fibrillation patients in general. We will also test the utility of the present model to predict incident AF.

4.5.1 Limitations

Our ascertainment of AF status relied on clinical diagnoses. A limitation of this approach is that we would not capture as yet clinically unrecognized AF cases. As a result, some of the participants labelled as controls in our study may have low burden or paroxysmal AF that is not yet clinically identified. The impact of such misclassification would be attenuation rather than spurious high performance of our models. Additionally, the models were not validated externally limiting the generalizability of our results.

4.5.2 Conclusions

In this study of the UK Biobank participants we demonstrated that an ECG-based model had lower accuracy to detect AF in female subjects compared to males. The inclusion of CMR radiomics combined with ECG increased the model performance in women. Especially CMR derived radiomics shape features of the LA had robust role in the betterment of our models, suggesting the critical role of atrial remodeling in the disease mechanism of AF. The main universal implication is that a combined approach of ECG and atrial imaging might lead to better assessment of female participants suspected of AF. As a further layer of our analysis we selected prevalent AF patients with normal ECG tests, here, we found that all models got benefit from adding radiomics. But again, the clearest case was for women with the inclusion of radiomics with ECG features.

TABLE 4.1: Baseline characteristics describing the whole population, population without AF, patients with AF in general and sex-specific. In bold, the characteristics with not significant differences.

Characteristics	**Whole population (n=32,121)**	**Subjects without AF (n=31,424)**	**Patients with AF (n=495)**	**p-value AF vs non-AF**	**AF in women (n=152)**	**AF in men (n=343)**	**p-value AF in women vs men**
Age mean (std)	63.27 (±15.2)	62.32 (±14.8)	68.43 (±6.23)	<0.001	68.04 (±6.54)	68.61 (±6.55)	
Female sex n (%)	16,658 (51.86%)	16,462 (52.32%)	152 (30.40%)	<0.001	152 (100%)	0 (0%)	0.29
Townsend deprivation index median (IQR)	**-1.95 (±3.30)**	**-2.64 (±3.30)**	**-2.74 (±3.54)**	0.66	-2.73 (±3.50)	-2.74 (±3.55)	0.43
Body mass index mean (kg/m^2)	**26.57 (±4.35)**	**26.55 (±4.34)**	**27.79 (±4.53)**	<0.001	27.91 (±5.41)	27.74 (±4.09)	0.73
Current smoker n (%)	**2032 (6.32%)**	**1993 (6.34%)**	**26 (5.25%)**	0.32	5 (3.28%)	21 (6.12%)	0.19
Hypertension status n (%)	**4397 (13.68%)**	**4177 (13.29%)**	**165 (33.33%)**	<0.001	48 (31.57%)	117 (34.11%)	0.58
High cholesterol status n (%)	**7272 (22.63%)**	**7055 (22.45%)**	**164 (33.13%)**	<0.001	42 (27.63%)	122 (35.56%)	0.08
IPAQ (MET minutes/week median [IQR])	2271 [2360]	1528 [2350]	1532 [2545]	0.84	1543 [2772]	1515 [2373]	0.99
Education level n (%)							
Left school age 14 or younger	421 (1.31%)	414 (1.31%)	5 (1.01%)	0.56	2 (1.31%)	3 (0.87%)	0.002
Left school age 15 or older	2260 (7.03%)	2198 (6.99%)	49 (9.9%)	<0.001	18 (11.84%)	28 (8.16%)	
High school diploma	4229 (13.16%)	4138 (13.16%)	56 (11.29%)		27 (17.76%)	32 (9.32%)	
Sixth form qualification	1820 (5.66%)	1758 (5.68%)	24 (4.84%)		5 (3.28%)	19 (5.53%)	
Professional qualification	8953 (27.87%)	8745 (27.82%)	143 (28.88%)		48 (31.57%)	95 (27.69%)	
Higher education University degree	14,438 (44.94%)	14,144 (45.01%)	218 (44.04%)	0.56	52 (34.21%)	166 (48.39%)	0.002
Alcohol intake n (%)	**16,658 (51.86%)**	**15,930 (50.59%)**	**268 (55.63%)**	<0.001	13 (8.55%)	15 (4.37%)	<0.001
Never	1,595 (4.96%)	1,505 (4.78%)	28 (5.65%)	0.32	20 (13.16%)	15 (4.37%)	0.19
Special occasions only	2,032 (6.32%)	1,993 (6.34%)	35 (7.07%)	0.21	31 (20.39%)	19 (5.53%)	0.94
1–2 times a month	993 (3.09%)	8133 (25.88%)	109 (22.02%)	<0.001	31 (20.39%)	78 (22.74%)	0.58
3–4 times a week	4,397 (13.68%)	8,896 (28.30%)	133 (26.86%)	0.21	36 (23.68%)	97 (28.27%)	0.08
Daily or almost daily	2,727 (8.49%)	8,096 (25.94%)	150 (30.30%)	0.66	52 (34.21%)	119 (34.69%)	0.02

TABLE 4.2: Average performance of the models for all AF patients adjusted by sex. The standard deviation is indicated in parenthesis

	ECG	**Radiomics**	**ECG + radiomics**
		All	
F1_score	0.82	0.71	0.81
Accuracy	0.84	0.74	0.81
Sensitivity	0.77	0.66	0.77
Specificity	0.91	0.82	0.86
AUC	0.86 ($\pm$0.04)	0.82 ($\pm$0.03)	0.87 ($\pm$0.04)
		Women	
F1_score	0.72	0.72	0.78
Accuracy	0.74	0.73	0.78
Sensitivity	0.68	0.69	0.79
Specificity	0.80	0.77	0.77
AUC	0.77 ($\pm$0.13)	0.81 ($\pm$0.09)	0.87 ($\pm$0.05)
		Men	
F1_score	0.84	0.73	0.82
Accuracy	0.85	0.75	0.82
Sensitivity	0.81	0.69	0.82
Specificity	0.89	0.80	0.82
AUC	0.88 ($\pm$0.05)	0.82 ($\pm$0.04)	0.89 ($\pm$0.06)

TABLE 4.3: Average performance of the models for all AF patients adjusted by sex and age only, and sex age and other comorbidities. The standard deviation is indicated in parenthesis.

	Adjusted by sex and age (p > 0.05)			**Adjusted by sex and comorbidities (p > 0.05)**		
	ECG	**Radiomics**	**ECG + radiomics**	**ECG**	**Radiomics**	**ECG + radiomics**
Women						
F1_score	0.72	0.73	0.8	0.71	0.73	0.77
Accuracy	0.75	0.75	0.8	0.73	0.75	0.78
Sensitivity	0.67	0.67	0.79	0.68	0.68	0.76
Specificity	0.83	0.83	0.81	0.78	0.82	0.8
AUC	0.79 (±0.09)	0.82 (±0.03)	0.88 (±0.07)	0.78 (±0.13)	0.83 (±0.04)	0.85 (±0.05)
Men						
F1_score	0.85	0.74	0.84	0.86	0.73	0.84
Accuracy	0.85	0.75	0.84	0.86	0.74	0.84
Sensitivity	0.79	0.7	0.81	0.82	0.69	0.81
Specificity	0.92	0.8	0.87	0.9	0.79	0.87
AUC	0.89 (±0.05)	0.83 (±0.06)	0.89 (±0.06)	0.89 (±0.04)	0.82 (±0.06)	0.89 (±0.05)

TABLE 4.4: Average performance of the models when the AF patients have a normal sinus rhythm and normal ECG. The standard deviation appears indicated in parenthesis.

	ECG	**Radiomics**	**ECG + radiomics**
		All	
F_score	0.49	0.55	0.6
Accuracy	0.53	0.53	0.6
Sensitivity	0.51	0.57	0.62
Specificity	0.56	0.49	0.57
AUC	0.54 ($\pm$0.11)	0.59 ($\pm$0.12)	0.61 ($\pm$0.08)
		Women	
F_score	0.60	0.6	0.68
Accuracy	0.54	0.59	0.66
Sensitivity	0.65	0.66	0.72
Specificity	0.44	0.52	0.6
AUC	0.54 ($\pm$0.23)	0.67 ($\pm$0.16)	0.72 ($\pm$0.15)
		Men	
F_score	0.49	0.58	0.5
Accuracy	0.49	0.58	0.54
Sensitivity	0.49	0.57	0.49
Specificity	0.51	0.6	0.59
AUC	0.45 ($\pm$0.14)	0.59 ($\pm$0.19)	0.56 ($\pm$0.19)

TABLE 4.5: Feature selection for AF for all participants for the electro-radiomics model.

Feature selection	Feature type	ROI	Phase	Importance
tachy				95.7189
brady				95.3946
'LA_shape_SurfaceArea_ED'	Shape	LA	ED	94.4181
'LA_shape_VoxelVolume_ED'	Shape	LA	ED	93.3183
'LA_shape_MeshVolume_ED'	Shape	LA	ED	93.0671
medianRR				92.9818
'LA_shape_Sphericity_ED'	Shape	LA	ED	92.2226
meanRR				91.1313
'LA_shape_SurfaceVolumeRatio_ED'	Shape	LA	ED	89.5061
DistCennS				86.7612
'LA_gldm_DependenceNonUniformity_ED'	Texture	LA	ED	86.5382
'LA_shape_Maximum2DDiameterColumn_ED'	Shape	LA	ED	86.0905
'LA_shape_Maximum2DDiameterSlice_ED'	Shape	LA	ED	85.0250
'LA_shape_Maximum3DDiameter_ED'	Shape	LA	ED	84.4272
'LA_firstorder_Energy_ED'	First-Order	LA	ED	84.0768
'LA_firstorder_TotalEnergy_ED'	First-Order	LA	ED	83.5177
'LA_shape_MajorAxisLength_ED'	Shape	LA	ED	83.3500
'LA_ngtdm_Strength_ED'	Texture	LA	ED	77.9337
'LA_shape_MinorAxisLength_ED'	Shape	LA	ED	77.4875
'LA_glrlm_GrayLevelNonUniformity_ED'	Texture	LA	ED	75.8153
'LA_glszm_GrayLevelNonUniformity_ED'	Texture	LA	ED	74.2248
'LA_ngtdm_Busyness_ED'	Texture	LA	ED	74.0887
DlvI2				73.8246
DlvI8				73.8246
'LA_glrlm_RunLengthNonUniformity_ED'	Texture	LA	ED	72.1410
'LA_ngtdm_Coarseness_ED'	Texture	LA	ED	69.0288
'LA_gldm_GrayLevelNonUniformity_ED'	Texture	LA	ED	67.3082
DlvI3				65.5733
DlvI9				65.5733

TABLE 4.6: Feature selection for AF for women for the electro-radiomics model.

Feature selection	Feature type	ROI	Phase	Importance
meanRR				23.8432
'LA_shape_Maximum2DDiameterColumn_ED'	Shape	LA	ED	23.2997
Cvlc6				23.2829
brady				22.6090
'LA_shape_Sphericity_ED'	Shape	LA	ED	22.1852
medianRR				21.1193
tachy				20.5118
'LA_shape_MinorAxisLength_ED'	Shape	LA	ED	20.4190
'LA_firstorder_TotalEnergy_ED'	First-Order	LA	ED	20.3422
'LA_shape_VoxelVolume_ED'	Shape	LA	ED	20.2966
'LA_shape_MeshVolume_ED'	Shape	LA	ED	20.2541
'LA_shape_SurfaceArea_ED'	Shape	LA	ED	20.2347
DlvI2				20.1606
DlvI8				20.1606
'LA_shape_SurfaceVolumeRatio_ED'	Shape	LA	ED	20.0550
'LA_firstorder_Energy_ED'	First-Order	LA	ED	19.5947
'LA_gldm_DependenceNonUniformity_ED'	Texture	LA	ED	19.3729
'LA_glrlm_RunLengthNonUniformity_ED'	Texture	LA	ED	19.2391
'LA_shape_MajorAxisLength_ED'	Shape	LA	ED	19.0635
'LA_glrlm_GrayLevelNonUniformity_ED'	Texture	LA	ED	18.6670
DistCennS				18.3854
'RA_shape_MajorAxisLength_ED'	Shape	RA	ED	17.3390
'LA_shape_Maximum2DDiameterColumn_ES'	Shape	LA	ES	17.3050
DlvI3				17.0507
DlvI9				17.0507
'LA_glszm_GrayLevelNonUniformity_ED'	Texture	LA	ED	16.8258
'LA_ngtdm_Coarseness_ED'	Texture	LA	ED	16.6224
'LA_ngtdm_Strength_ED'	Texture	LA	ED	16.3346
'RA_shape_Maximum2DDiameterSlice_ED'	Shape	RA	ED	15.9814
'RA_shape_Maximum3DDiameter_ED'	Shape	RA	ED	15.9814

TABLE 4.7: Feature selection for AF for men for the electro-radiomics model.

Feature selection	Feature type	ROI	Phase	Importance
'LA_shape_Sphericity_ED'	Shape	LA	ED	67.1634
'LA_shape_VoxelVolume_ED'	Shape	LA	ED	66.8851
'LA_shape_MeshVolume_ED'	Shape	LA	ED	66.6944
'LA_shape_SurfaceArea_ED'	Shape	LA	ED	65.4330
'LA_shape_Maximum2DDiameterSlice_ED'	Shape	LA	ED	64.5759
'LA_shape_Maximum3DDiameter_ED'	Shape	LA	ED	64.5759
tachy				64.0718
brady				63.9600
'LA_shape_MajorAxisLength_ED'	Shape	LA	ED	63.9363
'LA_gldm_DependenceNonUniformity_ED'	Texture	LA	ED	63.8068
Cvlc6				63.3610
medianRR				63.3370
'LA_firstorder_TotalEnergy_ED'	First-Order	LA	ED	61.8981
'LA_shape_SurfaceVolumeRatio_ED'	Shape	LA	ED	61.3160
'LA_firstorder_Energy_ED'	First-Order	LA	ED	60.8644
DistCennS				60.3365
meanRR				59.1486
'LA_shape_Maximum2DDiameterColumn_ED'	Shape	LA	ED	58.8432
'LA_glrlm_GrayLevelNonUniformity_ED'	Texture	LA	ED	58.6194
'LA_ngtdm_Busyness_ED'	Texture	LA	ED	56.8700
'LA_ngtdm_Coarseness_ED'	Texture	LA	ED	53.7045
'LA_shape_MinorAxisLength_ED'	Shape	LA	ED	53.6843
'LA_glrlm_RunLengthNonUniformity_ED'	Texture	LA	ED	52.9183
'LA_gldm_GrayLevelNonUniformity_ED'	Texture	LA	ED	51.8700
'RA_shape_Sphericity_ED'	Shape	RA	ED	50.8045
'LA_ngtdm_Strength_ED'	Texture	LA	ED	50.1410
'RA_shape_MajorAxisLength_ED'	Shape	RA	ED	49.2244
'LA_glszm_GrayLevelNonUniformity_ED'	Texture	LA	ED	49.1139
pNN50				48.8595
'RA_shape_VoxelVolume_ED'	Shape	RA	ED	48.5482

TABLE 4.8: Feature selection for AF for women for the electro-radiomics model in all partitions in the nested-cross validation indicating the number of repetitions in each feature in each different partition.

Feature selection for AF in women in electro-radiomics model in all nested-cross validation partitions	Feature type	ROI	Phase	Number of Iterations
'meanRR'	Shape	LA	ED	10
'LA_shape_Maximum2DDiameterColumn_ED'	Shape	LA	ED	10
'crvdi'	Shape	LA	ED	10
'blah6'	Shape	LA	ED	10
'tachy'	Shape	LA	ED	10
'medianRR'	Shape	LA	ED	8
'LA_shape_MinorAxisLength_ED'	Shape	LA	ED	10
'LA_sfirstorder_TotalEnergy_ED'	First-Order	LA	ED	10
'LA_shape_VoxelVolume_ED'	Shape	LA	ED	10
'LA_shape_MeshVolume_ED'	Shape	LA	ED	10
'LA_shape_SurfaceArea_ED'	Shape	LA	ED	10
'Dlvl2'	Shape	LA	ED	10
'Dlvl8'	Shape	LA	ED	9
'LA_shape_SurfaceVolumeRatio_ED'	Shape	LA	ED	10
'LA_sfirstorder_Energy_ED'	First-Order	LA	ED	10
'LA_gldm_DependenceNonUniformity_ED'	Texture	LA	ED	10
'LA_glrlm_RunLengthNonUniformity_ED'	Texture	LA	ED	10
'LA_shape_MajorAxisLength_ED'	Shape	LA	ED	10
'LA_glrlm_GrayLevelNonUniformity_ED'	Texture	LA	ED	8
'DistCent5'	Texture	LA	ED	5
'RA_shape_MajorAxisLength_ED'	Shape	RA	ED	10
'LA_shape_Maximum2DDiameterColumn_ES'	Shape	LA	ES	9
'Dlvl9'	Texture	LA	ED	7
'Dlvl15'	Texture	LA	ED	5
'LA_glzsm_GrayLevelNonUniformity_ED'	Texture	LA	ED	5
'LA_ngtdm_Coarseness_ED'	Texture	LA	ED	7
'LA_ngtdm_Strength_ED'	Texture	LA	ED	8
'RA_shape_Maximum2DDiameterSlice_ED'	Shape	RA	ED	9
'RA_shape_Maximum2DDiameter_ED'	Shape	RA	ED	8
'LA_shape_Maximum3DDiameter_ED'	Shape	LA	ED	8
'LA_ngtdm_Busyness_ED'	Texture	LA	ED	2
'LA_shape_MajorAxisLength_ES'	Shape	LA	ES	4
'RA_shape_Maximum2DDiameterColumn_ED'	Shape	RA	ED	2
'LA_shape_Maximum2DDiameterSlice_ES'	Shape	LA	ES	1
'edgesbin2n1'	Shape	LA	ES	1
'LA_shape_Sphericity_ES'	Shape	LA	ES	1
'LA_shape_Maximum3DDiameter_ES'	Shape	LA	ES	1

Chapter 5

Prediction of incident cardiovascular events using machine learning and CMR radiomics

5.1 Introduction

Cardiovascular disease (CVD) is the most common cause of morbidity and mortality worldwide (Aparicio, Benjamin, Callaway, et al., 2021). Accurate risk stratification has a key role in ensuring appropriately targeted preventive strategies. Existing disease prediction algorithms reliant on demographic and clinical variables have been proposed for prediction of selected major CVDs (Himmelreich, Veelers, Lucassen, et al., 2020; Sahle et al., 2017; Flueckiger et al., 2018). Cardiovascular magnetic resonance (CMR) is the reference modality for quantification of cardiovascular structure and function and is widely used in clinical and research settings (Schulz-Menger, Bluemke, Bremerich, et al., 2020). The rich phenotyping provided by CMR allows characterisation of pre-clinical organ-level remodelling (Sekaran, Crowley, Souza, et al., 2017). Therefore, there is growing interest in the integration of imaging biomarkers into CVD prediction algorithms (Leiner, Rueckert, Suinesiaputra, et al., 2019). However, existing approaches to CMR image analysis are limited to simplistic volumetric measurements or qualitative assessments (Raisi-Estabragh et al., 2020b). These conventional CMR metrics (left ventricular ejection fraction or maximal end-diastolic wall thickness) have shown potential for the early detection of cardiac deterioration and the characterisation of subclinical diseases (Petersen et al., 2017). Radiomics is a quantitative image analysis method, which allows extraction of highly detailed information about ventricular shape and myocardial character, thereby providing new information from existing standard of care images (Raisi-Estabragh et al., 2021a). Radiomics features may be used as predictor variables in clinical models, often developed using machine learning (ML) methods. A key advantage of radiomics analysis over unsupervised ML algorithms is the interpretability of the models; that is, the radiomics features can be traced back to the heart's morphological and tissue level alterations (Kolossváry et al., 2018). CMR radiomics is in the early stages of its development and thus far existing work has largely focused on demonstrating feasibility of the technique for disease discrimination (Neisius et al., 2020; Cetin et al., 2020). The CMR radiomics analysis is more mature within oncology and in this context, radiomics models have been successful for prediction of incident health events (Bera, Braman, Gupta, et al., 2021). The value of CMR radiomics models for incident CVD prediction has not been previously studied. In this chapter, we aim

to evaluate the feasibility and clinical utility of CMR radiomics for the prediction of four key incident CVDs: atrial fibrillation (AF), heart failure (HF), myocardial infarction (MI), stroke. To evaluate the incremental value of CMR radiomics over existing approaches, we hierarchically built supervised ML models incorporating traditional vascular risk factors (VRFs) and conventional CMR metrics.

5.2 Methods

5.2.1 Population and Setting

The UK Biobank (UKB) is an extensive cohort study that comprises over half a million individuals recruited between 2006 and 2010. The UKB provides a rich source of health data including comprehensive medical history, risk factors, biomarkers, and physical measurements (*UK Biobank: Protocol for a large-scale prospective epidemiological resource* 2007). The UKB imaging study commenced in 2015 and aims to scan 100,000 participants from the original dataset, and includes CMR (Littlejohns, Holliday, Gibson, et al., 2020). Participants' incident outcomes are tracked through the national data sources, including Hospital Episode Statistics (HES) and death registers to provide continuous longitudinal follow-up (Raisi-Estabragh et al., 2020b).

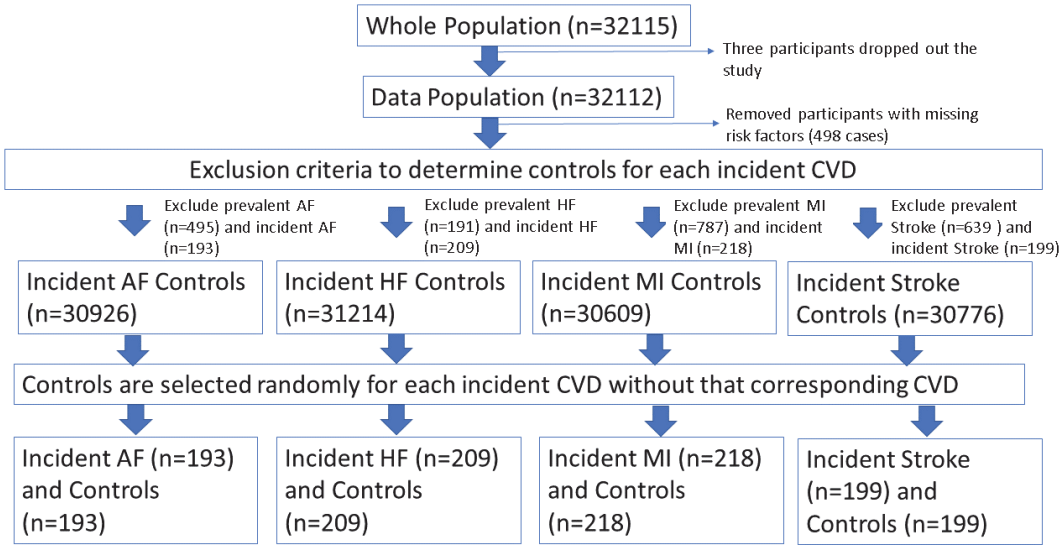

FIGURE 5.1: Definition of the study sample. Abbreviations: AF, atrial fibrillation; HF, heart failure; MI, myocardial infarction

5.2.2 Ethical approval

This study complies with the Declaration of Helsinki; the work was covered by the ethical approval for UKB studies from the National Health Service (NHS) National Research Ethics Service on 17th June 2011 (Ref 11/NW/0382) and extended on 18th June 2021 (Ref 21/NW/0157) with written informed consent obtained from all participants.

5.2.3 Definition of the Study Sample

From the UK Biobank, most of the particpants starts with a healthy condition developing diseases along the time. We identified individuals who experienced incident AF (N=193), HF (N=209), MI (N=218) or stroke (N=199) until the censoring date, 28th February 2021. Outcomes were ascertained through linked HES data with diseases defined according to the standardised international classification of disease (ICD) codes (Supplementary Table A.1). Individuals with the outcome of interest at imaging were not included. We selected comparator groups for each outcome (AF, HF, MI, stroke) comprising an equal number of randomly selected subjects who did not develop the outcome of interest during follow-up to eliminate class imbalance bias (Figure 5.1).

5.2.4 Vascular risk factors

We selected VRFs based on biological plausibility and reported associations in the literature, including the following variables: age, sex, body mass index, material deprivation, education, current smoking, alcohol intake, physical exercise, high cholesterol, diabetes mellitus, and hypertension (Visseren, Mach, Smulders, et al., 2021). The definition used for the ascertainment of high cholesterol, diabetes mellitus and hypertension is given in Supplementary Table A.1.

5.2.5 Conventional CMR measures

All CMR scans were completed in dedicated UKB imaging centres using 1.5 Tesla scanners (MAGNETOM Aera, Syngo Platform VD13A, Siemens Healthcare, Erlangen, Germany) under pre-defined acquisition protocols (Petersen, Matthews, Francis, et al., 2016). Standard long-axis images and a short axis stack covering both ventricles from base to apex were captured using balanced steady-state free precession sequence (Petersen, Matthews, Francis, et al., 2016. CMR examinations of the first 5,065 UKB participants were assessed manually using CVI42 post-processing software (Version 5.1.1, Circle Cardiovascular Imaging Inc., Calgary, Canada) (Petersen et al., 2017). This analysis set was used to develop a fully automated quality controlled pipeline and extract the contours for the 32,121 CMR studies (Bai, Sinclair, Tarroni, et al., 2018; Attar, Pereañez, Gooya, et al., 2019). The following conventional CMR indices were considered during our analysis: LV end-diastolic volume (LVEDV), LV end-systolic volume (LVESV), RV end-diastolic volume (RVEDV), RV end-systolic volume (RVESV), LV stroke volume (LVSV), RV stroke volume (RVSV), LV ejection fraction (LVEF), RV ejection fraction (RVEF), LV mass (LVM). For ease of interpretation, we gave LV and RV ventricular volumes and masses in body surface area standardized format.

5.2.6 Background of CMR Radiomics

CMR radiomics is a novel image analysis technique permitting the computation of multiple indices of shape and texture (Raisi-Estabragh et al., 2020b). Three classes of features are extracted: shape, first-order and texture-based features. First-order features are histogram-based and related to the distribution of the grey level values in the tissue. Shape features describe geometrical properties of the organ, such as volume, diameter, minor/major axis and sphericity. Texture features are derived from images that encode the global texture information, using mathematical formulae based on the spatial arrangement of pixels. Radiomic features can appreciate the

heart's complexity in detail by revealing patterns invisible to the naked eye. Thus, it furnishes a nearly limitless supply of imaging biomarkers with potential added value over conventional CMR metrics. However, caution should be taken regarding the reproducibility of different features (Raisi-Estabragh et al., 2021b).

5.2.7 Radiomics Feature extraction

Radiomics workflow is illustrated in Figure 5.2. We used the short axis stack contours for conventional image analysis to define three regions of interest (ROI) for radiomics analysis: RV cavity, LV cavity, LV myocardium in ES and ED phases. We calculated these features from the 3D volumes of the ROIs. The open-source PyRadiomics platform (version 2.2.0.) was adopted to extract Radiomics features. The grey value discretisation was performed using a binwidth of 25 to pull the intensity-based and texture radiomics features. A total of 262 radiomics features were included from each CMR study (LV shape n=26, RV shape n=26, MYO shape n=26, LV myocardium first-order n=36, LV myocardium texture n=148).

5.2.8 Radiomics feature selection

Sequential Feature Forward Selection (SFFS) algorithm (Kudo and Sklansky, 2000) was applied to select the most relevant subset of features to improve computational efficiency or reduce the model's generalisation error . SFFS starts with zero feature and finds the one that maximises a score when an estimator is trained on this single feature. This procedure is repeated until the total number of features is reached or there is no improvement. The score selected was given from a Support Vector Machine model (SVM) (Noble, 2006; Chandra and Bedi, 2021). The objective of SVM is to maximise the margin between cases and controls, which is defined as the distance between the separating hyperplane (decision boundary) and the training samples that are closest to this hyperplane, as shown in Figure 5.3.

5.2.9 Statistical analysis

Data analysis and graph visualisation were performed using Matlab (version 2001b), R (version 4.1.2, R package: gplots package heatmap.2 function) and RStudio (version 2022.02.3) programs. We assessed the intercorrelation between conventional CMR metrics and radiomics features using Pearson's correlation. Due to the large number of radomics features, we grouped the inter-correlated variables into six clusters using hierarchical clustering, as per our previous publication (Raisi-Estabragh et al., 2021b). We created hierarchical models to understand the influence of vascular risk factors (VRF), conventional CMR indices and radiomics features and their integrated use in the prediction of incident CVDs (AF, HF, MI and stroke). The first three models assess the performance of VRF, conventional CMR indices, and CMR radiomics separately. Next, we combined categories as follows: VRF-CMR indices, VRF-radiomics, and CMR indices-radiomics. Finally, we merged all three components into an integrative model: VRF-CMR indices-radiomics. The summary of the process is shown in Figure 2. Training data sets are used to train and tune the parameters of the model then a separate testing set is used to assess the performance of the model to see that the model built is able to generalise to unseen data. SVM is used for classification. We chose SVM due to its properties: good performance in real-world applications, computationally efficient, robust in high dimension, and sound in theoretical foundations. In order to tune the SVM parameters brute force exhaustive

search also known as greedy optimisation is used. The model is then trained with the parameters optimised. This procedure of tuning and training is performed five times each with different partitions of training (80%) and test (20%) samples to reduce overfitting. The average error of the testing folds determines the performance of the model. We determined model performance using receiver operating characteristic (ROC) curve and area under the curve (AUC) scores. To assess the model accuracy, the mean accuracy, sensitivity, specificity, and AUC are reported. Welch's t-test and Chi-Squared test were used for group-wise comparisons for continous and categorical values, respectively.

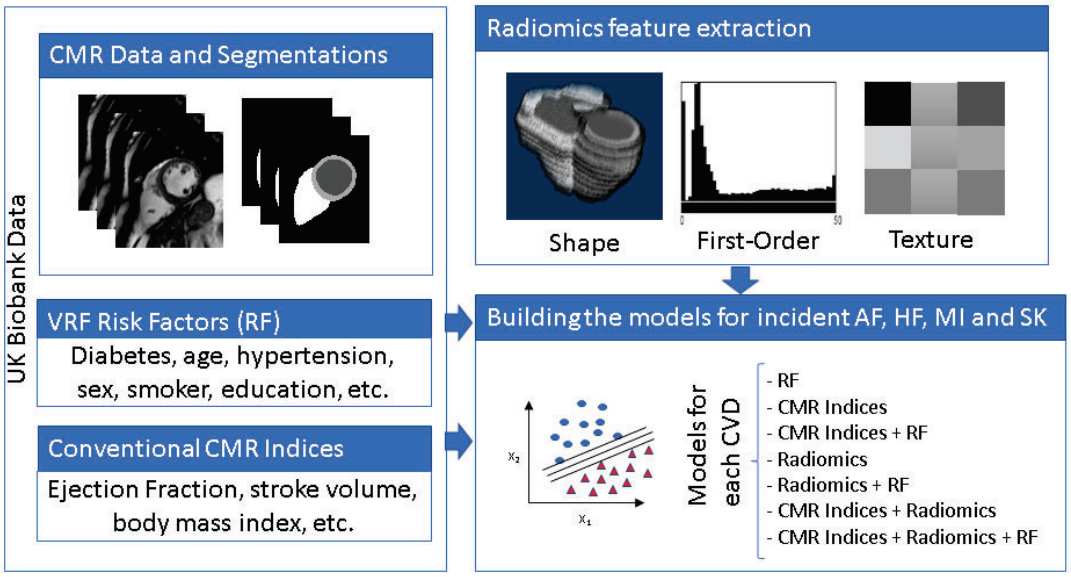

FIGURE 5.2: Flowchart to create the models for incident CVD. Abbreviations: CMR,cardiac magnetic resonance imaging; CVD,cardiovascular disease; VRF, vascular risk factor

5.3 Results

5.3.1 Baseline characteristics

The subjects' characteristics are summarised in Tables 5.1 and 5.2. CMR data was available for 32,121 UKB participants. For the whole imaging set, the average age was 63.3(±7.5) years, and the sample included 51.9% women. Over 3.7(±1.3) years of prospective follow-up, 193 participants had incident AF, 209 incident HF, 218 incident MI, and 199 incident stroke. Men were more likely to experience all incident CVDs considered. As expected, individuals who experienced incident CVD events had a greater overall risk factor burden. Conventional CMR metrics differed among at-risk groups and the whole imaging set: participants, who later developed AF, HF, MI or stroke had on average higher LVMi ($p<0.05$). The HF group had larger LVEDVi, and reduced LVEF ($p<0.05$) compared to the whole imaging set.

5.3.2 Correlation between CMR metrics and radiomics features

Figure 5.4 shows the correlation pattern between conventional CMR metrics and the imaging set's radiomics features. Overall, size radiomics features showed the strongest correlation with conventional metrics. Moreover, some parameters from

TABLE 5.1: Baseline characteristics

Characteristics	**Whole imaging set** (n=32121)	**Incident Atrial Fibrillation** (n=193)	**Incident Heart Failure** (n=209)	**Incident Myocardial Infarction** (n=218)	**Incident Stroke** (n=199)
Age mean (std)	63.3 (±7.5)	66.9 (±6.4)	68.7 (±6.2)	66.2(±7.3)	67(±8)
Female sex n(%)	16658(51.7%)	59 (30.6%)	72 (34.4%)	66 (30.3%)	75 (37.7%)
Townsend deprivation index median(IQR)	-2.0 (3.3)	-2.6 (2.9)	-2.6 (2.9)	-2.5 (3.8)	-3.0 (2.5)
Body mass index mean (kg/m2)	26.6 (±4.4)	27.0 (±4.4)	28.3 (±4.9)	27.7 (±4.0)	27.0 (±3.5)
Current smoker n (%)	2032 (6.3%)	13 (6.7%)*	15 (7.2%)*	20 (9.2%)*	11 (5.5%)*
Diabetes status n (%)	993(3.1%)	10 (5.2%)*	15 (7.2%)	11 (5.0%)*	8 (4.0%)*
Hypertension status n (%)	4397(13.7%)	54 (28.0%)	79 (37.8%)	60 (27.5%)	42 (21.1%)
High cholesterol status n (%)	7272(22.6%)	49 (25.4%)*	84 (40.2%)	64 (29.4%)	61 (30.7%)
IPAQ (MET minutes/week) median [IQR]	1528 [2360]	1519 [2892]*	1470 [2574]*	1281 [2262]*	1706 [2419]*
Education level n (%)					
Left school age 14 or younger	421 (1.3%)	2 (1.0%)*	3 (1.4%)	2 (0.9%)	3 (1.5%)**
Left school age 15 or older	2260 (7.0%)	13 (6.7%)*	33 (15.8%)	26 (11.9%)	15 (7.5%)
High school diploma	4229 (13.2%)	32 (16.6%)*	37 (17.7%)	39 (17.9%)	24 (12.1%)
Sixth form qualification	1820 (5.7%)	11 (5.7%)*	11 (5.3%)	13 (6.0%)	12 (6.0%)
Professional qualification	8953 (27.9%)	64 (33.2%)*	57 (27.3%)	68 (31.2%)	52 (26.1%)
Higher education	14438 (44.9%)	71 (36.8%)*	68 (32.5%)	70 (32.1%)	93 (46.7%)

TABLE 5.2: Baseline characteristics

Characteristics	Whole imaging set (n=32121)	Incident Atrial Fibrillation (n=193)	Incident Heart Failure (n=209)	Incident Myocardial Infarction (n=218)	Incident Stroke (n=199)
Alcohol intake n(%)					
Never	1547 (4.8%)	14 (7.3%)	15 (7.2%)	14 (6.4%)	8 (4.0%)
Special occasions only	2646 (8.2%)	9 (4.7%)	20 (9.2%)	15 (7.5%)	
1–3 times a month	3452 (10.7%)	17 (8.8%)	26 (12.4%)	26 (11.9%)	23 (11.6%)
1–2 times a week	8224 (25.8%)	37 (19.2%)	55 (26.3%)	51 (23.4%)	43 (21.6%)
3–4 times a week	5908 (18.3%)	45 (23.3%)	58 (27.8%)	61 (28.0%)	55 (27.6%)
Daily or almost daily	7098 (22.1%)	51 (26.4%)	45 (21.5%)	46 (21.1%)	55 (27.6%)
CMR Indices mean(± std)					
LVEDVi, ml/m2	78.5 (±14.2)	84.2 (±21.7)	88.0 (±24.9)	80.2 (±13.9)*	81.3 (±17.3)*
LVESVi, ml/m2	32.0 (±8.8)	36.2 (±16.1)	42.6 (±21.0)	33.7 (±10.3)	35.0 (±12.0)
LVSVi, ml/m2	46.5 (±8.5)	48.0 (±11.7)	45.5 (±11.0)	46.5 (±8.4)*	46.4 (±9.1)*
LVMi, g/m2	45.6 (±8.9)	50.8 (±11.8)	53.7 (±14.7)	49.8 (±9.4)	49.8 (±10.6)
LVEF, %	59.5 (±6.2)	57.8 (±8.5)*	53.2 (±10.3)	58.5 (±7.6)*	57.6 (±7.0)
RVEDVi, ml/m2	82.9 (±15.5)	86.8 (±17.9)	82.6 (±17.8)*	82.6 (±15.0)*	83.7 (±15.1)*
RVESVi, ml/m2	35.7 (±9.4)	38.9 (±11.0)	37.7 (±11.7)	36.1 (±9.4)*	36.9 (±9.4)*
RVSVi, ml/m2	47.2 (±8.9)	47.9 (±10.6)*	44.9 (±10.2)	46.5 (±8.8)*	46.9 (±8.4)*
RVEF, %	57.2 (±6.2)	55.5 (±7.1)	54.7 (±7.9)	56.6 (±6.5)*	56.3 (±5.8)

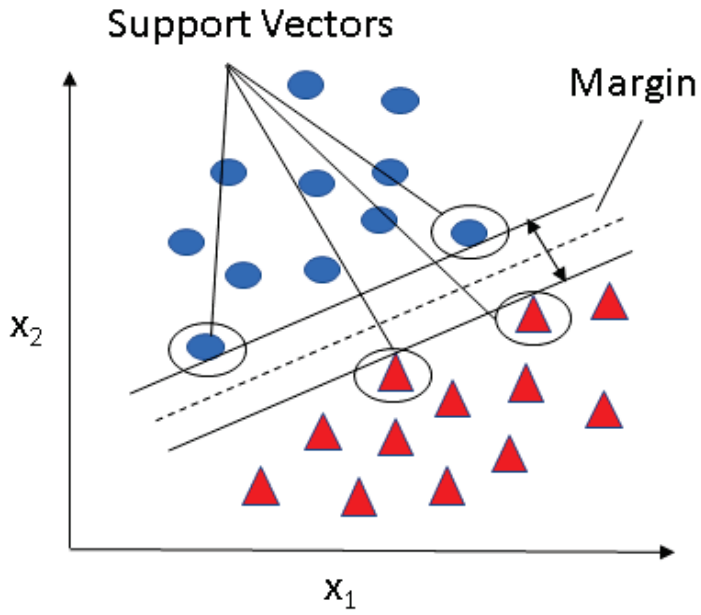

FIGURE 5.3: SVM process of maximising the margin. The objective of the support vector machine model is to maximise the margin between cases and controls, which is defined as the distance between the separating hyperplane (decision boundary) and the training samples that are closest to this hyperplane, which is the so-called support vectors (marked with circles)

the local uniformity and shape groups also correlated with conventional metrics. Contrary to that, the majority of global intensity, local dimness and global variance features showed inconsistent correlation patterns with CMR indices. Thus, although there is some overlap of conventional and radiomics CMR metrics, there are many areas where radiomics features provide new information.

5.3.3 Identification of metrics for each CVD outcome

The features selected for each model are shown in Tables 5.4, 5.5 and 5.6. Feature importance is shown as the accuracy given by the SVM algorithm for each standalone feature. The SFFS algorithm chose hypertension for all predictive models, its standalone accuracy was similar among incident outcomes, except for stroke which was lower (Accuracy: AF vs HF vs MI vs Stroke – 0.59 vs 0.62 vs 0.58 vs 0.55). Sex was included in all but the HF models. LVM and LVSV were the two conventional features consistently selected by the SFFS. The accuracy of LVM alone was higher in all models compared to LVSV. The identified radiomics signatures for each incident outcome are depicted in Appendix Supplementary Tables A.2, A.3, A.4, A.5 and A.6. Overall, ventricular shape and myocardial texture feature dominated all models and there was only a marginal role for first-order features. Indeed, HF and MI prediction models included only shape and texture features. Radiomics features derived from the LV blood pool and myocardium dominated all prediction models. Notably, when conventional CMR metrics and radiomics features were included alongside each other, the latter were selected more frequently than the former. Shape features depicting the "Maximum diameter" presented the most discriminative power

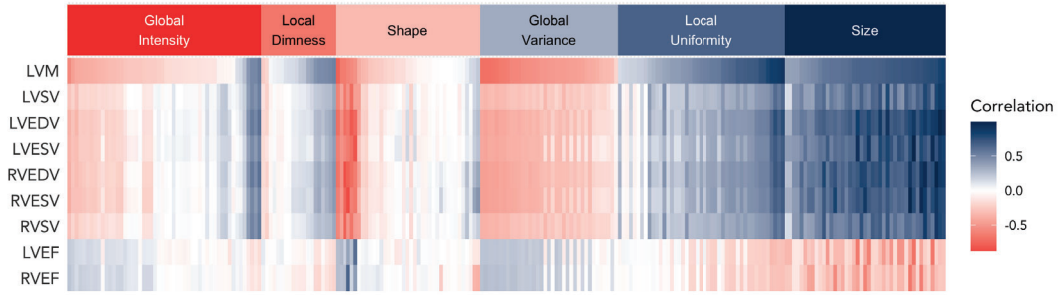

FIGURE 5.4: Correlation matrix of conventional CMR indices vs radiomics features in the whole sample. The correlation matrix illustrates correlation of each radiomics feature on the x-axis with the conventional CMR metrics indicated on the y-axis. Due to the large number of radiomics features, we grouped the inter-correlated variables into six clusters using hierarchical clustering using Ward's algorithm. Abbreviations: LVEDV, left ventricular end-diastolic volume; LVEF, left ventricular ejection fraction; LVESV, left ventricular end-systolic volume; LVM, left ventricular mass; RVEDV, right ventricular end-diastolic volume; RVESV, right ventricular end-systolic volume; RVSV, right ventricular stroke volume

in AF, alongside texture features of non-uniformity. In the HF model, shape features (maximum diameter, minor axis and volume) presented the greatest selective power, whilst in the MI model, the texture features, such as coarseness or large area emphasis, were more prominent.

Abbreviations: ROI, region of interest; SVM model alone: support vector machine model performance showing the mean and standard deviation using each radiomic feature individually; LV, left-ventricle; RV, right-ventricle; MYO, left ventricle myocardium; ED, end-diastolic.

5.3.4 The degree of discrimination achieved for each incident CVD

Results from the hierarchical models are summarized in Table 5.3. The average error of the testing folds determines the performance of the model. Radiomics models alone yielded slightly better discrimination and higher sensitivity than VRFs or conventional CMR models in each outcome. AF and HF prediction models performed generally better than MI and stroke prediction models. The addition of radiomics features improved the performance of VRF models in AF (AUC: 0.67 vs 0.76) and HF (AUC: 0.73 vs 0.83) prediction (Figure 5.5). Moreover, VRFs and radiomics features' combination reached better performance than VRFs and conventional CMR metrics in AF, HF and stroke prediction models.We reached the best performance in the incident AF prediction model combining VRFs, CMR indices, and radiomics features (Table 5.3). In Table A.7, we have added an additional experiment defining the healthy controls as subjects not having any cardiovascular disease or stroke at the baseline visit and during follow-up to see if the models behave in the same way. The results followed the same pattern for all the models except in the sensitivity which was lower. Additionally, the models stabilized with 40 features in the univariate feature selection. We could conclude that the performance of our model is rather similar regardless of the comparator groups, suggesting that the patterns we pick up are stable.

5.4 Discussion

In this study, we demonstrate the feasibility of CMR derived radiomics features to predict incident AF, HF, MI, and stroke. Additionally, using hierarchically built SVM models, we demonstrate the incremental value of CMR radiomics features for risk prediction over VRFs and conventional CMR metrics.

TABLE 5.3: Performance metrics for different diagnostic methods

	VRF	CMR	Radiomics	VRF + CMR	VRF + Radiomics	CMR + Radiomics	VRF + CMR + Radiomics
AF							
Accuracy	0.67(±0.03)	0.66(±0.03)	0.68(±0.05)	0.67(±0.04)	0.69(±0.06)	0.70(±0.07)	0.71(±0.08)
Sensitivity	0.69(±0.04)	0.68(±0.02)	0.77(±0.06)	0.68(±0.1)	0.73(±0.07)	0.76(±0.1)	0.72(±0.1)
Specificity	0.64(±0.05)	0.63(±0.09)	0.60(±0.06)	0.64(±0.05)	0.70(±0.08)	0.66(±0.03)	0.70(±0.08)
AUC	0.67(±0.05)	0.68(±0.04)	0.73(±0.06)	0.67(±0.06)	0.76(±0.06)	0.73(±0.07)	0.76(±0.07)
HF							
Accuracy	0.66(±0.03)	0.70(±0.02)	0.71(±0.03)	0.74(±0.02)	0.77(±0.02)	0.70(±0.06)	0.79(±0.02)
Sensitivity	0.63(±0.04)	0.61(±0.01)	0.82(±0.06)	0.80(±0.06)	0.74(±0.04)	0.63(±0.08)	0.73(±0.04)
Specificity	0.69(±0.06)	0.82(±0.05)	0.65(±0.05)	0.66(±0.06)	0.79(±0.04)	0.75(±0.1)	0.85(±0.03)
AUC	0.73(±0.03)	0.74(±0.02)	0.75(±0.02)	0.82(±0.03)	0.83(±0.03)	0.76(±0.8)	0.84(±0.02)
MI							
Accuracy	0.67(±0.02)	0.67(±0.02)	0.70(±0.06)	0.69(±0.01)	0.67(±0.05)	0.67(±0.05)	0.71(±0.04)
Sensitivity	0.69(±0.06)	0.58(±0.08)	0.75(±0.05)	0.70(±0.05)	0.69(±0.04)	0.64(±0.07)	0.76(±0.05)
Specificity	0.58(±0.03)	0.75(±0.04)	0.64(±0.1)	0.66(±0.04)	0.66(±0.06)	0.73(±0.08)	0.65(±0.05)
AUC	0.70(±0.03)	0.73(±0.04)	0.75(±0.04)	0.73(±0.03)	0.72(±0.04)	0.71(±0.04)	0.76(±0.04)
Stroke							
Accuracy	0.58(±0.03)	0.61(±0.01)	0.64(±0.03)	0.65(±0.04)	0.63(±0.03)	0.64(±0.03)	0.64(±0.03)
Sensitivity	0.63(±0.03)	0.60(±0.04)	0.81(±0.05)	0.61(±0.02)	0.51(±0.07)	0.81(±0.05)	0.74(±0.06)
Specificity	0.52(±0.03)	0.62(±0.06)	0.45(±0.03)	0.69(±0.03)	0.74(±0.04)	0.45(±0.03)	0.64(±0.03)
AUC	0.58(±0.02)	0.65(±0.03)	0.68(±0.04)	0.61(±0.04)	0.63(±0.05)	0.68(±0.04)	0.63(±0.05)

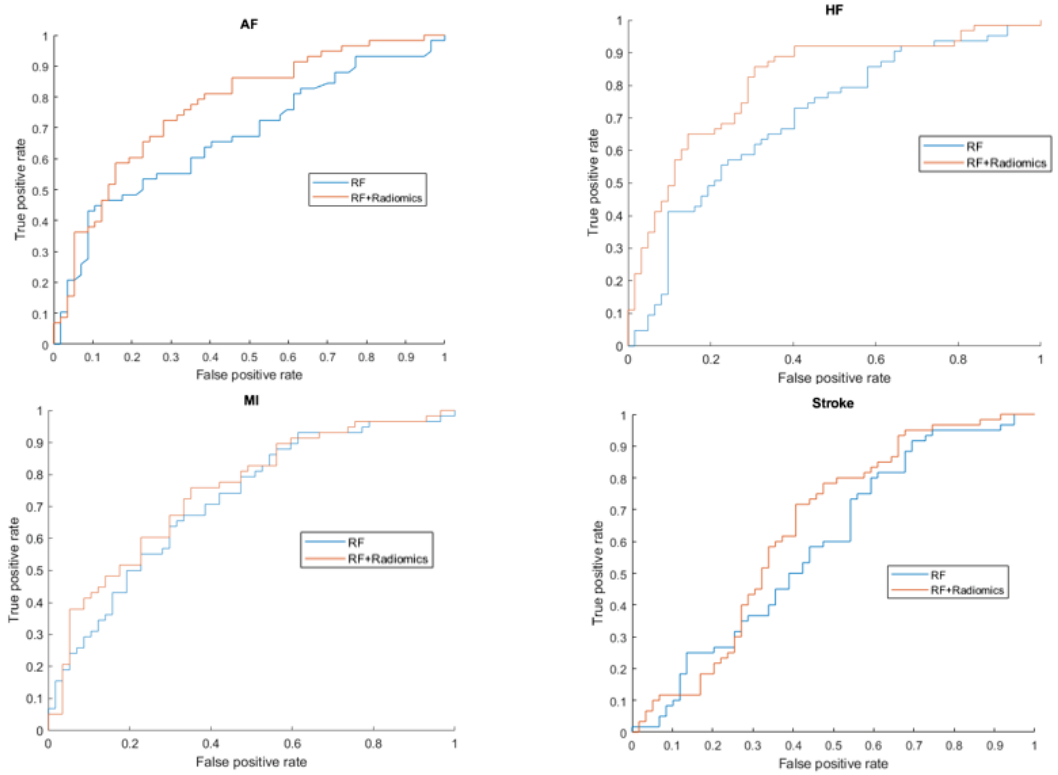

FIGURE 5.5: ROC curves showing the discriminative power of vascular risk factors alone and the combination of vascular risk factors and radiomics feature in all incident outcome prediction model. The combination of vascular risk factors (VRFs) and radiomics features (orange) reached better performance in the prediction of AF and HF compared to VRF alone (blue) ($p < 0.05$). Abbreviations: AF, atrial fibrillation; HF, heart failure; MI, myocardial infarction

5.4.1 Comparison with existing literature

To the best of our knowledge, this is the first study to demonstrate the value of CMR radiomics models for incident CVD prediction. Previous research supports the utility of CMR radiomics in the differential diagnosis of left ventricular hypertrophy (Schofield et al., 2019), especially the diagnosis of HCM (Neisius et al., 2020; Baeßler, Mannil, Maintz, et al., 2018; Antonopoulos, Boutsikou, Simantiris, et al., 2021). Cetin et al., 2020 have shown the technique's potential to identify imaging signatures associated with cardiovascular risk factors such as diabetes or hypertension . Furthermore, Raisi-Estabragh et al. demonstrated the independent associations of CMR phenotypes with sex, age, and important VRFs (Raisi-Estabragh et al., 2021b). Recently, Ma et al. concluded that a non-contrast T1 map-based radiomics nomogram is suitable for predicting major adverse cardiac events in patients with acute MI (Ma, Ma, Wang, et al., 2021). We built hierarchical models to test the utility and added benefit of including radiomics features in predicting AF, HF, MI and stroke using the SFFS algorithm. Not surprisingly, hypertension proved a crucial predisposing factor linked to all considered outcomes. This finding is consistent with the overwhelming evidence showing that among all risk factors for CVD, hypertension is associated with the strongest causal link to adverse outcomes (Schnabel, Sullivan, Levy, et al., 2008; Gosmanova, Mikkelsen, Molnar, et al., 2016; Fuchs and Whelton, 2020; Ekundayo et al., 2013; Rathore, 2018). Sex was selected for inclusion in all predictive models, except for HF, a finding that is in line with the results from major epidemiological studies (Lloyd-Jones, Larson, Leip, et al., 2002; Bleumink, Knetsch, Sturkenboom, et al., 2004) showing that the lifetime risk of HF is comparable among males and females. Of note, we did not differentiate subgroups of HF, which clearly show sex-specific differences as emphasised by Lam, Arnott, Beale, et al., 2019. Left ventricular hypertrophy (most commonly assessed by LVM increase) is a remarkable prognostic marker that incorporates a broad range of pathologies, such as hypertrophic and infiltrative cardiomyopathies, although it is most commonly caused by chronic pressure and volume overload (Stewart, Lavie, Shah, et al., 2018). Early studies have recognised increased LVM as a risk factor for stroke in the Framingham Heart

Study (Bikkina et al., 1995). LVM has been widely utilised ever since due to its ability to predict a variety of clinical outcomes (Stewart, Lavie, Shah, et al., 2018). Whilst conventional metrics quantify LVM according to mass or wall thickness, radiomics analysis can additionally quantify the distribution and pattern of myocardial signal intensities within the LV myocardium. As such, radiomics features extracted from the myocardium may provide more granular distinction of health and disease in comparison to conventional CMR indices where, rather crudely, the single most discriminatory feature for all risk factors was higher LVM (Cetin et al., 2020). Indeed, Schofield et al. showed that texture radiomics features derived from bSSFP sequences can differentiate between the aetiologies of LV hypertrophy (Schofield et al., 2019). These findings suggest that radiomics has the capability to enrich risk information beyond the limits of LVM. In our study, texture features were identified as the most defining model predictors, highlighting the clinical relevance of these metrics. Finally, we illustrated that radiomics features derived from CMR could provide incremental discriminative value over VRFs and CMR indices in the prediction of incident AF and HF. The HF model showed the most robust improvement with the addition of radiomics features, whilst stroke prediction showed only a slight improvement in the hierarchical models. This might be partially due to the aetiology: diseases such as dilated cardiomyopathy (the most common non-ischaemic cause of HF (Antonopoulos, Boutsikou, Simantiris, et al., 2021)) that primarily affect the global muscular structure of the heart may be better captured by CMR radiomics. In contrast, MI typically comprises more focal areas of myocardial injury and stroke is a primary cerebral illness.

5.4.2 Clinical interpretation of radiomics findings

Shape features and texture radiomics features presented the most discriminative value in AF prediction models. The most prominent shape feature was the maximum diameters of the LV and the ventricular wall in different phases of the cardiac cycle. This refers to the notion that the adverse remodelling of the heart described by larger chamber sizes and hypertrophy predispose AF. Alterations of the non-uniformity levels ("dependence non-uniformity" and "gray level non-uniformity") are referring to changes in the heterogeneity of intensity values, which might reflect on the adverse changes in tissue composition of the myocardial structure. Similarly, "large area low gray level emphasis" suggests larger myocardial regions with low signal intensity (dimmer) pixels. Indeed, LV diastolic dysfunction has been linked to an increased risk of AF in the general population (Kim, Shim, Park, et al., 2016), and more recently Tian et al. demonstrated the association between adverse LV remodelling and AF among HCM patients (Tian, Cui, Yang, et al., 2018). In the HF models, shape features, derived from the myocardium, LV and RV demonstrated strong discriminatory value. This can be explained by adverse and often biventricular remodelling that characterises HF patients. Our results suggested that apart from the diameter of a given slice, the elongation of the heart (depicted by "minor axis") also provide additional information.

5.4.3 Limitations

Although our analysis is performed with different partitions of data to have a model independent to the samples by minimising the case of over-fitting, the model might still be biased to the participants obtained in the UKB. In this proof-of-concept study, we limited our investigations to LV and RV metrics derived from bSSFP images. The clinical utility of this proof-of-concept study is limited in its current state: 1) CMR is not a routine examination 2) CMR should not be performed for the sole purpose of risk stratification. However, we believe it is reasonable to postulate that the radiomics models may be a useful enhancement to existing CMR scans performed with a clinical indication; and may improve risk stratification in the future. Moreover, no external validation has been performed, and the case-control design leaves significant risk of residual confounding. Of note, only 5% of the UK Biobank population was studied and a 2.5% event rate in this hypothesis generating study. Thus, the predictiveness of the model if these radiomic metric were deployed in the general cohort remains unanswered.

5.5 Conclusions

We demonstrated the feasibility of using CMR derived radiomics features to predict key cardiovascular outcomes. Radiomics features provided additional information over VRFs, although the improvement was only marginal compared to conventional CMR metrics. The improvement was most prominent in AF and HF prediction, which highlight that the performance of radiomics models is dependent on the disease aetiology and mechanism.

TABLE 5.4: Radiomics features selected for Atrial fibrillation.

Incident cardiovascular outcome	Radiomics feature	Feature type	ROI	Phase	SVM model alone
Atrial fibrillation	Maximum 2D diameter slice	Shape	MYO	ES	0.67(±0.07)
	Energy	First-Order	MYO	ES	0.57(±0.03)
	Maximum 2D diameter column	Shape	LV	ES	0.58(±0.01)
	Maximum 2D diameter row	Shape	MYO	ES	0.60(±0.07)
	Dependence non-Uniformity	Texture	MYO	ES	0.65(±0.08)
	Inverse difference moment	Texture	MYO	ED	0.58(±0.06)
	Large area low gray level emphasis	Texture	MYO	ED	0.59(±0.06)
	Large area low gray level emphasis	Texture	MYO	ES	0.59(±0.03)
	Maximum 2D Diameter Row	Shape	LV	ES	0.56(±0.04)
	Surface area	Shape	LV	ES	0.63(±0.07)
	Maximum 2D diameter slice	Shape	LV	ED	0.62(±0.05)
	Maximum 3D diameter	Shape	MYO	ES	0.61(±0.05)
	Sum of squares	Texture	MYO	ES	0.55(±0.02)
	Zone variance	Texture	MYO	ED	0.64(±0.09)
	Maximum 2D diameter row	Shape	MYO	ED	0.58(±0.06)
	Energy	First-Order	LV	ED	0.58(±0.03)
	Gray level non-uniformity	Texture	MYO	ES	0.65(±0.04)
	Run percentage	Texture	MYO	ED	0.60(±0.08)
	Major axis	Shape	MYO	ES	0.63(±0.06)

TABLE 5.5: Radiomics features selected for heart failure and myocardial infarction.

Incident cardiovascular outcome	Radiomics feature	Feature type	ROI	Phase	SVM model alone
Heart failure	Maximum 2D diameter slice	Shape	MYO	ES	0.68(±0.06)
	Minor axis	Shape	LV	ES	0.66(±0.06)
	Volume	Shape	RV	ED	0.56(±0.05)
	Large area low gray level emphasis	Texture	MYO	ES	0.58(±0.02)
	Volume	Shape	LV	ES	0.58(±0.06)
	Informal measure of correlation1	Texture	MYO	ED	0.57(±0.07)
	Small dependence emphasis	Texture	MYO	ED	0.52(±0.05)
	Gray level non-uniformity	Texture	MYO	ED	0.64(±0.07)
	Surface area	Shape	MYO	ED	0.63(±0.03)
Myocardial infarction	Coarseness	Texture	MYO	ES	0.64(±0.02)
	Maximum 2D diameter row	Shape	RV	ED	0.54(±0.05)
	Dependence variance	Texture	MYO	ES	0.52(±0.03)
	Inverse variance	Texture	MYO	ED	0.56(±0.02)
	Large area emphasis	Texture	MYO	ED	0.62(±0.02)
	Gray level variance	Texture	MYO	ED	0.52(±0.04)
	Sphericity	Shape	RV	ES	0.53(±0.04)
	Sphericity	Shape	MYO	ED	0.61(±0.02)
	Complexity	Texture	MYO	ES	0.56(±0.04)

TABLE 5.6: Radiomics features selected for the incident cardiovascular outcome of Stroke.

Radiomics feature	**Feature type**	**ROI**	**Phase**	**SVM model alone**
Surface area to volume ratio	Shape	MYO	ED	0.64(±0.02)
Median	First-Order	MYO	ES	0.57(±0.06)
Busyness	Texture	MYO	ES	0.50(±0.04)
Large area low gray level emphasis	Texture	MYO	ES	0.55(±0.04)
Gray level non-uniformity	Texture	MYO	ES	0.53(±0.01)
Root mean squared	First-Order	MYO	ES	0.54(±0.05)
Large area low gray level emphasis	Texture	MYO	ED	0.57(±0.04)
Mean	First-Order	MYO	ES	0.55(±0.06)
Large dependence low gray level emphasis	Texture	MYO	ED	0.57(±0.04)
Sphericity	Shape	LV	ED	0.52(±0.05)
Contrast	Texture	MYO	ED	0.56(±0.03)
Gray level non-uniformity	Texture	MYO	ES	0.61(±0.01)
Difference entropy	Texture	MYO	ES	0.57(±0.04)
Energy	Texture	MYO	ES	0.48(±0.03)
Sphericity	Shape	MYO	ED	0.59(±0.04)
Joint average	Texture	MYO	ES	0.56(±0.05)
Range	First-Order	MYO	ED	0.56(±0.07)
Large area emphasis	Texture	MYO	ED	0.60(±0.01)
Sum entropy	Texture	MYO	ES	0.54(±0.02)

Chapter 6

Radiomics analysis of CMR images for the detection of genetic and familial cases in excessive trabeculation of the left ventricle

6.1 Introduction

Excessive trabeculation of the left ventricle (ETLV) is a ventricular phenotype identified by prominent myocardial trabeculation, with ongoing debate surrounding is clinical significance.(Petersen et al., 2023; Casas, Rodríguez-Palomares, and Ferreira-González, 2022). From one hand, there is substantial evidence that in certain individuals it is a normal (i.e. physiological) trait or a reversible physiological response to increased afterload, such as in exercise or pregnancy, with no clinical implications (Gati et al., 2014; Gati et al., 2013). For other patients, ETLV may be associated with a cardiomyopathy, often referred to as left-ventricular non-compaction (LVNC), secondary to a pathogenic genetic variant and leading to abnormal LV volumes or ejection fraction (LVEF) and increased risk of major adverse cardiac events (MACE)(Casas et al., 2021).

Current diagnostic criteria for ETLV are based on the ratio between compacted and non-compacted myocardium thickness (Petersen et al., 2005; Jenni, Oechslin, and Loo, 2007), do not consider alterations in ventricular function nor ultrastructural characteristics of the hypertrabeculated region, and do not differentiate physiological variations in trabeculation from cardiomyopathy. The identification of image-based biomarkers of pathological ETLV beyond systolic dysfunction may optimise the management of patients with ETLV. Along with long-term follow-up, genetic testing and familial screening offer a possibility to differentiate physiological and pathological ETLV (Oechslin and Jenni, 2018). However, this approach is associated with a high healthcare cost, genetic testing shows a positive result in only 30-40% of the patients (Arbustini et al., 2016), while familial screening requires substantial effort and cooperation, which is not always feasible. In this context, image-based descriptors may be helpful in assessing the pre-test probability of a positive genetic or familial result, providing a valuable alternative when these tests are not feasible, or in the initial evaluation of these patients without having to wait for long-term follow-up.

Radiomics is an automatic image analysis technique aiming at objectifying measures of shape and texture patterns. In cardiology, some of the radiomics measures overlap with standard imaging measures of the heart, such as LV volumes, while others encode statistical measures of brightness distribution and patterns. Taken together, these features have been shown to offer insights into cardiovascular pathophysiology, identifying unknown (Griethuysen et al., 2017b) and established (Izquierdo et al., 2021) phenotypes, and predicting the occurrence of adverse events (Cetin et al., 2018; Rauseo et al., 2021). Nonetheless, radiomics measures, often called "features", suffer for inter-image variability arising from differences in sequences and scanners (Raisi-Estabragh et al., 2020b) which can be partially overcome by image normalization (Campello et al., 2022). Despite image normalization, radiomics studies should involve multicentre data, where their high representation capacity is confronted by realistic image variability.

The aim of this study was to evaluate the capacity of radiomics in (i) identifying a genetic substrate or familial aggregation and (ii) estimating the risk of adverse events in a large multi-centre dataset of patients meeting ETLV diagnostic criteria. This study further addresses the robustness of such classification to inter-centre variability in image characteristics.

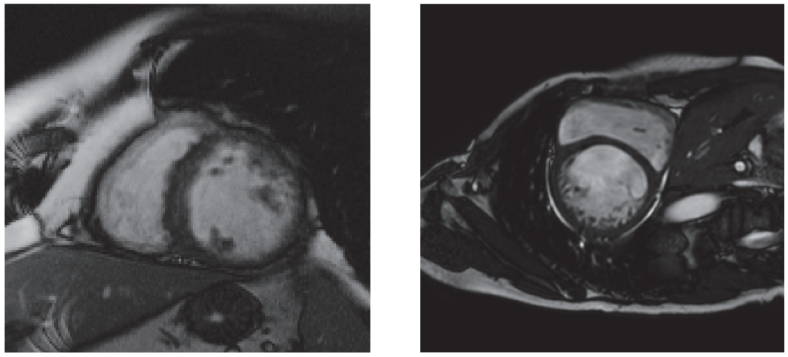

FIGURE 6.1: MRI imaging samples of LVNC subjects included in our study.

6.2 Methodology

6.2.1 Dataset

Study cohort

An observational, retrospective, longitudinal, multicentre cohort study was designed to identify individuals diagnosed with ETLV followed at 11 Spanish referral centres in the period between 2000 and 2018. A total of 347 subjects' cine short-axis (SAX) cardiovascular magnetic resonance (CMR) images were obtained along with their contours, which were manually delineated by an experienced cardiologist from VdH using the software CVi42. To ensure consistency in the Standard Operational Procedure (SOP) for segmentation [Fig.6.2], all the contouring procedures were performed at VdH location rather than at their origin centers. The CMR images were provided in DICOM format and were converted to NifTi format using the research group's internal software to facilitate the post-processing steps, which included the extraction of end-diastolic (ED) and end-systolic (ES) frame locations required for further radiomics extraction. Clinical data was provided, including genetics, demographics and distinct cardiac measures such left ventricle ejection fraction (LVEF) and follow-up endpoint named MACE (Major Adverse Cardiac Event). Both are described in the following subsections. The following table shows the distribution of subject per origin center. Diagnosis was made at the referral inherited cardiac disease unit based on Jenni criteria (Jenni, Oechslin, and Loo, 2007) on echocardiography images and was validated by cardiac magnetic resonance (CMR) according to Petersen and Jacquier criteria (Petersen et al., 2005; Jacquier et al., 2010). Fractal analysis was also performed (Captur et al., 2013). Patients were followed up on a regular yearly basis, treated according to clinical guidelines (Ponikowski et al., 2016; Yancy et al., 2013) and follow-up was censored after last contact with the outpatient clinic.

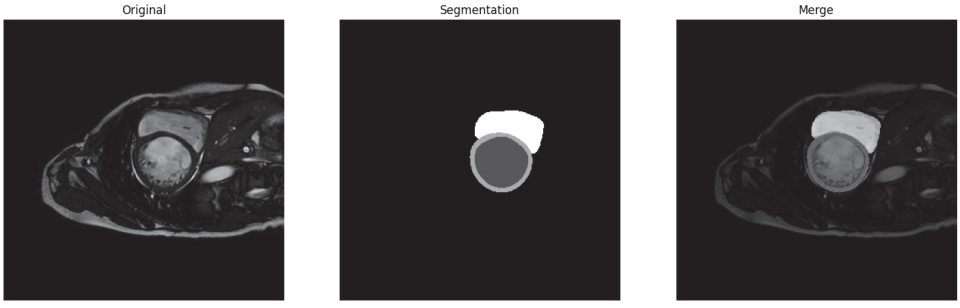

FIGURE 6.2: Example of Segmentarion Operational Procedure (SOP) for our study. Trabeculations are not segmented separately or considered part of the myocardium (MYO).

TABLE 6.1: Sample sizes at different clinical centers.

Clinical center	Code	Sample size
Hospital Universitari Vall d'Hebron	HUVH	132
Hospital Universitario Virgen de las Nieves	HVN	18
Hospital Virgen de la Arrixaca	HVAM	5
Hospital Universitario Virgen de la Victoria	HUVV	3
Hospital Universitario Puerta de Hierro	HUPH	35
Hospital Universitario Germans Trias i Pujol	HUGTP	13
Hospital Son Llàtzer	HSLL	7
Hospital La Fe	HLF	21
Hospital Josep Trueta	HJT	3
Complexo Hospitalario Universitario A Coruña	CHUAC	78
Complejo Asistencial Universitario Salamanca	CAUSA	29

Genetics and family aggregation

Active family screening was encouraged in all probands. Positive family aggregation was defined if at least one additional first-degree relative fulfilled the imaging criteria for ETLV. Genetic testing was indicated according to the criteria of each center and consisted of a next-generation sequencing panel of more than 200 genes related to inherited cardiovascular diseases. A study was considered positive if a pathogenic or likely pathogenic variant was identified (Richards et al., 2015).

Adverse events

The clinical endpoints of the study (MACE) was a combination of heart failure hospitalization, heart transplantation, LV assist device implantation, cardiac resynchronization therapy implantation, aborted sudden cardiac death, ventricular fibrillation, sustained or non-sustained ventricular tachycardia, appropriate implantable cardioverter-defibrillator therapy, systemic embolic stroke or transient ischemic attack, embolic myocardial infarction, peripheral artery embolism, or death by any cause.

6.2.2 Normalization and radiomics extraction

CMR images were centralised for core lab blinded evaluation at the Vall d'Hebron University Hospital (Barcelona, Spain). Short-axis cine CMR images were obtained and end-systolic (ES) and end-diastolic (ED) segmentations of the endo- and epicardium were performed by an experienced cardiologist using clinical software (CVi42, Circle Cardiovascular Imaging). To address the variability originating from different centers, protocols and vendors (Figure 1), piece-wise linear histogram matching normalization was performed (Campello et al., 2022), to reduce radiomics variability. Pre-processing steps were performed prior to radiomics extraction from end-diastole (ED) and end-systole (ES) frames. A total of 462 radiomics features were extracted from the LV and MYO regions using PyRadiomics (Griethuysen et al., 2017a), without any additional filtering steps, including bin size of 25.

Shape radiomics features are a subset of radiomics features that quantify and measure morphological characteristics in segmented contours, independent of the gray-level intensity distribution. These features focus on capturing surface and volume properties that are commonly calculated using established clinical indices in routine cardiac magnetic resonance (CMR) examinations. Examples of shape features include simple metrics such as volume, elongation, surface area, as well as more advanced metrics like sphericity. By analyzing shape features, it is possible to assess and understand the morphological cardiac attributes associated with specific diseases. This information can be valuable in diagnosing and stratifying the risk of patients with left ventricular noncompaction (LVNC).

In contrast, first-order radiomics features are derived from the statistical analysis of the intensity histogram of the image region defined by the mask. These features describe the distribution of pixel/voxel intensities and provide insights into grayscale changes in the left ventricle (LV) or LV myocardial tissue, which is relevant in LVNC cases. First-order statistics include straightforward metrics such as median or mean, as well as mathematically advanced measurements like entropy, energy, or kurtosis. By examining first-order features, it becomes possible to identify and quantify grayscale variations associated with trabeculations in LVNC subjects, offering important diagnostic information.

Texture radiomics features constitute another category of radiomics features, designed to capture subtle changes in the distribution of gray-scale pixel values. These features analyze patterns, trends, and relationships between neighboring gray-scale changes using sophisticated matrix calculations. Texture features can be classified into different types, including Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), Neighboring Gray Tone Difference Matrix (NGTDM), and Gray Level Dependence Matrix (GLDM). By employing texture features, it becomes possible to characterize the spatial distribution of trabeculations in the LVNC population, providing valuable insights into the arrangement and patterns of these structures.

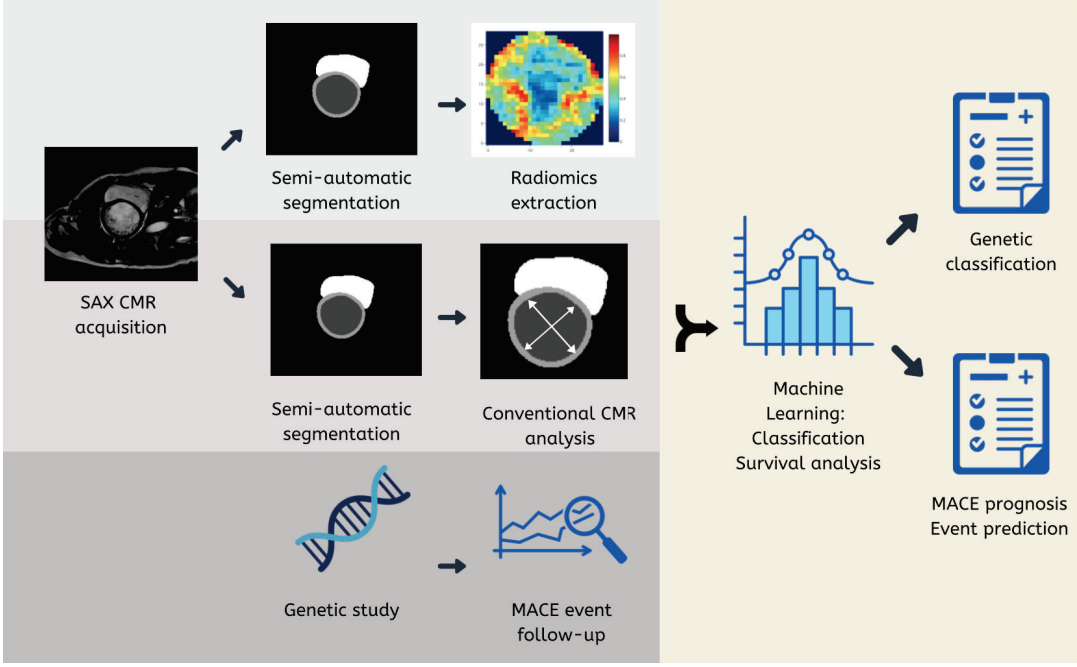

FIGURE 6.3: In the upper sequence, the methodology for radiomic feature extraction from cardiac magnetic resonance (CMR) imaging is elucidated, commencing with the initial acquisition and preceding the contour delineation. The intermediate sequence delineates the computation and assessment of traditional CMR indices, encompassing the left ventricular ejection fraction (LVEF), Petersen's coefficient, Jacquier's coefficient, and Captur analysis. The lower sequence details the genetic and familial scrutiny undertaken by the participants, alongside the longitudinal tracking of their clinical outcomes, with particular attention to the occurrence of major adverse cardiac events (MACE). On the right, the schematic represents the integration of the derived variables into various configurations for the development of Machine Learning models aimed at enhancing diagnostic accuracy, specifically for genetic classification and prognostication of MACE.

Machine Learning for genetic classification

In this study, we employed two different models, namely Imbalanced Random Forest (IMB-RF) and XGBoost (XGB), for the classification task. To enhance the performance of these models, a feature selection (FS) technique called Recursive Feature Elimination (RFE) was incorporated. Additionally, the models were integrated into a Nested Cross-Validation (NestedCV) pipeline, allowing for robust hyperparameter tuning through GridSearch.

One crucial aspect we considered during the model development was the imbalance present in the dataset. Imbalanced datasets often pose challenges for classification models as they contain a significant disparity in the number of instances between classes. To address this issue, we carefully designed the train-validation-test splitting process to take into account the class imbalance. This approach ensured that each split maintained the original class distribution, preventing biased model evaluation.

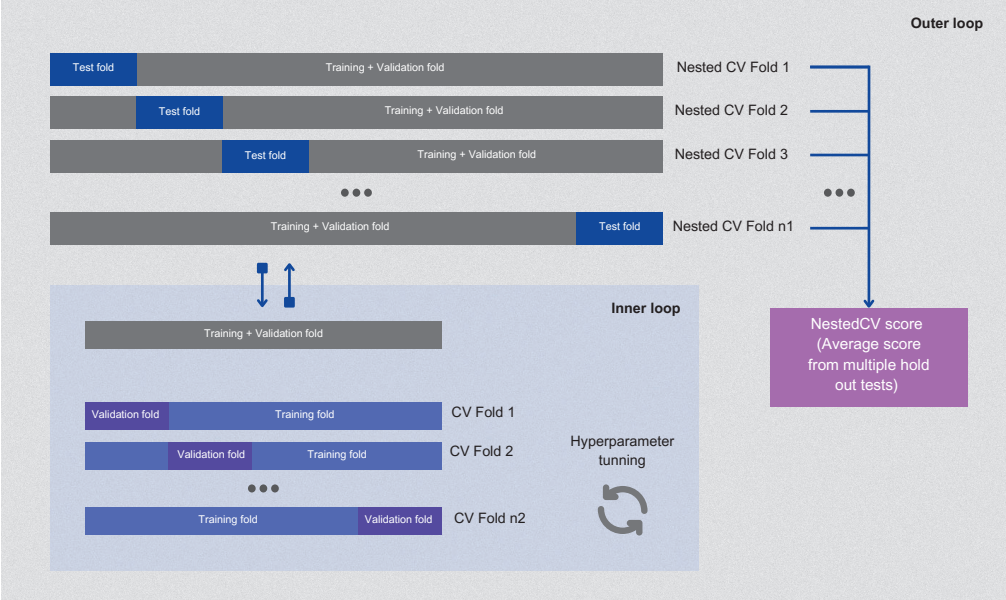

FIGURE 6.4: Architecture used for survival analysis.

To further enhance the model performance, we explored various scaling algorithms. However, interestingly, we observed that decision tree-based algorithms, including IMB-RF and XGB, are inherently invariant to scaling. This implies that the scaling of features had no substantial impact on the performance of these models. Therefore, we concluded that incorporating scaling algorithms in the pipeline did not yield any significant improvements.

Overall, our approach encompassed the following steps: feature selection using RFE, utilization of IMB-RF and XGB models, integration into a NestedCV pipeline for reliable evaluation, and consideration of dataset imbalance during the train-validation-test splitting. Furthermore, we confirmed the invariance of decision tree-based algorithms to scaling, thereby validating our choice to exclude scaling algorithms from the pipeline.

6.2.3 Machine Learning scheme for survival analysis prediction

In this section of our study, we delve into the structured process of constructing our machine learning models, designed for two distinct analyses: survival analysis and genetic classification. We provide a meticulous account of the methodologies employed to not only ascertain the performance of these models but also to evaluate their generalization capabilities.

We introduce an innovative approach to address the challenges posed by imbalanced datasets in predictive modeling. Our algorithm operates in several phases, encompassing preprocessing steps, feature selection, model training, and evaluation.

In the pre-processing stage, we tackled potential issues stemming from multicollinearity by removing variables with high correlation. This approach ensured that our models did not suffer from unnecessary redundancy and instability. Following this, we implemented data scaling to bring all features onto the same scale, a critical step in algorithms that rely on distance calculations or gradient-based optimization.

Subsequently, we applied Sequential Forward Floating Selection (SFFS) for feature selection. The SFFS algorithm operates by iteratively adding or removing features based on their contribution to the model performance until the addition or removal of features does not lead to an improvement of the prediction score. This technique allowed us to identify a subset of features that provided the best performance, improving the interpretability of our model and reducing overfitting.

For model training, given the imbalanced nature of our data, we used an Imbalanced Random Forest algorithm. This ensemble learning method creates multiple decision trees and merges them to get a more accurate and stable prediction. But unlike standard Random Forest, this algorithm adjusts the class distribution of the bootstrap sample in each iteration to handle the imbalanced data problem.

The evaluation of model performance was carried out using a nested cross-validation (CV) scheme, which provides an unbiased estimate of the model performance. In the outer loop, the data was split into a training and test set, while in the inner loop, model hyperparameters were tuned. This approach

avoided leaking information from the test set into the model during the tuning phase, hence offering a more robust estimation of the model's ability to generalize to unseen data.

We assessed model performance using the concordance index (C-index), an appropriate measure for survival models, or when the outcome is time-to-event data. The C-index measures the proportion of all usable patient pairs in which the predictions and outcomes are concordant, providing an overall measure of prediction accuracy.

Our model's top 10 performing features were identified based on their relevance and contribution to the model's predictive performance, offering insights into the key factors driving the outcomes.

6.3 Results

6.3.1 Quality control

In this study, we performed a meticulous subset selection procedure on our original dataset to ensure consistent measurement of Ejection Fraction (EF) values. We compared EF values derived from expert clinicians' MRI measurements (MRI-FEVI) with EF values obtained through radiomics analysis (RAD-FEVI). We excluded patients who exhibited a discrepancy of 15% or more between these two measurements. Discrepancies may arise from the semi-automatic segmentation of cardiac cavities or intraobserver variability when calculating MRI-FEVI. By employing this rigorous selection process as a crucial quality control step, we obtained a refined dataset consisting of 338 subjects with matching RAD-FEVI and MRI-FEVI values. Only 9 subjects were excluded during this process (see Fig.6.7), and none of the subjects experiencing MACE were removed. In light of MRI-FEVI being considered the benchmark gold standard measure, it is crucial to verify the absence of discrepancies between MRI and RAD FEVI. Any inconsistencies could potentially affect radiomics calculations and other related implications. Through this analysis, we ensure that our radiomics automatic extraction process was executed accurately.

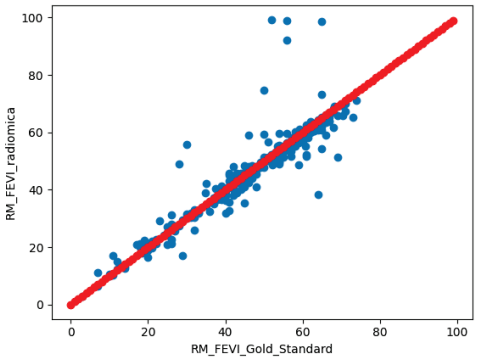

FIGURE 6.5:
Original dataset

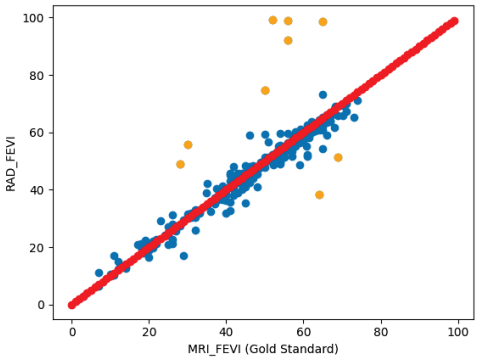

FIGURE 6.6:
Subset selection

FIGURE 6.7: Discrepancy between MRI-FEVI and RAD-FEVI

6.3.2 Results for genetic/familiar positive

After ensuring the quality and accuracy of the extracted radiomic data, it is imperative to meticulously select the subjects for the designated experiments, adhering to predefined criteria and research objectives. The initial experiments are designed to identify a genetic substrate or familial aggregation, focusing on uncovering hereditary patterns and genetic predispositions. The cohort primarily consisted of middle-aged patients, with a slight male predominance and a relatively-low prevalence of common cardiovascular risk factors. Familial screening or genetic tests were performed in 256 (76%) participants, with 100 (39%) of them presenting positive genetic findings and 58 (23%) a positive familial screening. In total, 157 (61%) patients tested positive for either familiar or genetic studies. During a follow-up of 42,5 (32.40) months, 38 (14.84%) experienced MACE. Regarding identification of genotype and family aggregation, patients presenting with a positive familial or genetic study had higher prevalence of hypertension, dyslipidemia and wide QRS, and were more likely to smoke (Table 6.2). Moreover, this group showed lower EDV and ESV LV volumes and higher prevalence of LGE, with no differences in LVEF.

Table 6.3 reports the results of the models developed to identify positive genotype and/or family aggregation. Results showed that LV ejection fraction has limited capacity to discern genotype and family aggregation (F1 score in between 40 and 60%), and that radiomics significantly improve the detection capacity, resulting in a F1 score in between 55% and 72%. Precision, also called positive predictive value, and recall, also known as sensitivity, further show the added value of including radiomic features in the analysis. Of note, the superiority of predictions including radiomics was consistent across the different tests and models.

The 10 most important features in the decision of the best model (XGB model based on LVEF and radiomics) showed that shape features are not among the most important ones (Table 6.4), despite the difference in LV geometry identified by LV volumes (Table 6.2). In contrast, textural features played the

TABLE 6.2: Clinical and demographic characteristics of the cohort depending on the results of the genotype and family aggregation studies. P-value presented for the comparisons between positive and negative genetic or familial groups. Values are percentage or mean (standard deviation). BMI = body mass index, EDV and ESV = left ventricular (LV) end diastolic and end systolic volume, EF = ejection fraction, LGE = late gadolinium enhancement, wide-QRS = QRS duration >120 ms.

Genetic/familiar experiment subset	Whole cohort	Genetic or familial (-)	Genetic or familial (+)	p-value
Demographics				
N	338	99	157	-
Male percentage (%)	56	57	54	0.100
Age (y)	44.6 ±15.8	44.4 ±21.0	42.4 ±18.0	0.210
BMI (Kg/m)	25.11 ±4.78	24.31 ±4.61	24.92 ±4.09	0.330
Hypertension (%)	20.5	20.1	33.2	< 0.0001
Diabetes (%)	5.4	6.1	12.2	0.068
Dislipidaemia (%)	27.2	19,2	41.1	< 0.001
Smoking (%)	14.2	11.1	19.6	< 0.001
Wide-QRS (%)	25.8	23.6	33.1	< 0.005
CMR indexes				
EDLVV (ml)	174 ±70	189.12 ±74	178.79 ±63	< 0.030
ESLVV (ml)	87 ±63	106 ±72	96 ±60	0.012
LVEF (%)	47.6 ±14.4	49.30 ±14.6	51.74 ±14.9	0.190
LGE (%)	16	12.1	20.2	0.006
Petersen ratio	2.85 ±0.636	2.80 ±0.58	2.88 ±0.66	0.10
Jacquier ratio	16.65 ±9.34	17.35 ±10.54	16.21 ±8.49	0.722
Fractal apical	1.39 ±0.10	1.37 ±0.09	1.39 ±0.01	0.068

most predominant role, suggesting the detection of a textural signature of the genetic cardiomyopathy in the myocardium.

TABLE 6.3: Results (F1 score [F1], Precision [P], Recall [R]) of the prediction of positive genetic study (top), positive familial screening (middle) and their combination (bottom) for the models (B-RF and XGB). P-value presented for the comparisons between a model and the corresponding model including only LVEF.Values are mean (standard deviation). Bold for best results. LVEF = left ventricular (LV) ejection fraction

	B-RF				XGBoost			
	f1	P	R	p-value	f1	P	R	p-value
Positive genetic study								
LVEF	0.44 ±0.07	0.55	0.35	-	0.40 ±0.06	0.42	0.45	-
LGE	0.40 ±0.10	0.42	0.37	-	0.40 ±0.09	0.42	0.49	-
Radiomics	0.57 ±0.02	0.70	0.48	0.052	0.55 ±0.07	0.62	0.50	0.030
LVEF + Radiomics	0.56 ±0.03	0.67	0.48	0.041	0.55 ±0.08	0.58	0.52	0.036
LGE + Radiomics	0.57 ±0.01	0.70	0.48	0.051	0.55 ±0.04	0.62	0.50	0.028
Positive family screening								
LVEF	0.54 ±0.04	0.45	0.57	-	0.45 ±0.08	0.40	0.55	-
LGE	0.42 ±0.09	0.45	0.32	-	0.41 ±0.09	0.44	0.47	-
Radiomics	0.60 ±0.03	0.79	0.42	0.05	0.58 ±0.05	0.55	0.66	0.054
LVEF + Radiomics	0.70 ±0.10	0.78	0.42	0.052	0.72 ±0.01	0.55	0.65	0.061*
LGE + Radiomics	0.60 ±0.04	0.77	0.43	0.051	0.58±0.04	0.55	0.66	0.053**
Combination								
LVEF	0.59 ±0.03	0.82	0.44	-	0.60 ±0.05	0.64	0.45	-
LGE	0.37 ±0.02	0.44	0.35	-	0.37 ±0.04	0.44	0.49	-
Radiomics	0.70 ±0.10	0.84	0.61	0.014	0.72 ±0.01	0.72	0.72	0.011
LVEF + Radiomics	0.70 ±0.10	0.88	0.59	0.031	0.72 ±0.01	0.75	0.70	0.023*
LGE + Radiomics	0.70 ±0.09	0.849	0.60	0.012	0.72 ±0.02	0.72	0.72	0.012**

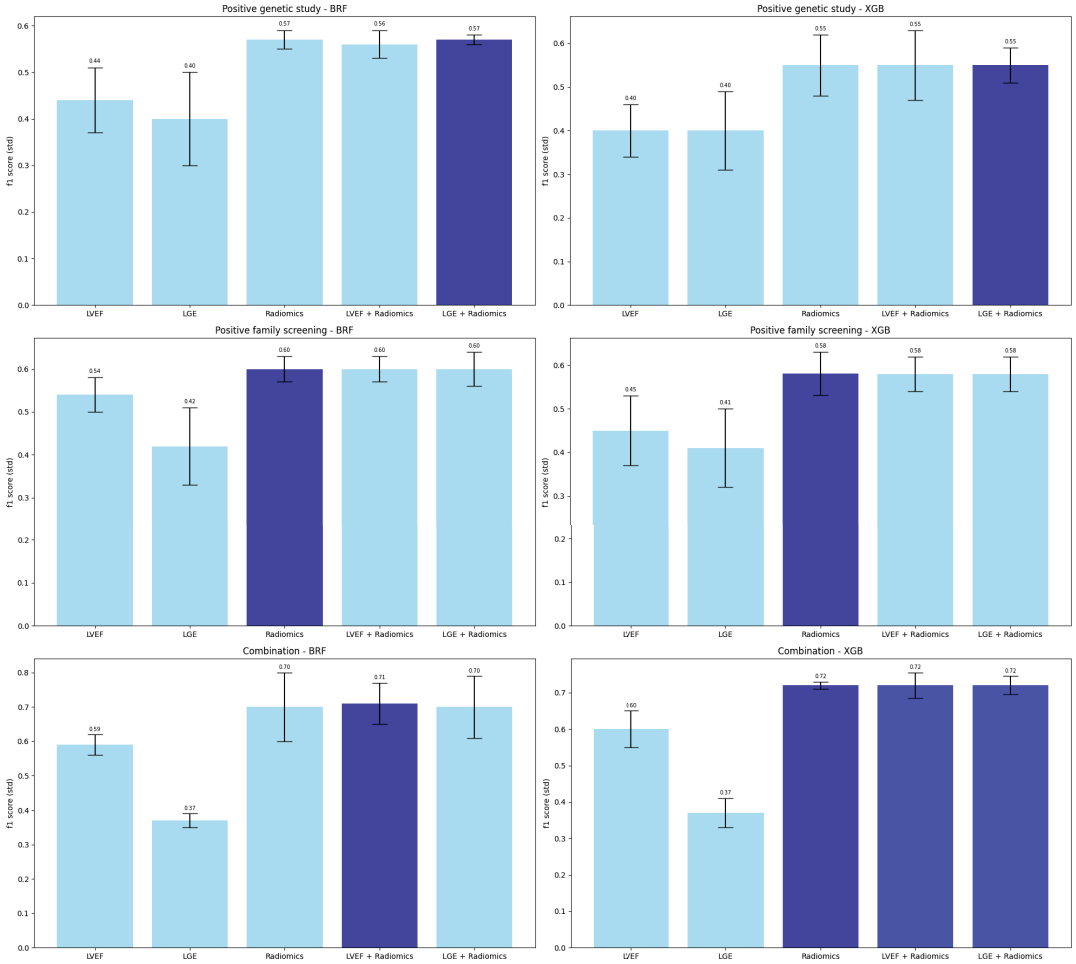

FIGURE 6.8: n the results of the genetic identification analysis, the plots reveal the tendency combining genetic studies with family screening enhances the differentiation of most variables, compared to when these methods are used independently. Furthermore, the integration of radiomics into this approach leads to an increase in F1 score values, indicating improved accuracy in the identification process

TABLE 6.4: Top 10 radiomics features in the prediction of positive genetic or family study. ED and ES = end diastolic and end systolic phases, LV = left ventricle, MYO = myocardial region.

Feature	Region	Type	Phase	Weight
Inverse Variance	LV	Texture	ES	0.16
Size Zone Non-Uniformity	LV	Texture	ED	0.13
Autocorrelation	MYO	Texture	ES	0.12
Long Run Emphasis	LV	Texture	ES	0.11
Run Variance	MYO	Texture	ES	0.11
Elongation	LV	Shape	ES	0.10
90 Percentile	LV	Tissue	ED	0.08
Dependence Entropy	MYO	Texture	ED	0.07
Contrast	LV	Texture	ED	0.07
Joint Entropy	LV	Texture	ES	0.06

6.3.3 Survival analysis time-to-event prediction for MACE

Patients experiencing a MACE had lower baseline LVEF and higher prevalence of comorbidities, smoking habit and larger LV (Table 6.5). Table 6.6 presents the results obtained for the task of identifying individuals with ETLV experiencing MACE during follow-up in both classification (left) and survival free from MACE (i.e. time-to-event, right). Regarding the classification task, LVEF alone yielded an F1 score of 69%, while radiomics alone achieved a score of 75%. When combining radiomics with LVEF, a noteworthy 10% improvement in F1 score (79%) beyond the results obtained by LVEF alone was achieved. Similarly, in the survival (i.e. "time-to-event") analysis, radiomics achieved a C-index of 0.72, outperforming the best model using LVEF (C-index of 0.67). Furthermore, integrating LVEF and radiomics yielded the best predictive performances, resulting in a C-index of 0.75. The findings are consistent across models and indicate that the inclusion of radiomics significantly enhances the identification of patients at risk beyond LVEF.

Shape features dominated the discrimination (Table 6), being the six most important features, with metrics related to short-axis size (diameter, area, volume) and shape (flatness, i.e., the squared ratio of largest and shortest distances). Notably, the following features describe textural information, suggesting that there are CMR textural differences between patients with normal ETLV and cardiomyopathy associated with MACE, which cannot be captured by LVEF.

TABLE 6.5: Characteristics of the cohort depending on the subsequent development of MACE.

Characteristics	MACE (-)	MACE (+)	p-value MACE + vs -
N	284	53	-
Male (%)	56	57	0.100
Age (y)	44.6 (15.8)	44.6 (15.8)	0.350
BMI (Kg/m^2)	24.69 (4.70)	25.88 (5.13)	0.590
Hypertension (%)	17.1	38.5	<0.001
Diabetes (%)	3.6	15.4	<0.001
Dyslipidaemia (%)	22.2	53.8	<0.001
Smoking (%)	12.3	24.5	0.016
Wide-QRS (%)	21.8	47.2	<0.001
Genetic and familiar studies			
Positive genetic (%)	24.92	32.07	0.800
Positive familiar (%)	22.98	25.41	0.130
CMR			
LV EDV (ml)	178 (62)	237 (106)	0.003
LV ESV (ml)	94 (58)	166 (101)	0.011
LV EF (%)	50.0 (12.6)	35.9 (17.5)	<0.001
LGE (%)	13.14	33.36	0.001

6.3.4 Robustness to unseen clinical centers

Generalization is a cornerstone of machine learning models, especially when they are applied in multicenter domains, such in our case. When a model generalizes well, it indicates that it has learned the underlying patterns and relationships in the data, rather than memorizing the specific details of the training set. This is essential in multicenter settings, where data can vary widely due to diverse demographics, equipment, procedures, and operational policies. A model that is overfit to the training data from one center may fail to predict accurately when deployed in a different center, leading to poor decision-making and outcomes. In figure 6.9 we can observe the differences between acquisitions from different centers. For normalization (or harmonization), we used the histogram matching normalization technique employed in Campello et al., 2022. In the publication, they recognized that such variability can compromise the integrity and reproducibility of radiomics-based models and proposed the use of image normalization as a potential solution. The pilot study specifically explored various techniques of image normalization to determine their efficacy in minimizing the discrepancies in radiomics features across different medical imaging centers. The findings indicate that certain normalization techniques can indeed enhance the consistency and reliability of radiomics data, suggesting that they could be adopted as standard practices in multi-center radiomics studies.

For a comprehensive assessment, individual analysis was conducted for each center with a sample size exceeding 20 cases. To achieve statistical significance, cases from centers 6 through 11 were consolidated into a single group. Our approach involved training the models on n-1 centers, ensuring consistent architecture mirroring that shown within the inner loop of Fig. 3.1 in the Survival Machine Learning model, specifically the Random Survival Forest (RSF). We adopted a nested cross-validation framework for this inner loop setup. This methodology allowed us to segregate a distinct test dataset from a specific center, positioning us to accurately measure the model's proficiency when faced with data from a center it hasn't been exposed to previously. This rigorous evaluation aims to provide insights into the model's adaptability and generalizability across different centers.

Table 6.8 presents the results of the generalizability tests, where the performance of the time-to-event prediction model (including LVEF and radiomics) to data from unseen clinical centers was quantified. Compared to results on Table 6.6 (C-index of 0.75), the results showed a slight decline in performance. This was particularly marked in the group containing data from a multitude of centers (bottom line), where a modest c-index of 0.54 was obtained.

TABLE 6.6: F1 score and C-index for the prediction of MACE with and without time-to-event analysis, respectively. P-value between LVEF/LGE and each specific row within the addition of radiomics. Bold for best results. LVEF = left ventricular (LV) ejection fraction. LGE = Late Gadolinium Enhanced. (*): Statistical comparison against LVEF. (**): Statistical comparison against LGE.

	Classification			**Time-to-event**		
	B-RF	**XGBoost**	**p-value**	**FS-SVM**	**RSF**	**p-value**
LVEF*	0.69 (0.02)	0.69 (0.01)	-	0.65 (0.06)	0.67 (0.03)	-
LGE**	0.63 (0.03)	0.64 (0.02)	-	0.55 (0.03)	0.56 (0.03)	-
Radiomics	0.75 (0.03)	0.75 (0.02)	0.06*	0.72 (0.04)	0.72 (0.04)	0.05*
LVEF + radiomics	0.79 (0.01)	0.79 (0.02)	0.01*	0.75 (0.06)	0.75 (0.04)	0.02*
LGE + radiomics	0.75 (0.04)	0.74 (0.05)	0.06**	0.71 (0.05)	0.71 (0.03)	0.009**

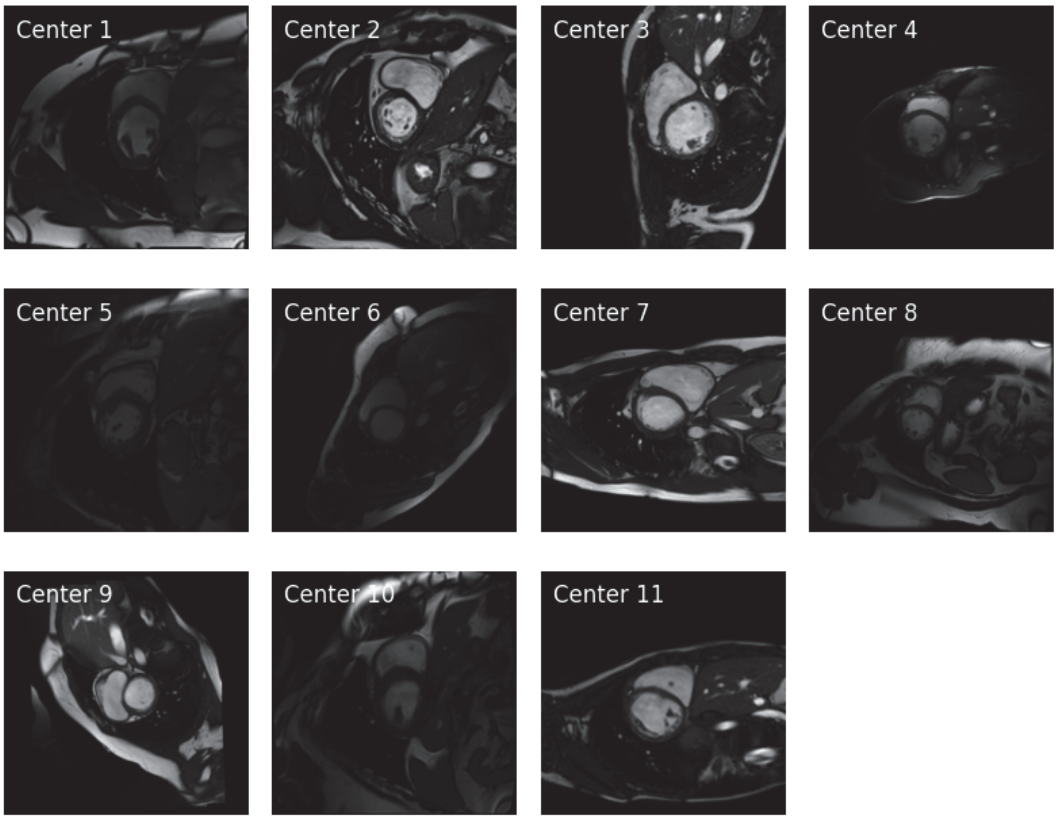

FIGURE 6.9: Individual instances of data offered by each center can be distinguished by aspects such as histogram characteristics, for example, contrast. The process of normalization becomes a critical factor in mitigating the variability between different centers

TABLE 6.7: Top 10 radiomics for the classification of MACE with (right weights) and without (left weights) time-to-event. ED and ES = end diastolic and end systolic phases, LV = left ventricle, MYO = myocardial region. WC=Weights in classification. WS=Weights in Survival analysis

Feature	Region	Type	Phase	WC	WS
Maximum 2D Diameter Slice	LV	Shape	ES	0.20	0.20
Minor Axis Length	LV	Shape	ES	-	0.17
Flatness	MYO	Shape	ED	0.20	0.16
Mesh Volume	LV	Shape	ED	0.14	0.14
Surface Area	LV	Shape	ED	0.13	0.14
Minor Axis Length	LV	Shape	ES	-	0.10
Large Dependence High GLE	MYO	Texture	ED	0.09	0.09
Cluster Shade	MYO	Texture	ES	0.09	0.09
Inverse Variance	LV	Texture	ES	-	-
Root Mean Squared	LV	Tissue	ES	-	0.09
Grey Level Variance	MYO	Texture	ED	0.07	0.07
Joint Entropy	LV	Texture	ES	-	0.05
Contrast	MYO	Texture	ED	-	0.07

TABLE 6.8: Generalizability test. The score presented is C-index. H = Harmonized. NH = Non-harmonized.

Centre	Sample size	H score (C-index)	NH score (C-index)
Centre 1	132	0.67	0.65
Centre 2	78	0.60	0.58
Centre 3	35	0.73	0.72
Centre 4	29	0.68	0.68
Centre 5	21	0.62	0.59
Centres 6-11	49	0.54	0.47

6.4 Discussion

The main results of this study are that in a large multi-center cohort of patients with ETLV radiomics (i) may identify the signature of a genetic or familial cardiomyopathy from non-enhanced, routine CMR images and (ii) may provide added prognostic value beyond LVEF. On the other hand, the results underline the challenges posed to such a high-throughput, automatic approach by the significant variability of images collected from several independent clinical centers.

There is an ongoing debate on whether ETLV is a morphologic trait or a distinct cardiomyopathy (Arbelo et al., 2023). Previous studies have proposed comprehensive patient evaluations, including symptoms, electrocardiogram, imaging, family screening and genetic studies, among others, to differentiate physiological from pathological ETLV (Casas et al., 2021; Gati et al., 2014; Caselli et al., 2015). In particular, the current CMR diagnostic criteria have a poor specificity for detecting actual cardiomyopathy cases and probably lead to an overdiagnosis of the entity(Protonotarios and Elliott, 2019), while the different criteria are inconsistent and present poor reproducibility (Ivanov et al., 2017).

The identification of cardiomyopathy has traditionally involved assessing reduced EF or abnormal cardiac structure and pacing. Genetic or familial association studies offer an alternative avenue, especially appealing for early detection of patients who may present with preserved systolic function, and to identify patients who require periodic follow-ups from those who could be safely discharged. Considering the limited availability and high cost of performing these studies, the results presented here highlight the potential offered by radiomics for the identification of positive genetic or familial association may be of clinical interest. The here-reported value of radiomics in the identification of patients harboring a positive genotype or having a positive familial aggregation was notable and consistent throughout multiple tests and model architectures. Specifically, radiomics predictions were mostly

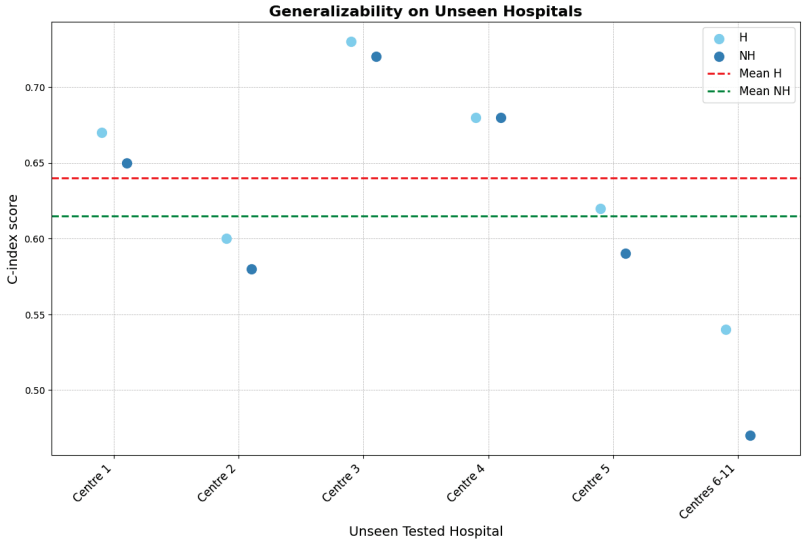

FIGURE 6.10: The scatter plot demonstrates the variability of the generalization performance of a predictive model across multiple centers. The centers are evaluated on their ability to maintain a consistent C-index score, which is a measure of the model's predictive accuracy. Category 'H' generally exhibits higher generalizability, with a mean C-index score (red dashed line) that is consistently above that of category 'NH' (green dashed line). Notably, Centre 3 shows the highest C-index score for category 'H', suggesting that the model performs best in this center. On the other hand, Centers 6-11, when aggregated, indicate a lower generalization capacity for both categories. This could imply a need for model recalibration or adaptation to improve performance in these centers. The data suggests that while the model has varying levels of generalizability, there is room for optimization, particularly in centers where the C-index score falls below the mean

based on texture descriptors, highlighting a textural signature associated with a genetic cardiomyopathy. To the best of the authors knowledge, only one previous study tested the possibility to identify a genetic positive result from cine CMR, demonstrating that a larger extent of hypertrabeculation was associated with a higher likelihood of a positive genetic result [25]. No multicenter studies nor studies using deep phenotyping techniques are available in this regard, while no other radiomics applications in ETLV or in other genetic cardiomyopathies are available. Further studies are certainly needed.

To the best of the authors' knowledge, only one small and single-center study tested the possibility of identifying a genetic positive result from cine CMR, reporting that a larger extent of ETLV was associated with a higher likelihood of a positive genetic result (Waning et al., 2021). This result was not reproduced in the present cohort (Table 1), since neither Petersen, Jacquier nor fractal values differed between patients with a positive or negative result from the genetic and familial studies, suggesting that the degree of ETLV alone does not define the presence of an actual cardiomyopathy, aligning with previous reports (Grigoratos et al., 2019; Amzulescu et al., 2015). No multicenter studies nor studies using deep phenotyping techniques are available in ETLV patients; further studies in these patients are needed. Notably, conventional and deep image analysis techniques were successfully used for imaging-based predictions of genotypes in other cardiomyopathies (Morita et al., 2021; Bos et al., 2014; Gruner et al., 2013).

In the present cohort radiomics demonstrated a 10% improvement in risk stratification of patients beyond LVEF or LGE. While it is worth underlying that other factors have been shown to modulate risk in patients with ETLV (Casas et al., 2021), LVEF remains the cornerstone of clinical management of cardiomyopathy patients, including those with ETLV (Aung et al., 2020), while previous studies underlined the prognostic impact of having a positive LGE finding (Casas et al., 2021). Thus, present data may reflect the potential of radiomics features to provide additional information regarding the cardiomyopathic nature and risk profile of these patients, which is clinically relevant considering the high heterogeneity of this population (Petersen et al., 2023). Also, our findings may allow to avoid the administration of contrast media, if confirmed in future prospective studies.

These results are in line with a previous, single-center and substantially smaller study, where radiomics showed added value for MACE prediction beyond clinical data (Han et al., 2023). Nonetheless, while promising, present results also underline the challenges of obtaining a robust model in a multicenter imaging context, where the variability in image appearance is vast. Indeed, despite the use of state-of-the-art image normalization (piece-wise linear histogram matching) to account for center variability, a significant drop in performance was obtained when data of the center of interest were omitted. This underscores the need for future work focusing on dealing with multi-center variability in radiomics for the development of robust machine learning models.

6.4.1 Limitations

Despite the significance of the study, it presents certain limitations. Although this study is based on data from 11 clinical centers, they were all from Spanish Institutions. Therefore, the generalizability of the results to a broader international context as well as the fairness of the proposed approach in terms of ethnicity should be evaluated in further studies. Another limitation arises from the limited number of confounders included in the event prediction tests. In particular, certain prognostic markers, such as LGE CMR, were not included in models. Larger studies are needed to test whether radiomics provide added diagnostic or prognostic information beyond those established predictors.

6.4.2 Conclusions

In a multicenter cohort study of individuals diagnosed with excessive trabeculation of the left ventricle radiomics analysis of standard, non-enhanced cine CMR images provided added value beyond left ventricular ejection fraction in the identification of a genetic or familial substrate and of adverse prognosis. Textural radiomics features were instrumental to recognize a genetic or familial substrate, while shape features dominated the identification of adverse prognosis.

Chapter 7

Conclusions

7.1 Generic conclusions

This PhD thesis offers innovative contributions in the integration of radiomics with machine learning (ML), representing a transformative approach to address significant clinical challenges in the diagnosis, prognosis of complex cardiovascular diseases, and forecasting of adverse events in complex cardiovascular diseases. More precisely, the main contributions of this dissertation are:

- In the research outlined in Chapter 3, we proposed a fully automatic CMR radiomics-based ML pipeline for distinguishing between complex and overlapping conditions like hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), and left ventricular non-compaction (LVNC). The accuracy of the proposed radiomics-based models aligns with that of conventional CMR metrics, offering the added advantage of completely automated diagnoses. This method eliminates the necessity for manual trabeculae delineation by experts, as required in prior studies, streamlining the diagnostic process.

- In Chapter 4, we demonstrated the the efficacy of integrating radiomics with ML for the prognostication of Atrial Fibrillation (AF). The ECG-based model demonstrated a reduced capability in identifying AF in female subjects compared to male counterparts. Enhancing this model with CMR radiomics, particularly when paired with ECG data, significantly improved its accuracy for female subjects. Notably, radiomic shape features derived from the left atrium (LA) via CMR played a pivotal role in enhancing our model's effectiveness, highlighting the importance of atrial remodeling in the pathophysiology of AF. This suggests that integrating ECG data with atrial imaging techniques could result in more accurate assessments for women who are suspected of having AF. Further analysis targeted patients with prevalent AF but normal ECG readings, where the incorporation of radiomics consistently improved model performance across the board.

- In chapter 5, our research proved the effectiveness of utilizing radiomics features extracted from CMR to forecast critical cardiovascular outcomes (CVD). These features offer supplementary insights beyond traditional vascular risk factors (VRFs), albeit with a relatively modest enhancement when juxtaposed with standard CMR measurements. This augmentation in predictive accuracy was notably significant in cases of atrial fibrillation (AF) and heart failure (HF), underscoring the notion that the efficacy of radiomics-based models is contingent upon the specific pathogenesis and underlying mechanisms of the condition in question.

- Finally, in chapter 6 we proved that in a study spanning multiple centers, involving participants identified with left ventricular excessive trabeculation, an analysis using radiomics from standard, non-enhanced cine CMR images revealed additional value beyond the assessment of left ventricular ejection fraction for detecting a genetic or familial basis and predicting adverse cardiac events. The study found that textural radiomics features were crucial in identifying genetic or familial connections, whereas shape-based radiomics features played a key role in predicting adverse prognostic outcomes.

7.2 Impact of the research

Integrating radiomics into the clinical workflow offers a promising avenue to augment the prognostic evaluation of cardiovascular conditions. By combining radiomic features with established clinical biomarkers, healthcare professionals can achieve a more comprehensive and nuanced understanding

of a patient's disease state. This multidimensional analysis enables the identification of subtle patterns and associations that may elude conventional diagnostic methods.

The synergy between radiomics and clinical biomarkers facilitates the development of personalized treatment plans tailored to the individual characteristics of a patient's disease. This approach not only holds the potential to improve the accuracy of cardiovascular disease predictions but also offers insights into the mechanisms driving complex cardiovascular diseases. Consequently, radiomics could play a pivotal role in advancing precision medicine for cardiovascular care, leading to improved patient outcomes and optimized treatment strategies.

In summary, our thesis showed that the union of radiomics with traditional clinical biomarkers signifies a transforming advancement in cardiovascular diagnostics, underscoring the power of innovative technologies to refine patient care. This approach not only enhances diagnostic precision but also introduces a tailored, predictive model for the effective management of cardiovascular diseases, including atrial fibrillation, Stroke, Heart Failure, Myocardial infarction, HCM, DCM and left ventricular non-compaction (LVNC). As radiomics continues to advance, its potential to redefine our methodologies for diagnosing, understanding, and treating complex conditions such as LVNC becomes increasingly evident, promising a future where medical interventions are more accurately aligned with individual patient profiles.

7.3 Future work

Despite the importance of our findings, there are some limitations to the current work, related to disadvantages of using artificial intelligence (AI) in medicine, specially associated to the radiomics pipelines and its adoption in clinical practice:

- **Reliability and Reproducibility:** The process of acquiring images across multiple centers and using different scanner brands introduces significant variability in radiomics data, which can impede the reliability and reproducibility of research findings. This variability arises primarily due to differences in acquisition protocols and scanner technologies, which can affect the integrity of radiomics features and, consequently, the trends and interpretations derived from the data.

  To mitigate these challenges and enhance data consistency, it is crucial to standardize image acquisition processes. Standardization involves establishing uniform criteria that can be universally applied across various settings, ensuring that the extracted radiomics models retain their applicability and accuracy irrespective of the location or equipment used. Such standardization not only facilitates more robust comparative studies but also improves the clinical utility of radiomics models.

  Additionally, the disparity in scanner brands and settings can substantially alter the acquired data, potentially leading to inconsistent radiomics features. Numerous studies have explored the impact of these variables on radiomics analysis, highlighting the necessity for normalization techniques (Raisi-Estabragh et al., 2020c; Campello et al., 2022). In our research, we employed specific normalization strategies to address these issues. While these methods provided some level of control over data variability, the quest for more refined techniques remains. Ongoing research and development of advanced normalization methods are imperative to further minimize data discrepancies and bolster the reliability of radiomics analyses across different studies and clinical applications.

  Continued advancements in this field could lead to the development of more sophisticated standardization and normalization protocols (Zwanenburg et al., 2020), thereby enhancing the predictive power and clinical relevance of radiomics. Engaging in cross-disciplinary collaborations and leveraging technological innovations are critical steps toward achieving these goals.

- **Inter-observer variability in segmentation:** The role of segmentation in radiomics extraction is critical, as it significantly influences the quality and reliability of the derived data. While existing guidelines aim to standardize segmentation practices, considerable variability can still occur due to differences in how individual clinicians interpret and apply these criteria. This variability is particularly pronounced in radiomics features that describe morphological aspects of the tissue, which are sensitive to the precise boundaries defined during the segmentation process.

  To address the challenge of inter-observer variability, the scientific community is actively investigating the potential of automated segmentation techniques (Cardenas et al., 2019; Campello et al., 2021; Martín-Isla et al., 2023). Automation aims to minimize human bias and inconsistency

by standardizing the segmentation process, thus enhancing the reproducibility of radiomics analyses across different studies and settings. Several studies have demonstrated the efficacy of automated tools in reducing variability and improving the precision of radiomics features.

In our study, we have implemented a centralized approach to segmentation to further control variability. By centralizing the segmentation process, we ensured that all images were processed under consistent guidelines and by a dedicated team, reducing the influence of individual clinician variability on the segmentation outcomes. This approach not only improves the consistency of our radiomics data but also provides a robust foundation for comparing our results with other studies that may employ different segmentation protocols.

Ongoing research into more advanced automated segmentation methods is essential. As these technologies evolve, they hold the promise of enabling even more consistent and accurate radiomics extraction, potentially revolutionizing the field by allowing for the seamless integration of radiomics into clinical practice.

- **Interpretability:** Radiomics involves a range of variables; some are straightforward and easy to grasp, whereas others require specialized tools or expertise for interpretation. To enhance the adoption of radiomics in clinical settings, it's crucial to link these variables with visual characteristics that are easily interpretable (Abbasian Ardakani et al., 2022). In the medical field, the ability to clearly explain and interpret findings is especially important for patient understanding.

  To improve the adoption of radiomics in clinical settings, there is a crucial need to make radiomic features more interpretable. This entails mapping complex radiomic variables to visual characteristics that clinicians can easily understand and explain. Simplifying the interpretation of radiomic data without compromising its analytical depth is essential for effective communication between healthcare providers and patients. Enhancing interpretability not only aids in clinical decision-making but also helps in educating patients about their conditions, fostering a better understanding and increasing trust in the use of advanced imaging technologies in their treatment plans (Lambin et al., 2017).

  Moreover, developing user-friendly software tools that can visualize radiomic data in an intuitive manner will further facilitate its integration into routine clinical practice. These tools should be designed to bridge the gap between the complex quantitative data that radiomics provides and the qualitative insights that clinicians need *euCanSHare* n.d.). By offering visual representations of radiomic findings, such tools can make these data points accessible to clinicians without specialized training in data science, thereby broadening the scope of radiomics' applicability in diverse medical specialties.

- **Validation:** A significant challenge in radiomics is the validation of computational models, which are typically trained on specific datasets. When these models are applied to different patient groups or external datasets, their performance can vary significantly. The intricacy and high dimensionality of radiomic features, when combined with the often limited sample sizes, increase the likelihood of developing models that fail to generalize effectively beyond the environments for which they were originally designed.

  Validation challenges in radiomics are compounded by the variability inherent in imaging data. Different imaging modalities, such as CT, MRI, or PET, produce data with unique characteristics and scales, which can affect the performance of radiomic models when applied across modalities not included in the training set. Furthermore, variations in imaging protocols between institutions can introduce additional discrepancies that complicate the generalization of these models. Standardization of imaging protocols and the use of harmonization techniques are critical steps towards improving the robustness and transferability of radiomic models across diverse clinical settings (Lambin et al., 2017).

  Moreover, the statistical methods used to evaluate model performance also play a crucial role in the validation process. Traditional metrics such as accuracy, sensitivity, and specificity might not be entirely sufficient for assessing the performance of radiomic models due to their high-dimensional nature. Advanced statistical techniques, such as cross-validation and bootstrapping, are often recommended to provide more reliable performance estimates. Additionally, incorporating external validation through multicentric studies can further enhance the credibility and generalizability of radiomic assessments.

- **Lack of Prospective Studies:** Radiomics research is often based on retrospective data analysis, which limits the ability to apply findings predictively in clinical practice. For radiomics to be

integrated into clinical decision-making, there is a need for prospective studies that validate the predictive capabilities of radiomic features in real-time clinical scenarios.

The reliance on retrospective data in radiomics poses significant limitations because these datasets may not fully capture the diversity and complexity of future patient populations. Retrospective studies are typically constrained by the data's existing biases and the conditions under which the data were collected, which may not accurately represent current clinical practices or patient demographics. This can lead to models that perform well on historical data but falter in prospective, real-world scenarios. Therefore, conducting prospective radiomic studies is crucial as they allow researchers to validate and refine models under contemporary clinical conditions and with real-time data collection (Yip and Aerts, 2016).

Prospective radiomic studies also offer the opportunity to evaluate the clinical utility of radiomic features before they are implemented in routine practice. These studies can provide insights into how radiomic models influence clinical outcomes and decision-making processes. Furthermore, prospective trials can facilitate the standardization of imaging protocols and data collection methods, ensuring that the radiomic features extracted are reliable and applicable across different clinical settings. This shift from retrospective to prospective analysis in radiomics research is essential for moving the field towards practical, evidence-based applications in personalized medicine.

- **Effectiveness:** The practical effectiveness of radiomics in enhancing clinical outcomes is still an area of active investigation. While radiomics has shown potential in research phases, translating these findings into clinical practice and demonstrating real-world benefits remains a challenge. Effective implementation requires evidence from controlled trials that assess the impact on patient management, healthcare costs, and overall clinical efficacy. The transition from research to clinical practice in radiomics is fraught with challenges, particularly in demonstrating tangible improvements in patient outcomes. While research phases often show promising results, these findings do not always translate into effective clinical tools due to various barriers such as technological integration, clinical workflow adaptation, and healthcare professional training. Moreover, the absence of robust clinical trials focusing on patient-centered outcomes makes it difficult to assess the true impact of radiomics. Therefore, to establish radiomics as a valuable clinical tool, there is a pressing need for controlled trials that not only assess diagnostic and predictive accuracy but also focus on how radiomics can improve patient management, reduce healthcare costs, and enhance overall treatment efficacy (Gillies, Kinahan, and Hricak, 2016).

  Further compounding the challenge is the need for multi-disciplinary collaboration to effectively integrate radiomics into clinical practice. The development and implementation of radiomic strategies require concerted efforts from radiologists, oncologists, data scientists, and IT professionals. This collaboration is essential for designing systems that are both clinically relevant and technically feasible (Yip and Aerts, 2016). Additionally, educating healthcare professionals about the benefits and limitations of radiomics will be crucial for its acceptance and effective use. As such, continuous professional development and training programs must be part of the implementation strategy to ensure that radiomics tools are used optimally to achieve the best clinical outcomes.

Appendix A

Appendix A: Supplementary Material

TABLE A.1: Supplementary Table I: Disease definitions.

Source	UKB Field ID: code	Description
Myocardial infarction		
Self-report	20002	Heart attack/myocardial infarction
Algorithm	42002	Date of myocardial infarction
ICD10	I21	Acute myocardial infarction
	I22	Subsequent myocardial infarction
	I23	Certain current complications following acute myocardial infarction
First occurrences	131298	Acute myocardial infarction
	131300	Subsequent myocardial infarction
	131302	Certain current complications following acute myocardial infarction
Diagnosed by doctor	6150 : 1	Heart attack
ICD9	3894	Age heart attack diagnosed
	410	Acute myocardial infarction
	411	Other acute and subacute forms of ischaemic heart disease
	412	Old myocardial infarction
Heart failure		
Self-report	20002	Heart failure/pulmonary oedema
ICD10	I500	Congestive heart failure
	I501	Left ventricular failure
	I509	Heart failure, unspecified
First occurrences	131354	heart failure
Atrial fibrillation		
Self-report	20002	Atrial fibrillation
ICD10	I480	Paroxysmal atrial fibrillation
	I481	Persistent atrial fibrillation
	I482	Chronic atrial fibrillation
Stroke		
Self-report	20002	Stroke
Algorithm	42002	Ischaemic stroke
	42006	Brain haemorrhage
	42008	Date of stroke
	42100	Date of ischaemic stroke
Diagnosed by doctor	6150 : 3	Stroke
ICD10	40506	Age stroke diagnosed
	I60	Intracranial haemorrhage
	I61	Other nontraumatic intracranial haemorrhage
	I62	Other nontraumatic intracranial haemorrhage
ICD9	164	Stroke, not specified as haemorrhage or infarction
	431	Intracerebral haemorrhage
	434	Other and unspecified intracranial haemorrhage
	436	Acute but ill-defined cerebrovascular disease
First occurrences	131362	Intracerebral haemorrhage
	131364	Other nontraumatic intracranial haemorrhage
	131366	Cerebral infarction
	131368	Stroke, not specified as haemorrhage or infarction
Diabetes		
Diagnosed by doctor	2443	Diabetes diagnosed by doctor
High cholesterol		
Medications	6177, 6153 : 1	Cholesterol lowering medication
Biochemistry	30690	Cholesterol > 7 mmol/L
Hypertension		
Medications	6177, 6153 : 2	Blood pressure medication

ICD10 codes are drawn from fields 41270, 41280, 41234 and 41259 ICD9 codes are drawn from fields

41271, 41281, 41234 and 41259 Where a 3-digit code is given, this includes all 4-digit sub-codes, for example, I21 includes I210, I211 and I212 etc.

TABLE A.2: Supplementary Table 2: Atrial fibrillation.

AF Model	Features Selected (ordered as selected)	Type	ROI	Phase	SVM (alone)
VRF					
	Sex				0.66(±0.05)
	Age				0.59(±0.04)
	Hypertension				0.59(±0.06)
CMR					
	LVM				0.66(±0.07)
	RVF				0.59(±0.03)
	LVEDV				0.59(±0.02)
	LVSV				0.56(±0.05)
CMR+VRF					
	LVM				0.66(±0.07)
	Hypertension				0.59(±0.03)
	Age				0.59(±0.06)
	RVEDV				0.62(±0.04)
	LVEDV				0.59(±0.03)
Radiomics					
	Maximum 2D diameter slice	Shape	MYO	ES	0.57(±0.07)
	Maximum 2D diameter column	Shape	LV	ES	0.58(±0.01)
	Maximum 2D diameter row	Shape	MYO	ES	0.60(±0.07)
	Dependence non-Uniformity	Texture	MYO	ED	0.65(±0.08)
	Inverse difference moment	Texture	MYO	ED	0.58(±0.06)
	Large area low gray level emphasis	Texture	MYO	ED	0.59(±0.03)
	Large area low gray level emphasis	Texture	MYO	ES	0.59(±0.04)
	Maximum 2D Diameter row	Shape	LV	ES	0.63(±0.07)
	Surface area	Shape	LV	ED	0.62(±0.05)
	Maximum 2D diameter slice	Shape	LV	ES	0.61(±0.05)
	Maximum 3D diameter	Shape	MYO	ES	0.61(±0.05)
	Sum of squares	Texture	MYO	ES	0.55(±0.02)
	Zone variance	Texture	MYO	ED	0.64(±0.09)
	Maximum 2D diameter row	Shape	MYO	ED	0.58(±0.06)
	Energy	First-Order	LV	ED	0.68(±0.03)
	Gray level non-uniformity	Texture	MYO	ES	0.65(±0.07)
	Run percentage	Texture	MYO	ED	0.60(±0.08)
	Major axis	Shape	MYO	ES	0.67(±0.07)
Radiomics+VRF					
	Maximum 2D diameter slice	Shape	MYO	ES	0.63(±0.06)
	Age				0.59(±0.04)
	Small area low gray level emphasis	Texture	MYO	ES	0.58(±0.04)
	Long run low gray level emphasis	Texture	MYO	ES	0.45(±0.02)
	Percentage non-uniformity	Texture	MYO	ES	0.60(±0.03)
	Long run high gray level emphasis	Texture	MYO	ED	0.52(±0.04)
	Gray level non-uniformity emphasis	Texture	MYO	ES	0.58(±0.02)
	Hypertension				0.59(±0.06)
	Gray level non-uniformity normalized	Texture	MYO	ES	0.61(±0.03)
	Large area low gray level emphasis	Texture	MYO	ES	0.59(±0.04)
	Maximum 2D diameter slice	Shape	MYO	ES	0.57(±0.07)
	Surface area	Shape	RV	ED	0.64(±0.03)
CMR+Radiomics					
	LVM				0.66(±0.04)
	Contrast	Texture	MYO	ED	0.59(±0.03)
	Sphericity	Shape	RV	ES	0.59(±0.05)
	Major axis	Shape	LV	ES	0.63(±0.06)
	Maximum 2D diameter slice	Shape	MYO	ED	0.62(±0.05)
CMR+Radiomics+VRF					
	Maximum 2D diameter slice	Shape	MYO	ES	0.67(±0.07)
	Age				0.59(±0.04)
	Small area low gray level emphasis	Texture	MYO	ES	0.48(±0.04)
	Long run low gray level emphasis	Texture	MYO	ES	0.55(±0.02)
	Percentile 90th	Texture	MYO	ES	0.54(±0.06)
	Gray level non-uniformity	Texture	MYO	ES	0.58(±0.04)
	Long run high gray level emphasis	Texture	MYO	ED	0.52(±0.04)
	Hypertension				0.59(±0.06)
	Gray level non-uniformity normalized	Texture	MYO	ES	0.51(±0.02)
	Large area low gray level emphasis	Texture	MYO	ES	0.59(±0.03)

Abbreviations: CMR, cardiac magnetic resonance imaging; VRF, vascular risk factor, ROI, region of interest, SVM model alone: support vector machine model performance showing the mean and standard deviation using each radiomic feature individually; LV, left-ventricle; RV, right-ventricle; MYO, left ventricle myocardium; ED, end-diastolic, EF, ejection fraction, EDV end-diastolic volume, ESV, end-systolic volume, LV, left ventricle, RV right ventricle, SV stroke volume.

TABLE A.3: Supplementary Table 4: Myocardial infarction.

MI Model	Features Selected (ordered as selected)	Type	ROI	Phase	SVM (alone)
VRF					
	Sex				0.66 (±0.04)
	Hypertension				0.58 (±0.03)
	Body Surface Area				0.56 (±0.02)
CMR					
	LVM				0.65 (±0.02)
	LVSV				0.55 (±0.06)
	LVEDV				0.57 (±0.04)
	RVEDV				0.57 (±0.04)
CMR+VRF					
	Sex				0.66 (±0.04)
	Body mass index				0.56 (±0.03)
	Hypertension				0.59 (±0.02)
	Body Surface Area				0.58 (±0.03)
	LVEDV				0.56 (±0.05)
Radiomics					
	Coarseness	Texture	MYO	ES	0.54 (±0.06)
	Maximum 2D diameter row	Shape	RV	ED	0.54 (±0.02)
	Dependence variance	Texture	MYO	ED	0.52 (±0.03)
	Inverse variance	Texture	MYO	ED	0.56 (±0.02)
	Large area emphasis	Texture	MYO	ED	0.52 (±0.04)
	Gray level variance	Texture	MYO	ES	0.53 (±0.02)
	Sphericity	Shape	MYO	ES	0.61 (±0.04)
	Complexity	Texture	MYO	ES	0.56 (±0.04)
Radiomics+VRF					
	Sex				0.66 (±0.04)
	Small dependence low gray level emphasis	Texture	MYO	ED	0.56 (±0.03)
	Hypertension				0.58 (±0.03)
	Body Surface Area				0.58 (±0.03)
	Maximum 2D diameter slice	Shape	MYO	ES	0.60 (±0.05)
	Maximum 2D diameter uniformity	Texture	MYO	ES	0.54 (±0.04)
	Max 2D diameter slice	Shape	LV	ES	0.64 (±0.04)
	Zone entropy	Texture	MYO	ES	0.53 (±0.06)
CMR+Radiomics					
	LVM	Shape	RV	ED	0.65 (±0.02)
	Least axis	Shape	RV	ED	0.54 (±0.04)
	Major axis	Shape	RV	ES	0.59 (±0.03)
	Surface area	Shape	MYO	ES	0.60 (±0.03)
	Maximum 2D diameter row	Shape	MYO	ES	0.57 (±0.03)
	Gray level non-uniformity normalized	Texture	MYO	ED	0.57 (±0.01)
	Large area high gray level emphasis	Texture	MYO	ES	0.58 (±0.02)
	Volume	Shape	MYO	ES	0.63 (±0.03)
	Maximum 2D diameter slice	Shape	MYO	ES	0.64(±0.05)
CMR+Radiomics+VRF					
	Sex				0.66 (±0.04)
	Small dependence low gray level emphasis	Texture	MYO	ED	0.56 (±0.03)
	Hypertension				0.58 (±0.03)
	Maximum 2D diameter slice	Shape	MYO	ES	0.60 (±0.05)
	Maximum 2D diameter slice non-uniformity	Texture	MYO	ES	0.54 (±0.04)
	Zone Entropy	Texture	MYO	ES	0.58 (±0.06)
	Maximum	First-Order	MYO	ES	0.53 (±0.02)

Abbreviations: CMR, cardiac magnetic resonance imaging; VRF, vascular risk factor, ROI, region of interest, SVM model alone: support vector machine model performance showing the mean and standard deviation using each radiomic feature individually; LV, left-ventricle; RV, right-ventricle; MYO, left ventricle myocardium; ED, end-diastolic, EF, ejection fraction, EDV end-diastolic volume, ESV, end-systolic volume, LV, left ventricle, RV right ventricle, SV stroke volume.

TABLE A.4: Supplementary Table 3: Heart Failure

HF Model	Features Selected (ordered as selected)	Type	ROI	Phase	SVM (alone)
VRF					
	Age				0.65(±0.08)
	Body surface area				0.61(±0.03)
	Hypertension				0.62(±0.04)
	Diabetes				0.53(±0.02)
	High cholesterol				0.52(±0.06)
	Body mass index				0.51(±0.05)
CMR					
	LVEDV				0.66(±0.05)
	LVM				0.64(±0.05)
	RVEDV				0.57(±0.05)
	LVEF				0.66(±0.06)
	LVSV				0.62(±0.04)
	RVSV				0.51(±0.03)
	LVESSV				0.66(±0.05)
CMR+VRF					
	Age				0.65(±0.08)
	Body mass index				0.61(±0.03)
	Body surface area				0.61(±0.06)
	Hypertension				0.62(±0.04)
	Diabetes				0.53(±0.03)
	LVEDV				0.59(±0.02)
Radiomics					
	Maximum 2D diameter slice	Shape	MYO	ES	0.68(±0.06)
	Minor axis	Shape	LV	ES	0.66(±0.06)
	Volume	Shape	RV	ED	0.56(±0.05)
	Large area low gray level emphasis	Texture	MYO	ES	0.58(±0.02)
	Informal measure of correlation1	Texture	MYO	ED	0.57(±0.07)
	Small dependence emphasis	Texture	MYO	ES	0.52(±0.05)
	Gray level non-uniformity	Texture	MYO	ED	0.63(±0.07)
	Surface area	Shape	MYO	ED	0.64(±0.03)
Radiomics+VRF					
	Maximum 2D diameter slice	Shape	MYO	ES	0.67(±0.06)
	Minor axis	Shape	LV	ES	0.66(±0.06)
	Age				0.59(±0.04)
	Hypertension				0.56(±0.04)
	Major axis	Shape	RV	ED	0.62(±0.05)
	Size zone non-uniformity normalized	Texture	MYO	ED	0.53(±0.06)
	Least axis	Shape	RV	ED	0.52(±0.05)
	Maximum 2D diameter slice	Shape	MYO	ES	0.68(±0.06)
	LVEDV	Shape	LV	ES	0.60(±0.05)
	Dependence non-uniformity	Texture	MYO	ES	0.52(±0.03)
CMR+Radiomics+VRF					
	Maximum 2D diameter slice	Shape	MYO	ES	0.66(±0.06)
	Minor axis	Shape	LV	ES	0.56(±0.03)
	LVEDV	Shape	MYO	ED	0.62(±0.05)
	Dependence non-uniformity	Texture	MYO	ED	0.62(±0.03)
	Hypertension				0.60(±0.05)
	Mean absolute deviation	First-Order	MYO	ES	0.64(±0.03)
	Run length non-uniformity	Texture	MYO	ES	0.63(±0.08)
	Age				0.56(±0.04)
	Complexity	Texture	MYO	ED	0.54(±0.04)
	Low gray level zone emphasis	Texture	MYO	ES	0.53(±0.04)

Abbreviations: CMR, cardiac magnetic resonance imaging; VRF, vascular risk factor; ROI, region of interest; SVM model alone: support vector machine model performance showing the mean and standard deviation using each radiomic feature individually; LV, left-ventricle; RV, right-ventricle; MYO, left ventricle myocardium; ED, end-diastolic; EF, ejection fraction; EDV, end-diastolic volume; ESV, end-systolic volume, LVEF, left ventricular ejection fraction; SV, stroke volume.

TABLE A.5: Supplementary Table 5: Stroke

Stroke Model	Features Selected (ordered as selected)	Type	ROI	Phase	SVM (alone)
VRF					
	Sex				0.62 (±0.03)
	Body surface area				0.58 (±0.05)
	Hypertension				0.55 (±0.03)
	Age				0.61 (±0.05)
CMR					
	LVM				0.61 (±0.02)
	RVSV				0.52 (±0.04)
	LVSV				0.52 (±0.03)
	LVEDV				0.52 (±0.05)
	LVESSV				0.55 (±0.03)
CMR+VRF					
	Sex				0.62 (±0.03)
	LVM				0.61 (±0.01)
	Hypertension				0.60 (±0.04)
	Age				0.61 (±0.05)
	RVEDV				0.56 (±0.05)
Radiomics					
	Median	First-Order	MYO	ES	0.57 (±0.06)
	Surface area to volume ratio	Shape	MYO	ES	0.64 (±0.02)
	Busyness	First-Order	MYO	ES	0.57 (±0.06)
	Large area low gray level emphasis	Texture	MYO	ES	0.55 (±0.04)
	Gray level non-uniformity	Texture	MYO	ES	0.63 (±0.01)
	Root mean squared	First-Order	MYO	ED	0.54 (±0.05)
	Large area low gray level emphasis	Texture	MYO	ED	0.57 (±0.04)
	Mean	First-Order	MYO	ES	0.55 (±0.06)
	Large dependence low gray level emphasis	Texture	MYO	ED	0.57 (±0.04)
	Sphericity	Shape	LV	ED	0.52 (±0.05)
	Contrast	Texture	MYO	ED	0.52 (±0.03)
	Gray level non-uniformity	Texture	MYO	ES	0.61 (±0.01)
	Difference entropy	Texture	MYO	ED	0.57 (±0.04)
	Energy	First-Order	MYO	ES	0.56 (±0.04)
	Sphericity	Shape	MYO	ES	0.59 (±0.04)
	Joint average	First-Order	MYO	ED	0.58 (±0.03)
	Range	First-Order	MYO	ED	0.56(±0.07)
	Large area emphasis	Texture	MYO	ED	0.60(±0.01)
	Sum entropy	Texture	MYO	ES	0.54(±0.02)
Radiomics+VRF					
	Surface area to volume ratio	Shape	MYO	ES	0.64 (±0.02)
	Median	First-Order	MYO	ES	0.57 (±0.06)
	Busyness	First-Order	MYO	ES	0.57 (±0.02)
	Large area high gray level emphasis	Texture	MYO	ES	0.56 (±0.05)
	Mean	First-Order	MYO	ED	0.54 (±0.01)
	Zone variance	Texture	MYO	ED	0.60 (±0.05)
	Large area emphasis Zone variance	Texture	MYO	ES	0.60 (±0.03)
	Sex				0.59 (±0.04)
	Busyness	Texture	MYO	ED	0.57 (±0.04)
	Gray level non-uniformity	Texture	MYO	ED	0.59 (±0.03)
	Root mean squared	First-Order	MYO	ES	0.55 (±0.04)
	Hypertension				0.55 (±0.03)
	Maximum 2D diameter slice	Shape	MYO	ED	0.55 (±0.03)
	Short run low gray level emphasis	Texture	MYO	ED	0.46(±0.05)
	Long run high gray level emphasis	Texture	MYO	ES	0.45(±0.03)
	Low gray level emphasis	Texture	MYO	ED	0.53(±0.04)

TABLE A.6: Supplementary Table 5: Stroke

CMR + Radiomics					
	Surface area to volume ratio	Shape	MYO	ED	0.64(±0.02)
	Median	First-Order	MYO	ES	0.57(±0.06)
	Busyness	Texture	MYO	ES	0.57(±0.04)
	Large area low gray level emphasis	Texture	MYO	ES	0.55(±0.01)
	Gray level non-uniformity	Texture	MYO	ES	0.63(±0.04)
	Root mean squared	First-Order	MYO	ES	0.54(±0.05)
	Large area low gray level emphasis	Texture	MYO	ED	0.57(±0.04)
	Mean	First-Order	MYO	ES	0.55(±0.06)
	Large dependence low gray level emphasis	Texture	MYO	ED	0.57(±0.04)
	Sphericity	Shape	LV	ED	0.52(±0.05)
	Contrast	Texture	MYO	ED	0.56(±0.03)
	Gray level non-uniformity	Texture	MYO	ES	0.61(±0.01)
	Difference entropy	Texture	MYO	ED	0.57(±0.04)
	Energy	First-Order	MYO	ES	0.48(±0.03)
	Sphericity	Shape	MYO	ES	0.59(±0.04)
	Joint average	First-Order	MYO	ED	0.56(±0.05)
	Range	First-Order	MYO	ED	0.56(±0.07)
	Large area emphasis	Texture	MYO	ED	0.60(±0.01)
	Sum entropy	Texture	MYO	ES	0.54(±0.02)
CMR + Radiomics + VRF					
	Surface area to volume ratio	Shape	MYO	ED	0.64(±0.02)
	Median	First-Order	MYO	ES	0.57(±0.06)
	Busyness	Texture	MYO	ES	0.57(±0.04)
	Age				0.61(±0.05)
	Large area high gray level emphasis	Texture	MYO	ES	0.55(±0.04)
	Mean	First-Order	MYO	ED	0.54(±0.01)
	Zone variance	Texture	MYO	ED	0.60(±0.05)
	Large area emphasis	Texture	MYO	ED	0.60(±0.01)
	Zone variance	Texture	MYO	ES	0.59(±0.05)
	Sex				0.62(±0.03)
	Busyness	Texture	MYO	ED	0.57(±0.04)
	Gray level non-uniformity	Texture	MYO	ED	0.59(±0.02)
	Root mean squared	First-Order	MYO	ES	0.56(±0.05)
	Hypertension				0.55(±0.03)
	Maximum 2D diameter slice	Shape	LV	ED	0.55(±0.03)
	Short run low gray level emphasis	Texture	MYO	ED	0.46(±0.05)
	Long run low gray level emphasis	Texture	MYO	ES	0.45(±0.03)
	Low gray level emphasis	Texture	MYO	ED	0.53(±0.04)

Abbreviations: CMR, cardiac magnetic resonance imaging; VRF, vascular risk factor; ROI, region of interest; SVM model alone: support vector machine model performance showing the mean and standard deviation using each radiomic feature individually; LV, left-ventricle; RV, right-ventricle; MYO, left ventricle myocardium; ED, end-diastolic; EF, ejection fraction; EDV, end-diastolic volume; ESV, end-systolic volume; LVEF, left ventricular ejection fraction; SV, stroke volume.

	VRF	CMR	Radiomics	VRF + CMR	VRF + Radiomics	CMR + Radiomics	VRF + CMR + Radiomics
AF							
Accuracy	0.66(±0.02)	0.65(±0.04)	0.66(±0.07)	0.67(±0.03)	0.72(±0.06)	0.67(±0.07)	0.73(±0.04)
Sensitivity	0.65(±0.06)	0.66(±0.05)	0.66(±0.08)	0.65(±0.07)	0.71(±0.04)	0.66(±0.07)	0.72(±0.06)
Specificity	0.66(±0.05)	0.63(±0.08)	0.65(±0.07)	0.69(±0.04)	0.73(±0.09)	0.67(±0.09)	0.75(±0.05)
AUC	0.69(±0.04)	0.68(±0.06)	0.71(±0.08)	0.73(±0.04)	0.77(±0.06)	0.71(±0.08)	0.77(±0.06)
HF							
Accuracy	0.71(±0.04)	0.69(±0.03)	0.68(±0.05)	0.75(±0.04)	0.77(±0.06)	0.68(±0.06)	0.78(±0.05)
Sensitivity	0.68(±0.04)	0.64(±0.03)	0.66(±0.07)	0.73(±0.08)	0.76(±0.09)	0.66(±0.04)	0.76(±0.08)
Specificity	0.74(±0.07)	0.74(±0.08)	0.70(±0.07)	0.77(±0.1)	0.77(±0.1)	0.71(±0.09)	0.79(±0.05)
AUC	0.78(±0.04)	0.74(±0.03)	0.77(±0.05)	0.84(±0.05)	0.84(±0.05)	0.76(±0.5)	0.85(±0.06)
MI							
Accuracy	0.64(±0.03)	0.63(±0.03)	0.64(±0.05)	0.64(±0.03)	0.67(±0.04)	0.64(±0.05)	0.66(±0.02)
Sensitivity	0.65(±0.03)	0.62(±0.09)	0.64(±0.01)	0.64(±0.09)	0.65(±0.05)	0.64(±0.09)	0.63(±0.03)
Specificity	0.63(±0.06)	0.63(±0.04)	0.64(±0.05)	0.64(±0.04)	0.69(±0.05)	0.64(±0.05)	0.69(±0.06)
AUC	0.71(±0.03)	0.68(±0.04)	0.69(±0.06)	0.70(±0.04)	0.74(±0.03)	0.68(±0.06)	0.74(±0.04)
Stroke							
Accuracy	0.62(±0.07)	0.59(±0.04)	0.63(±0.05)	0.63(±0.03)	0.64(±0.02)	0.63(±0.05)	0.64(±0.04)
Sensitivity	0.62(±0.08)	0.56(±0.05)	0.60(±0.08)	0.60(±0.07)	0.64(±0.07)	0.61(±0.06)	0.61(±0.08)
Specificity	0.62(±0.1)	0.61(±0.1)	0.66(±0.06)	0.64(±0.08)	0.62(±0.08)	0.65(±0.03)	0.66(±0.06)
AUC	0.67(±0.1)	0.62(±0.03)	0.68(±0.07)	0.69(±0.07)	0.69(±0.07)	0.67(±0.07)	0.69(±0.07)

TABLE A.7: Table 6 footnote: Abbreviations: CMR, cardiac magnetic resonance imaging; VRF, vascular risk factor; AF, atrial fibrillation; HF, heart failure; MI, myocardial infarction

Bibliography

"2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy" (Aug. 2014). In: *European Heart Journal* 35.39, pp. 2733–2779. DOI: 10.1093/eurheartj/ehu284. URL: https://doi.org/10.1093/eurheartj/ehu284.

A, Xin et al. (2023). "Non-Contrast Cine Cardiac Magnetic Resonance Derived-Radiomics for the Prediction of Left Ventricular Adverse Remodeling in Patients With ST-Segment Elevation Myocardial Infarction". In: *Korean Journal of Radiology* 24.9, p. 827. DOI: 10.3348/kjr.2023.0061. URL: https://doi.org/10.3348/kjr.2023.0061.

Abbasian Ardakani, Ali et al. (Mar. 2022). "Interpretation of radiomics features–A pictorial review". In: *Computer Methods and Programs in Biomedicine* 215, p. 106609. ISSN: 0169-2607. DOI: 10.1016/j.cmpb.2021.106609. URL: http://dx.doi.org/10.1016/j.cmpb.2021.106609.

Abd ElHafeez, Samar et al. (Nov. 2021). "Methods to Analyze Time-to-Event Data: The Cox Regression Analysis". In: *Oxidative Medicine and Cellular Longevity* 2021. Ed. by Alexandros Georgakilas, 1–6. ISSN: 1942-0900. DOI: 10.1155/2021/1302811. URL: http://dx.doi.org/10.1155/2021/1302811.

Acharya, U. R. et al. (2006). "Heart rate variability: A review". In: *Medical and Biological Engineering and Computing* 44. DOI: 10.1007/s11517-006-0119-0. URL: https://doi.org/10.1007/s11517-006-0119-0.

Aerts, Hugo J. W. L. et al. (June 2014). "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach". In: *Nature Communications* 5.1. DOI: 10.1038/ncomms5006. URL: https://doi.org/10.1038/ncomms5006.

Aizawa, Y., H. Watanabe, and K. Okumura (2017). "Electrocardiogram (ECG) for the prediction of incident atrial fibrillation: An overview". In: *Journal of Atrial Fibrillation* 10. DOI: 10.4022/jafib.1724.

Amano, Yasuo et al. (Sept. 2018). "Relationship between Extension or Texture Features of Late Gadolinium Enhancement and Ventricular Tachyarrhythmias in Hypertrophic Cardiomyopathy". In: *BioMed Research International* 2018, pp. 1–6. DOI: 10.1155/2018/4092469. URL: https://doi.org/10.1155/2018/4092469.

Amzulescu, Mihaela-Silvia et al. (Aug. 2015). "Prognostic Impact of Hypertrabeculation and Noncompaction Phenotype in Dilated Cardiomyopathy". In: *JACC: Cardiovascular Imaging* 8.8, 934–946. ISSN: 1936-878X. DOI: 10.1016/j.jcmg.2015.04.015. URL: http://dx.doi.org/10.1016/j.jcmg.2015.04.015.

Andersen, R. S., A. Peimankar, and S. Puthusserypady (2019). "A deep learning approach for real-time detection of atrial fibrillation". In: *Expert Systems with Applications* 115, pp. 465–473. DOI: 10.1016/j.eswa.2018.08.011.

Andreotti, F. et al. (2017). "Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG". In: *Computing in Cardiology*. Vol. 44, pp. 1–4.

Antonopoulos, AS, M Boutsikou, S Simantiris, et al. (2021). "Machine learning of native T1 mapping radiomics for classification of hypertrophic cardiomyopathy phenotypes". In: *Scientific Reports* 11, pp. 1–11. DOI: 10.1038/s41598-021-02971-z.

Aparicio, HJ, EJ Benjamin, CW Callaway, et al. (2021). "Heart Disease and Stroke Statistics-2021 Update: A Report from the American Heart Association". In: *Circulation*.

Arbelo, Elena et al. (Aug. 2023). "2023 ESC Guidelines for the management of cardiomyopathies". In: *European Heart Journal* 44.37, 3503–3626. ISSN: 1522-9645. DOI: 10.1093/eurheartj/ehad194. URL: http://dx.doi.org/10.1093/eurheartj/ehad194.

Arbustini, Eloisa et al. (Aug. 2016). "Left Ventricular Noncompaction". In: *Journal of the American College of Cardiology* 68.9, 949–966. ISSN: 0735-1097. DOI: 10.1016/j.jacc.2016.05.096. URL: http://dx.doi.org/10.1016/j.jacc.2016.05.096.

Athif, Mohamed, Pamodh Chanuka Yasawardene, and Chathuri Daluwatte (June 2018). "Detecting atrial fibrillation from short single lead ECGs using statistical and morphological features". In: *Physiological Measurement* 39.6, p. 064002. ISSN: 1361-6579. DOI: 10.1088/1361-6579/aac552. URL: http://dx.doi.org/10.1088/1361-6579/aac552.

Attar, R, M Pereañez, A Gooya, et al. (2019). "Quantitative CMR population imaging on 20,000 subjects of the UK Biobank imaging study: LV/RV quantification pipeline and its evaluation". In: *Medical Image Analysis* 56, pp. 26–42. DOI: 10.1016/j.media.2019.05.006.

Attia, Z. I. et al. (2019a). "Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram". In: *Nature Medicine* 25.1, pp. 70–74.

Attia, Zachi I et al. (2019b). "Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram". In: *Nature medicine* 25.1, pp. 70–74.

Aung, Nay et al. (Jan. 2020). "Prognostic Significance of Left Ventricular Noncompaction: Systematic Review and Meta-Analysis of Observational Studies". In: *Circulation: Cardiovascular Imaging* 13.1. ISSN: 1942-0080. DOI: 10.1161/circimaging.119.009712. URL: http://dx.doi.org/10.1161/CIRCIMAGING.119.009712.

Baalman, S. W. E. et al. (2020). "A morphology based deep learning model for atrial fibrillation detection using single cycle electrocardiographic samples". In: *International Journal of Cardiology* 316, pp. 176–183. DOI: 10.1016/j.ijcard.2020.03.031.

Baessler, Bettina et al. (Dec. 2019). "Radiomics allows for detection of benign and malignant histopathology in patients with metastatic testicular germ cell tumors prior to post-chemotherapy retroperitoneal lymph node dissection". In: *European Radiology* 30.4, pp. 2334–2345. DOI: 10.1007/s00330-019-06495-z. URL: https://doi.org/10.1007/s00330-019-06495-z.

Baeßler, B, M Mannil, D Maintz, et al. (2018). "Texture analysis and machine learning of non-contrast T1-weighted MR images in patients with hypertrophic cardiomyopathy—Preliminary results". In: *European Journal of Radiology* 102, pp. 61–67. DOI: 10.1016/j.ejrad.2018.03.013.

Bai, W, M Sinclair, G Tarroni, et al. (2018). "Automated cardiovascular magnetic resonance image analysis with fully convolutional networks". In: *Journal of Cardiovascular Magnetic Resonance* 20, pp. 1–12. DOI: 10.1186/s12968-018-0471-x.

Benítez-Parejo, N., M.M. Rodríguez del Águila, and S. Pérez-Vicente (2011). "Survival analysis and Cox regression". In: *Allergologia et Immunopathologia* 39.6, pp. 362–373. ISSN: 0301-0546. DOI: https://doi.org/10.1016/j.aller.2011.07.007. URL: https://www.sciencedirect.com/science/article/pii/S0301054611002679.

Bera, K, N Braman, A Gupta, et al. (2021). "Predicting cancer outcomes with radiomics and artificial intelligence in radiology". In: *Nature Reviews Clinical Oncology*. DOI: 10.1038/s41571-021-00560-7.

Bertelsen, L. et al. (2020). "Left atrial volume and function assessed by cardiac magnetic resonance imaging are markers of subclinical atrial fibrillation as detected by continuous monitoring". In: *Europace* 22. DOI: 10.1093/europace/euaa252.

Bikkina, M et al. (1995). "Left ventricular mass and risk of stroke in an elderly cohort. The Framingham heart study". In: *JAMA - Journal of the American Medical Association* 43, p. 202. DOI: 10.1111/j.1532-5415.1995.tb06399.x.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Bleumink, GS, AM Knetsch, MCJM Sturkenboom, et al. (2004). "Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime risk and prognosis of heart failure The Rotterdam Study". In: DOI: 10.1016/j.ehj.2004.06.038.

Boriani, G. et al. (2017). "The effects of gender on electrical therapies for the heart: Physiology, epidemiology, and access to therapies". In: *Europace* 19. DOI: 10.1093/europace/eux068.

Bos, J. Martijn et al. (June 2014). "Characterization of a Phenotype-Based Genetic Test Prediction Score for Unrelated Patients With Hypertrophic Cardiomyopathy". In: *Mayo Clinic Proceedings* 89.6, 727–737. ISSN: 0025-6196. DOI: 10.1016/j.mayocp.2014.01.025. URL: http://dx.doi.org/10.1016/j.mayocp.2014.01.025.

Breiman, Leo (2001). In: *Machine Learning* 45.1, pp. 5–32. DOI: 10.1023/a:1010933404324. URL: https://doi.org/10.1023/a:1010933404324.

Brundel, Bianca J. J. M. et al. (Apr. 2022). "Atrial fibrillation". In: *Nature Reviews Disease Primers* 8.1. ISSN: 2056-676X. DOI: 10.1038/s41572-022-00347-9. URL: http://dx.doi.org/10.1038/s41572-022-00347-9.

Buckberg, Gerald et al. (June 2018). "What Is the Heart? Anatomy, Function, Pathophysiology, and Misconceptions". In: *Journal of Cardiovascular Development and Disease* 5.2, p. 33. ISSN: 2308-3425. DOI: 10.3390/jcdd5020033. URL: http://dx.doi.org/10.3390/jcdd5020033.

Bumgarner, J. M. et al. (2018). "Automated Atrial Fibrillation Detection Algorithm Using Smartwatch Technology". In: *Journal of the American College of Cardiology* 71.

Campello, Víctor M et al. (2022). "Minimising multi-centre radiomics variability through image normalisation: a pilot study". In: *Scientific reports* 12.1, p. 12532.

Campello, V.M. et al. (2021). "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&ampMs Challenge". In: *IEEE Transactions on Medical Imaging*, pp. 1–1. DOI: 10.1109/tmi.2021.3090082. URL: https://doi.org/10.1109/tmi.2021.3090082.

Captur, Gabriella et al. (2013). "Quantification of left ventricular trabeculae using fractal analysis". In: *Journal of Cardiovascular Magnetic Resonance* 15.1, p. 36. DOI: 10.1186/1532-429x-15-36. URL: https://doi.org/10.1186/1532-429x-15-36.

Cardenas, Carlos E. et al. (July 2019). "Advances in Auto-Segmentation". In: *Seminars in Radiation Oncology* 29.3, 185–197. ISSN: 1053-4296. DOI: 10.1016/j.semradonc.2019.02.001. URL: http://dx.doi.org/10.1016/j.semradonc.2019.02.001.

Cardim, Nuno et al. (Feb. 2015). "Role of multimodality cardiac imaging in the management of patients with hypertrophic cardiomyopathy: an expert consensus of the European Association of Cardiovascular Imaging Endorsed by the Saudi Heart Association". In: *European Heart Journal – Cardiovascular Imaging* 16.3,

pp. 280–280. DOI: 10.1093/ehjci/jeu291. URL: https://doi.org/10.1093/ehjci/jeu291.

Casas, Guillem, José F. Rodríguez-Palomares, and Ignacio Ferreira-González (Dec. 2022). "Left ventricular noncompaction: a disease or a phenotypic trait?" In: *Revista Española de Cardiología (English Edition)* 75.12, 1059–1069. ISSN: 1885-5857. DOI: 10.1016/j.rec.2022.07.002. URL: http://dx.doi.org/10.1016/j.rec.2022.07.002.

Casas, Guillem et al. (2021). "Clinical Risk Prediction in Patients With Left Ventricular Myocardial Noncompaction". In: *Journal of the American College of Cardiology* 78.7, pp. 643–662. ISSN: 0735-1097. DOI: https://doi.org/10.1016/j.jacc.2021.06.016. URL: https://www.sciencedirect.com/science/article/pii/S0735109721053808.

Caselli, Stefano et al. (Sept. 2015). "Left Ventricular Noncompaction Diagnosis and Management Relevant to Pre-participation Screening of Athletes". In: *The American Journal of Cardiology* 116.5, 801–808. ISSN: 0002-9149. DOI: 10.1016/j.amjcard.2015.05.055. URL: http://dx.doi.org/10.1016/j.amjcard.2015.05.055.

Cawley, Gavin C. and Nicola L. C. Talbot (2010). "On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation". In: *Journal of Machine Learning Research* 11.70, pp. 2079–2107. URL: http://jmlr.org/papers/v11/cawley10a.html.

Cetin, Irem et al. (2018). "A Radiomics Approach to Computer-Aided Diagnosis with Cardiac Cine-MRI". In: *Lecture Notes in Computer Science*. Springer International Publishing, pp. 82–90. DOI: 10.1007/978-3-319-75541-0_9. URL: https://doi.org/10.1007/978-3-319-75541-0_9.

Cetin, Irem et al. (Nov. 2020). "Radiomics Signatures of Cardiovascular Risk Factors in Cardiac MRI: Results From the UK Biobank". In: *Frontiers in Cardiovascular Medicine* 7. DOI: 10.3389/fcvm.2020.591368. URL: https://doi.org/10.3389/fcvm.2020.591368.

Chabner, Bruce A (2016). "Clinical strategies for cancer treatment: the role of drugs". In: *Cancer biology and medicine* 13.4, p. 375.

Chandra, MA and SS Bedi (2021). "Survey on SVM and their application in image classification". In: *International Journal of Information Technology* 13. DOI: 10.1007/s41870-017-0080-1.

Chen, Tianqi and Carlos Guestrin (Aug. 2016). "XGBoost". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. DOI: 10.1145/2939672.2939785. URL: https://doi.org/10.1145/2939672.2939785.

Cheng, Sainan et al. (2018). "LGE-CMR-derived texture features reflect poor prognosis in hypertrophic cardiomyopathy patients with systolic dysfunction: preliminary results". In: *European radiology* 28.11, pp. 4615–4624.

Chugh, Sumeet S. et al. (Feb. 2014). "Worldwide Epidemiology of Atrial Fibrillation: A Global Burden of Disease 2010 Study". In: *Circulation* 129.8, 837–847. ISSN: 1524-4539. DOI: 10.1161/circulationaha.113.005119. URL: http://dx.doi.org/10.1161/CIRCULATIONAHA.113.005119.

Chyou, J. Y. et al. (2015). "Individual and Combined Risk Factors for Incident Atrial Fibrillation and Incident Stroke: An Analysis of 3 Million At-Risk US Patients". In: *Journal of the American Heart Association* 4.

Cortes, Corinna and Vladimir Vapnik (Sept. 1995). "Support-vector networks". In: *Machine Learning* 20.3, pp. 273–297. DOI: 10.1007/bf00994018. URL: https://doi.org/10.1007/bf00994018.

CVI42 Version x.x (Year). https://www.circlecvi.com/cvi42. Accessed: date. Circle Cardiovascular Imaging Inc.

Deo, Rohit C (2015). "Machine learning in medicine". In: *Circulation* 132.20, pp. 1920–1930.

Deo, Salil Vasudeo, Vaishali Deo, and Varun Sundaram (Jan. 2021). "Survival analysis—part 2: Cox proportional hazards model". In: *Indian Journal of Thoracic and Cardiovascular Surgery* 37.2, 229–233. ISSN: 0973-7723. DOI: 10.1007/s12055-020-01108-7. URL: http://dx.doi.org/10.1007/s12055-020-01108-7.

Ding, Haoran et al. (Sept. 2021). "Radiomics in Oncology: A 10-Year Bibliometric Analysis". In: *Frontiers in Oncology* 11. DOI: 10.3389/fonc.2021.689802. URL: https://doi.org/10.3389/fonc.2021.689802.

"Discrete Cosine Transform and Support Vector Machines for Classification Cardiac Atrial Arrhythmia and Cardiac Normal" (2020). In: *International Journal of Emerging Trends in Engineering Research* 8, pp. 5400–5407.

Donal, Erwan et al. (Aug. 2019). "Multimodality imaging in the diagnosis, risk stratification, and management of patients with dilated cardiomyopathies: an expert consensus document from the European Association of Cardiovascular Imaging". In: *European Heart Journal - Cardiovascular Imaging* 20.10, pp. 1075–1093. DOI: 10.1093/ehjci/jez178. URL: https://doi.org/10.1093/ehjci/jez178.

Dwivedi, Girish et al. (Mar. 2013). "Cardiovascular Imaging: New Directions in an Evolving Landscape". In: *Canadian Journal of Cardiology* 29.3, 257–259. ISSN: 0828-282X. DOI: 10.1016/j.cjca.2013.01.011. URL: http://dx.doi.org/10.1016/j.cjca.2013.01.011.

Dörr, M. et al. (2019). "The WATCH AF Trial: SmartWATCHes for Detection of Atrial Fibrillation". In: *JACC: Clinical Electrophysiology* 5.

Ekundayo, OJ et al. (2013). "Isolated Systolic Hypertension and Incident Heart Failure in Older Adults: A Propensity-Matched Study". In: *Hypertension* 23, pp. 1–7. DOI: 10.1161/HYPERTENSIONAHA.108.119792.Isolated.

Elliott, P. et al. (Dec. 2007). "Classification of the cardiomyopathies: a position statement from the european society of cardiology working group on myocardial and pericardial diseases". In: *European Heart Journal* 29.2, pp. 270–276. DOI: 10.1093/eurheartj/ehm342. URL: https://doi.org/10.1093/eurheartj/ehm342.

euCanSHare (n.d.). https://github.com/euCanSHare/dicom2nitfi.

Fan, X. et al. (2018). "Multiscaled Fusion of Deep Convolutional Neural Networks for Screening Atrial Fibrillation from Single Lead Short ECG Recordings". In: *IEEE Journal of Biomedical and Health Informatics* 22.6, pp. 1744–1753. DOI: 10.1109/JBHI.2018.2832530.

Faust, O. et al. (2018). "Automated detection of atrial fibrillation using long short-term memory network with RR interval signals". In: *Computers in Biology and Medicine* 102, pp. 327–335. DOI: 10.1016/j.compbiomed.2018.09.009.

Flora, Gagan D. and Manasa K. Nayak (Dec. 2019). "A Brief Review of Cardiovascular Diseases, Associated Risk Factors and Current Treatment Regimes". In: *Current Pharmaceutical Design* 25.38, 4063–4084. ISSN: 1381-6128. DOI: 10.2174/1381612825666190925163827. URL: http://dx.doi.org/10.2174/1381612825666190925163827.

Flueckiger, P et al. (2018). "Revised framingham stroke risk score, nontraditional risk markers, and incident stroke in a multiethnic cohort". In: *Stroke* 49, pp. 363–369. DOI: 10.1161/STROKEAHA.117.018928.

Flynn, Robert (Aug. 2012). "Survival analysis". In: *Journal of Clinical Nursing* 21.19pt20, 2789–2797. ISSN: 1365-2702. DOI: 10.1111/j.1365-2702.2011.04023.x. URL: http://dx.doi.org/10.1111/j.1365-2702.2011.04023.x.

Fuchs, FD and PK Whelton (2020). "High Blood Pressure and Cardiovascular Disease". In: *Hypertension*, pp. 285–292. DOI: 10.1161/HYPERTENSIONAHA.119.14240.

Fujikura, Kana (2022). "Cardiovascular Imaging: Opportunities and Challenges in Clinical Application". In: *Journal of Clinical Medicine* 2022.Special Issue. URL: https://www.mdpi.com/journal/jcm/special_issues/Cardiovascular_Imaging.

Fuster, Valentin et al. (Aug. 2006). "ACC/AHA/ESC 2006 Guidelines for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation):Developed in Collaboration With the European Heart Rhythm Association and the Heart Rhythm Society". In: *Circulation* 114.7. ISSN: 1524-4539. DOI: 10.1161/circulationaha.106.177292. URL: http://dx.doi.org/10.1161/CIRCULATIONAHA.106.177292.

Gati, Sabiha et al. (Feb. 2013). "Increased left ventricular trabeculation in highly trained athletes: do we need more stringent criteria for the diagnosis of left ventricular non-compaction in athletes?" In: *Heart* 99.6, 401–408. ISSN: 1468-201X. DOI: 10.1136/heartjnl-2012-303418. URL: http://dx.doi.org/10.1136/heartjnl-2012-303418.

Gati, Sabiha et al. (2014). "Reversible de novo left ventricular trabeculations in pregnant women: Implications for the diagnosis of left ventricular noncompaction in low-risk populations". In: *Circulation* 130, pp. 475–483. DOI: 10.1161/circulationaha.114.008554.

Gawande, Atul (2014). *Being mortal: medicine and what matters in the end*. Metropolitan Books.

Geelhoed, B. et al. (2020). "Assessment of causality of natriuretic peptides and atrial fibrillation and heart failure: A Mendelian randomization study in the FINRISK cohort". In: *Europace* 22.

Gillies, Robert J., Paul E. Kinahan, and Hedvig Hricak (Feb. 2016). "Radiomics: Images Are More than Pictures, They Are Data". In: *Radiology* 278.2, pp. 563–577. DOI: 10.1148/radiol.2015151169. URL: https://doi.org/10.1148/radiol.2015151169.

Gopinathannair, R. et al. (2020). "COVID-19 and cardiac arrhythmias: a global perspective on arrhythmia characteristics and management strategies". In: *Journal of Interventional Cardiac Electrophysiology* 59.

Gosmanova, EO, MK Mikkelsen, MZ Molnar, et al. (2016). "Association of Systolic Blood Pressure Variability With Mortality, Coronary Heart Disease, Stroke, and Renal Disease". In: *Journal of the American College of Cardiology* 68, pp. 1375–1386. DOI: 10.1016/j.jacc.2016.06.054.

Griethuysen, Joost J.M. van et al. (Oct. 2017a). "Computational Radiomics System to Decode the Radiographic Phenotype". In: *Cancer Research* 77.21, e104–e107. DOI: 10.1158/0008-5472.can-17-0339. URL: https://doi.org/10.1158/0008-5472.can-17-0339.

— (Oct. 2017b). "Computational Radiomics System to Decode the Radiographic Phenotype". In: *Cancer Research* 77.21, e104–e107. ISSN: 1538-7445. DOI: 10.1158/0008-5472.can-17-0339. URL: http://dx.doi.org/10.1158/0008-5472.CAN-17-0339.

Grigoratos, Chrysanthos et al. (Nov. 2019). "Meta-Analysis of the Prognostic Role of Late Gadolinium Enhancement and Global Systolic Impairment in Left Ventricular Noncompaction". In: *JACC: Cardiovascular Imaging* 12.11, 2141–2151. ISSN:

1936-878X. DOI: 10.1016/j.jcmg.2018.12.029. URL: http://dx.doi.org/10.1016/j.jcmg.2018.12.029.

Grothues, Frank et al. (2002). "Comparison of interstudy reproducibility of cardiovascular magnetic resonance with two-dimensional echocardiography in normal subjects and in patients with heart failure or left ventricular hypertrophy". In: *American Journal of Cardiology* 90.1, pp. 29–34.

Gruner, Christiane et al. (Feb. 2013). "Toronto Hypertrophic Cardiomyopathy Genotype Score for Prediction of a Positive Genotype in Hypertrophic Cardiomyopathy". In: *Circulation: Cardiovascular Genetics* 6.1, 19–26. ISSN: 1942-3268. DOI: 10.1161/circgenetics.112.963363. URL: http://dx.doi.org/10.1161/CIRCGENETICS.112.963363.

Guo, Y. et al. (2019). "Mobile Photoplethysmographic Technology to Detect Atrial Fibrillation". In: *Journal of the American College of Cardiology* 74.

Hagiwara, Yuki et al. (Oct. 2018). "Computer-aided diagnosis of atrial fibrillation based on ECG Signals: A review". In: *Information Sciences* 467, 99–114. ISSN: 0020-0255. DOI: 10.1016/j.ins.2018.07.063. URL: http://dx.doi.org/10.1016/j.ins.2018.07.063.

Han, Pei-Lun et al. (Oct. 2023). "Prognostic prediction of left ventricular myocardial noncompaction using machine learning and cardiac magnetic resonance radiomics". In: *Quantitative Imaging in Medicine and Surgery* 13.10, 6468–6481. ISSN: 2223-4306. DOI: 10.21037/qims-23-372. URL: http://dx.doi.org/10.21037/qims-23-372.

Henriques, T. et al. (2020). "Nonlinear methods most applied to heart-rate time series: A review". In: *Entropy* 22. DOI: 10.3390/e22030309.

Hershberger, Ray E., Dale J. Hedges, and Ana Morales (July 2013). "Dilated cardiomyopathy: the complexity of a diverse genetic architecture". In: *Nature Reviews Cardiology* 10.9, 531–547. ISSN: 1759-5010. DOI: 10.1038/nrcardio.2013.105. URL: http://dx.doi.org/10.1038/nrcardio.2013.105.

Himmelreich, JCL, L Veelers, WAM Lucassen, et al. (2020). "Prediction models for atrial fibrillation applicable in the community: A systematic review and meta-analysis". In: *Europace* 22, pp. 684–694. DOI: 10.1093/europace/euaa005.

Hossain, M. B. et al. (2019). "An Accurate QRS Complex and P Wave Detection in ECG Signals Using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise Approach". In: *IEEE Access* 7.

Howlett, P. J. et al. (2015). "Diagnosing Paroxysmal Atrial Fibrillation: Are Biomarkers the Solution to This Elusive Arrhythmia?" In: *BioMed Research International* 2015, 1–10. ISSN: 2314-6141. DOI: 10.1155/2015/910267. URL: http://dx.doi.org/10.1155/2015/910267.

Hsieh, Chaur-Heh et al. (Apr. 2020). "Detection of Atrial Fibrillation Using 1D Convolutional Neural Network". In: *Sensors* 20.7, p. 2136. ISSN: 1424-8220. DOI: 10.3390/s20072136. URL: http://dx.doi.org/10.3390/s20072136.

Iizuka, N. et al. (2003). "Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection". In: *Lancet* 361.

Inohara, T. et al. (2018). "Association of atrial fibrillation clinical phenotypes with treatment patterns and outcomes: a multicenter registry study". In: *JAMA Cardiology* 3.

Ivanov, Alexander et al. (Sept. 2017). "Prevalence and Prognostic Significance of Left Ventricular Noncompaction in Patients Referred for Cardiac Magnetic Resonance Imaging". In: *Circulation: Cardiovascular Imaging* 10.9. DOI: 10.1161/circimaging.117.006174. URL: https://doi.org/10.1161/circimaging.117.006174.

Izquierdo, Cristian et al. (Oct. 2021). "Radiomics-Based Classification of Left Ventricular Non-compaction, Hypertrophic Cardiomyopathy, and Dilated Cardiomyopathy in Cardiovascular Magnetic Resonance". In: *Frontiers in Cardiovascular Medicine* 8. DOI: 10.3389/fcvm.2021.764312. URL: https://doi.org/10.3389/fcvm.2021.764312.

Jacquier, A. et al. (Jan. 2010). "Measurement of trabeculated left ventricular mass using cardiac magnetic resonance imaging in the diagnosis of left ventricular non-compaction". In: *European Heart Journal* 31.9, pp. 1098–1104. DOI: 10.1093/eurheartj/ehp595. URL: https://doi.org/10.1093/eurheartj/ehp595.

Jenni, R, E N Oechslin, and B van der Loo (Jan. 2007). "Isolated ventricular non-compaction of the myocardium in adults". In: *Heart* 93.1, pp. 11–15. DOI: 10.1136/hrt.2005.082271. URL: https://doi.org/10.1136/hrt.2005.082271.

Jin, Qiyu (2023). "Radiomics and Its Feature Selection: A Review". In: *Symmetry* 15.10, p. 1834. DOI: 10.3390/sym15101834.

Kang, Cyra-Yoonsun and Joo Heung Yoon (May 2023). "Current challenges in adopting machine learning to critical care and emergency medicine". In: *Clinical and Experimental Emergency Medicine* 10.2, 132–137. ISSN: 2383-4625. DOI: 10.15441/ceem.23.041. URL: http://dx.doi.org/10.15441/ceem.23.041.

Kassim, Natasha A. et al. (Jan. 2017). "Gender differences in management and clinical outcomes of atrial fibrillation patients". In: *Journal of Cardiology* 69.1, 195–200. ISSN: 0914-5087. DOI: 10.1016/j.jjcc.2016.02.022. URL: http://dx.doi.org/10.1016/j.jjcc.2016.02.022.

Katzman, Jared L. et al. (Feb. 2018). "DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network". In: *BMC Medical Research Methodology* 18.1. DOI: 10.1186/s12874-018-0482-1. URL: https://doi.org/10.1186/s12874-018-0482-1.

Kim, Raymond J et al. (2009). "Late gadolinium-enhanced cardiac magnetic resonance imaging in cardiovascular disease: Relevance and future implications". In: *JACC: Cardiovascular Imaging* 2.5, pp. 535–544.

Kim, TH, CY Shim, JH Park, et al. (2016). "Left ventricular diastolic dysfunction is associated with atrial remodeling and risk or presence of stroke in patients with paroxysmal atrial fibrillation". In: *Journal of Cardiology* 68, pp. 104–109. DOI: 10.1016/j.jjcc.2015.10.008.

Kiranyaz, S., T. Ince, and M. Gabbouj (2016). "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks". In: *IEEE Transactions on Biomedical Engineering* 63.3, pp. 664–675. DOI: 10.1109/TBME.2015.2468589.

Kishi, S. et al. (2015). "Race-ethnic and sex differences in left ventricular structure and function: the Coronary Artery Risk Development in Young Adults (CARDIA) Study". In: *Journal of the American Heart Association* 4.

Ko, Darae et al. (Apr. 2016). "Atrial fibrillation in women: epidemiology, pathophysiology, presentation, and prognosis". In: *Nature Reviews Cardiology* 13.6, 321–332. ISSN: 1759-5010. DOI: 10.1038/nrcardio.2016.45. URL: http://dx.doi.org/10.1038/nrcardio.2016.45.

Kolossváry, Márton et al. (Dec. 2017). "Radiomic Features Are Superior to Conventional Quantitative Computed Tomographic Metrics to Identify Coronary Plaques With Napkin-Ring Sign". In: *Circulation: Cardiovascular Imaging* 10.12. DOI: 10.1161/circimaging.117.006843. URL: https://doi.org/10.1161/circimaging.117.006843.

Kolossváry, M et al. (2018). "Cardiac Computed Tomography Radiomics". In: *Journal of Thoracic Imaging* 33, pp. 26–34. DOI: 10.1097/RTI.0000000000000268.

Koskinas, Konstantinos C. et al. (Dec. 2015). "Intracoronary imaging of coronary atherosclerosis: validation for diagnosis, prognosis and treatment". In: *European Heart Journal* 37.6, pp. 524–535. DOI: 10.1093/eurheartj/ehv642. URL: https://doi.org/10.1093/eurheartj/ehv642.

Krittanawong, Chayakrit et al. (2017). "Artificial intelligence in precision cardiovascular medicine". In: *Journal of the American College of Cardiology* 69.21, pp. 2657–2664.

Kudo, M and J Sklansky (2000). "Comparison of algorithms that select features for pattern classifiers". In: *Pattern Recognition* 33. DOI: 10.1016/S0031-3203(99)00041-2.

Lam, CSP, C Arnott, AL Beale, et al. (2019). "Sex differences in heart failure". In: DOI: 10.1093/eurheartj/ehz835.

Lambin, Philippe et al. (2017). "Radiomics: the bridge between medical imaging and personalized medicine". In: *Nature Reviews Clinical Oncology* 14.12, pp. 749–762.

Larroza, Andrés et al. (Feb. 2018). "Texture analysis of cardiac cine magnetic resonance imaging to detect nonviable segments in patients with chronic myocardial infarction". In: *Medical Physics* 45.4, 1471–1480. ISSN: 2473-4209. DOI: 10.1002/mp.12783. URL: http://dx.doi.org/10.1002/mp.12783.

Leiner, T, D Rueckert, A Suinesiaputra, et al. (2019). "Machine learning in cardiovascular magnetic resonance: Basic concepts and applications". In: *Journal of Cardiovascular Magnetic Resonance* 21.

Lennon, Robert P., Keith A. Claussen, and Karl A. Kuersteiner (Mar. 2018). "State of the Heart". In: *Primary Care: Clinics in Office Practice* 45.1, 1–15. ISSN: 0095-4543. DOI: 10.1016/j.pop.2017.11.001. URL: http://dx.doi.org/10.1016/j.pop.2017.11.001.

Lip, G.Y.H. (Dec. 2001). "Paroxysmal atrial fibrillation". In: *QJM* 94.12, 665–678. ISSN: 1460-2393. DOI: 10.1093/qjmed/94.12.665. URL: http://dx.doi.org/10.1093/qjmed/94.12.665.

Litjens, G. et al. (2017). "A survey on deep learning in medical image analysis". In: *Medical image analysis* 42, pp. 60–88.

Littlejohns, TJ, J Holliday, LM Gibson, et al. (2020). "The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions". In: *Nature Communications* 11, pp. 1–12. DOI: 10.1038/s41467-020-15948-9.

Liu, Yang et al. (2020). "How to read articles that use machine learning: users' guides to the medical literature". In: *JAMA* 322.18, pp. 1806–1816.

Lloyd-Jones, DM, MG Larson, EP Leip, et al. (2002). "Lifetime risk for developing congestive heart failure: The Framingham Heart Study". In: *Circulation* 106, pp. 3068–3072. DOI: 10.1161/01.CIR.0000039105.49749.6F.

Luzum, Jasmine A and Edward Peterson (2016). "Pharmacogenomics of heart failure: a systematic review". In: *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 36.7, pp. 735–748.

Ma, Q, Y Ma, X Wang, et al. (2021). "A radiomic nomogram for prediction of major adverse cardiac events in ST-segment elevation myocardial infarction". In: *European Radiology* 31, pp. 1140–1150. DOI: 10.1007/s00330-020-07176-y.

Maceira, Alicia M et al. (2006). "Normalized left ventricular systolic and diastolic function by steady state free precession cardiovascular magnetic resonance". In: *Journal of Cardiovascular Magnetic Resonance* 8.3, pp. 417–426.

Magnussen, C. et al. (2017). "Sex Differences and Similarities in Atrial Fibrillation Epidemiology, Risk Factors, and Mortality in Community Cohorts". In: *Circulation* 136.

Maron, Barry J (1997). "Hypertrophic cardiomyopathy". In: *The Lancet* 350.9071, pp. 127–133.

Martin-Isla, Carlos et al. (Jan. 2020). "Image-Based Cardiac Diagnosis With Machine Learning: A Review". In: *Frontiers in Cardiovascular Medicine* 7. DOI: 10.3389/fcvm.2020.00001. URL: https://doi.org/10.3389/fcvm.2020.00001.

Martín-Isla, Carlos et al. (2023). "Deep learning segmentation of the right ventricle in cardiac mri: The m&ms challenge". In: *IEEE Journal of Biomedical and Health Informatics* 27.7, pp. 3302–3313.

Martinez-Alanis, M. et al. (2020). "Prediction of sudden cardiac death risk with a support vector machine based on heart rate variability and heartprint indices". In: *Sensors* 20.

McHugh, Mary L. (2013). "The Chi-square test of independence". In: *Biochemia Medica*, 143–149. ISSN: 1846-7482. DOI: 10.11613/bm.2013.018. URL: http://dx.doi.org/10.11613/BM.2013.018.

McManus, D. D. et al. (2010). "Longitudinal tracking of left atrial diameter over the adult life course: Clinical correlates in the community". In: *Circulation* 121.

Minnis, Meghan and Terence J. Quinn (Mar. 2024). "Stroke Literature Synopsis (Clinical)". In: *Stroke* 55.3. ISSN: 1524-4628. DOI: 10.1161/strokeaha.124.046105. URL: http://dx.doi.org/10.1161/STROKEAHA.124.046105.

Morita, Sae X. et al. (Aug. 2021). "Deep Learning Analysis of Echocardiographic Images to Predict Positive Genotype in Patients With Hypertrophic Cardiomyopathy". In: *Frontiers in Cardiovascular Medicine* 8. ISSN: 2297-055X. DOI: 10.3389/fcvm.2021.669860. URL: http://dx.doi.org/10.3389/fcvm.2021.669860.

Mozaffarian, Dariush et al. (2016). "Heart disease and stroke statistics—2016 update". In: *Circulation* 133.4, e38–e360.

Murat, F. et al. (2021). "Review of deep learning-based atrial fibrillation detection studies". In: *International Journal of Environmental Research and Public Health* 18. DOI: 10.3390/ijerph182111302.

Najjar, Reabal (2023). "Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging". In: *Diagnostics* 13.17, p. 2760. DOI: 10.3390/diagnostics13172760.

Neisius, Ulf et al. (Oct. 2019). "Radiomic Analysis of Myocardial Native T1 Imaging Discriminates Between Hypertensive Heart Disease and Hypertrophic Cardiomyopathy". In: *JACC: Cardiovascular Imaging* 12.10, pp. 1946–1954. DOI: 10.1016/j.jcmg.2018.11.024. URL: https://doi.org/10.1016/j.jcmg.2018.11.024.

Neisius, Ulf et al. (Jan. 2020). "Texture signatures of native myocardial Tsub1/sub as novel imaging markers for identification of hypertrophic cardiomyopathy patients without scar". In: *Journal of Magnetic Resonance Imaging* 52.3, pp. 906–919. DOI: 10.1002/jmri.27048. URL: https://doi.org/10.1002/jmri.27048.

Noble, WS (2006). "What is a support vector machine?" In: *Nature Biotechnology* 24.

Obermeyer, Ziad and Ezekiel J Emanuel (2016). "Predicting the Future — Big Data, Machine Learning, and Clinical Medicine". In: *The New England Journal of Medicine* 375.13, pp. 1216–1219.

Oechslin, Erwin and Rolf Jenni (Feb. 2018). "Left Ventricular Noncompaction". In: *Journal of the American College of Cardiology* 71.7, pp. 723–726. DOI: 10.1016/j.jacc.2017.12.031. URL: https://doi.org/10.1016/j.jacc.2017.12.031.

Oechslin, Erwin and Sabine Klaassen (Apr. 2019). "Left Ventricular Noncompaction". In: *Journal of the American College of Cardiology* 73.13, pp. 1612–1615. DOI: 10.1016/j.jacc.2018.11.064. URL: https://doi.org/10.1016/j.jacc.2018.11.064.

Organization, World Health (2019). *Cardiovascular diseases (CVDs)*. Tech. rep. World Health Organization.

— (2021). "Cardiovascular diseases (CVDs)". In: Accessed: 2024-04-17. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

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

Pellman, Jason and Farah Sheikh (Mar. 2015). *Atrial Fibrillation: Mechanisms, Therapeutics, and Future Directions*. DOI: 10.1002/cphy.c140047. URL: http://dx.doi.org/10.1002/cphy.c140047.

Pennell, D J (2004). "Cardiovascular magnetic resonance". In: *Circulation* 109.21, pp. 2614–2631.

Perez, M. V. et al. (2019). "Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation". In: *New England Journal of Medicine* 381.

Petersen, S. E. et al. (2017). "Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort". In: *Journal of Cardiovascular Magnetic Resonance* 19, pp. 1–19. DOI: 10.1186/s12968-017-0327-9.

Petersen, SE, PM Matthews, JM Francis, et al. (2016). "UK Biobank's cardiovascular magnetic resonance protocol". In: *Journal of Cardiovascular Magnetic Resonance* 18. DOI: 10.1186/s12968-016-0227-4.

Petersen, Steffen E. et al. (July 2005). "Left Ventricular Non-Compaction". In: *Journal of the American College of Cardiology* 46.1, pp. 101–105. DOI: 10.1016/j.jacc.2005.03.045. URL: https://doi.org/10.1016/j.jacc.2005.03.045.

Petersen, Steffen E. et al. (Mar. 2023). "Excessive Trabeculation of the Left Ventricle". In: *JACC: Cardiovascular Imaging* 16.3, 408–425. ISSN: 1936-878X. DOI: 10.1016/j.jcmg.2022.12.026. URL: http://dx.doi.org/10.1016/j.jcmg.2022.12.026.

Pinto, Yigal M. et al. (Jan. 2016). "Proposal for a revised definition of dilated cardiomyopathy, hypokinetic non-dilated cardiomyopathy, and its implications for clinical practice: a position statement of the ESC working group on myocardial and pericardial diseases". In: *European Heart Journal* 37.23, pp. 1850–1858. DOI: 10.1093/eurheartj/ehv727. URL: https://doi.org/10.1093/eurheartj/ehv727.

Pinto-Coelho, Luís (2023). "How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications". In: *Bioengineering* 10.12, p. 1435. DOI: 10.3390/bioengineering10121435.

Polidori, Tiziano et al. (June 2023). "Radiomics applications in cardiac imaging: a comprehensive review". In: *La radiologia medica* 128.8, pp. 922–933. DOI: 10.1007/s11547-023-01658-x. URL: https://doi.org/10.1007/s11547-023-01658-x.

Pölsterl, Sebastian (2020). "scikit-survival: A Library for Time-to-Event Analysis Built on Top of scikit-learn". In: *Journal of Machine Learning Research* 21.212, pp. 1–6. URL: http://jmlr.org/papers/v21/20-729.html.

Ponikowski, Piotr et al. (May 2016). "2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC". In: *European Heart Journal* 37.27, 2129–2200. ISSN: 1522-9645. DOI: 10.1093/eurheartj/ehw128. URL: http://dx.doi.org/10.1093/eurheartj/ehw128.

Probst, Philipp, Anne-Laure Boulesteix, and Bernd Bischl (Jan. 2019). "Tunability: Importance of Hyperparameters of Machine Learning Algorithms". In: *J. Mach. Learn. Res.* 20.1, 1934–1965. ISSN: 1532-4435.

Protonotarios, Alexandros and Perry M Elliott (June 2019). "Left ventricular non-compaction: have we reached the limits of conventional imaging?" In: *European Heart Journal* 41.14, 1437–1438. ISSN: 1522-9645. DOI: 10.1093/eurheartj/ehz352. URL: http://dx.doi.org/10.1093/eurheartj/ehz352.

Pu, Cailing et al. (Nov. 2022). "Identification of fibrosis in hypertrophic cardiomyopathy: a radiomic study on cardiac magnetic resonance cine imaging". In: *European Radiology* 33.4, pp. 2301–2311. DOI: 10.1007/s00330-022-09217-0. URL: https://doi.org/10.1007/s00330-022-09217-0.

Pujadas, E. Ruiz et al. (2020). "Prediction of incident cardiovascular events using machine learning and CMR radiomics". In: *Radiology*.

"Radiomics in Cardiovascular Imaging" (2023). In: *Frontiers Research Topic*. URL: https://www.frontiersin.org/research-topics/16868/radiomics-in-cardiovascular-imaging.

Raisi-Estabragh, Z. et al. (2021a). "Left atrial structure and function are associated with cardiovascular outcomes independent of left ventricular measures: a UK Biobank CMR study". In: *European Heart Journal - Cardiovascular Imaging*. DOI: 10.1093/ehjci/jeab266.

Raisi-Estabragh, Zahra et al. (2020a). "Cardiac magnetic resonance radiomics: basic principles and clinical perspectives". In: *European Heart Journal-Cardiovascular Imaging* 21.4, pp. 349–356.

— (Mar. 2020b). "Cardiac magnetic resonance radiomics: basic principles and clinical perspectives". In: *European Heart Journal - Cardiovascular Imaging* 21.4, pp. 349–356. DOI: 10.1093/ehjci/jeaa028. URL: https://doi.org/10.1093/ehjci/jeaa028.

Raisi-Estabragh, Zahra et al. (Dec. 2020c). "Repeatability of Cardiac Magnetic Resonance Radiomics: A Multi-Centre Multi-Vendor Test-Retest Study". In: *Frontiers in Cardiovascular Medicine* 7. ISSN: 2297-055X. DOI: 10.3389/fcvm.2020.586236. URL: http://dx.doi.org/10.3389/fcvm.2020.586236.

Raisi-Estabragh, Zahra et al. (May 2021b). "Associations of cognitive performance with cardiovascular magnetic resonance phenotypes in the UK Biobank". In: *European Heart Journal - Cardiovascular Imaging* 23.5, 663–672. ISSN: 2047-2412. DOI: 10.1093/ehjci/jeab075. URL: http://dx.doi.org/10.1093/ehjci/jeab075.

Rajkomar, Alvin, Jeffrey Dean, and Isaac Kohane (2019). "Machine learning in medicine". In: *New England Journal of Medicine* 380.14, pp. 1347–1358.

Raschka, S. (2018). "Model evaluation, model selection, and algorithm selection in machine learning". In: *ArXiv*. Preprint at https://arxiv.org/abs/1811.12808.

Rathore, V (2018). "Risk Factors of Acute Myocardial Infarction: A Review". In: *Eurasian Journal of Medical Investigation* 2, pp. 1–7. DOI: 10.14744/ejmi.2018.76486.

Rauseo, Elisa et al. (Sept. 2021). "New Imaging Signatures of Cardiac Alterations in Ischaemic Heart Disease and Cerebrovascular Disease Using CMR Radiomics". In: *Frontiers in Cardiovascular Medicine* 8. DOI: 10.3389/fcvm.2021.716577. URL: https://doi.org/10.3389/fcvm.2021.716577.

Richards, Sue et al. (May 2015). "Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology". In: *Genetics in Medicine* 17.5, 405–424. ISSN: 1098-3600. DOI: 10.1038/gim.2015.30. URL: http://dx.doi.org/10.1038/gim.2015.30.

Rizwan, A. et al. (2021). "A Review on the State of the Art in Atrial Fibrillation Detection Enabled by Machine Learning". In: *IEEE Reviews in Biomedical Engineering* 14. DOI: 10.1109/RBME.2020.3024984.

Rizzo, Stefania et al. (Nov. 2018). "Radiomics: the facts and the challenges of image analysis". In: *European Radiology Experimental* 2.1. DOI: 10.1186/s41747-018-0068-z. URL: https://doi.org/10.1186/s41747-018-0068-z.

Roden, Dan M and Julie A Johnson (2016). "Cardiovascular pharmacogenomics". In: *Circulation research* 119.5, pp. 772–793.

Roth, Gregory A et al. (2017). "Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015". In: *Journal of the American College of Cardiology* 70.1, pp. 1–25.

Russell, S. and P. Norvig (2016). *Artificial Intelligence: A Modern Approach*. Pearson.

Russell, Stuart and Peter Norvig (2020). *Artificial Intelligence: A Modern Approach*. Pearson.

S. Luz, E. J. da et al. (2016). "ECG-based heartbeat classification for arrhythmia detection: A survey". In: *Computers Methods and Programs in Biomedicine* 127, pp. 144–164. DOI: 10.1016/j.cmpb.2015.12.008.

Sahle, BW et al. (2017). "Risk Prediction Models for Incident Heart Failure: A Systematic Review of Methodology and Model Performance". In: *Journal of Cardiac Failure* 23, pp. 680–687. DOI: 10.1016/j.cardfail.2017.03.005.

Schlett, Christopher L. et al. (Sept. 2013). "Histogram Analysis of Lipid-Core Plaques in Coronary Computed Tomographic Angiography". In: *Investigative Radiology* 48.9, pp. 646–653. DOI: 10.1097/rli.0b013e31828fdf9f. URL: https://doi.org/10.1097/rli.0b013e31828fdf9f.

Schnabel, RB, LM Sullivan, D Levy, et al. (2008). "Development of a Risk Score for Incident Atrial Fibrillation in the Community; the Framingham Heart Study". In: *Circulation* 118, pp. 739–745. DOI: 10.1161/circ.118.suppl_18.s_1089-c.

Schofield, R. et al. (Feb. 2019). "Texture analysis of cardiovascular magnetic resonance cine images differentiates aetiologies of left ventricular hypertrophy". In: *Clinical Radiology* 74.2, pp. 140–149. DOI: 10.1016/j.crad.2018.09.016. URL: https://doi.org/10.1016/j.crad.2018.09.016.

Schultheiss, Heinz-Peter et al. (May 2019). "Dilated cardiomyopathy". In: *Nature Reviews Disease Primers* 5.1. ISSN: 2056-676X. DOI: 10.1038/s41572-019-0084-1. URL: http://dx.doi.org/10.1038/s41572-019-0084-1.

Schulz-Menger, J, DA Bluemke, J Bremerich, et al. (2020). "Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update: Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing". In: *Journal of Cardiovascular Magnetic Resonance* 22. DOI: 10.1186/s12968-020-00610-6.

Schwab, P. et al. (2017). "Beat by beat: Classifying cardiac arrhythmias with recurrent neural networks". In: *Computing in Cardiology*. Vol. 44.

Sekaran, NK, AL Crowley, FR de Souza, et al. (2017). "The Role for Cardiovascular Remodeling in Cardiovascular Outcomes". In: *Current Atherosclerosis Reports* 19.

Sendak, Mark et al. (Jan. 2019). "Machine Learning in Health Care: A Critical Appraisal of Challenges and Opportunities". In: *eGEMs (Generating Evidence amp; Methods to improve patient outcomes)* 7.1, p. 1. ISSN: 2327-9214. DOI: 10.5334/egems.287. URL: http://dx.doi.org/10.5334/egems.287.

Shaffer, F. and J. P. Ginsberg (2017). "An Overview of Heart Rate Variability Metrics and Norms". In: *Frontiers in Public Health* 5. DOI: 10.3389/fpubh.2017.00258.

Shur, Joshua D. et al. (Oct. 2021). "Radiomics in Oncology: A Practical Guide". In: *RadioGraphics* 41.6, pp. 1717–1732. DOI: 10.1148/rg.2021210037. URL: https://doi.org/10.1148/rg.2021210037.

Smoleń, D. (2017). "Atrial fibrillation detection using boosting and stacking ensemble". In: *Computing in Cardiology*. Vol. 44.

Sohns, C. and N. F. Marrouche (2020). "Atrial fibrillation and cardiac fibrosis". In: *European Heart Journal* 41. DOI: 10.1093/eurheartj/ehz786.

Somani, Sulaiman et al. (Feb. 2021). "Deep learning and the electrocardiogram: review of the current state-of-the-art". In: *EP Europace* 23.8, 1179–1191. ISSN: 1532-2092. DOI: 10.1093/europace/euaa377. URL: http://dx.doi.org/10.1093/europace/euaa377.

Sposato, Luciano A. et al. (Mar. 2022). "Atrial Fibrillation Detected After Stroke and Transient Ischemic Attack: A Novel Clinical Concept Challenging Current Views". In: *Stroke* 53.3. ISSN: 1524-4628. DOI: 10.1161/strokeaha.121.034777. URL: http://dx.doi.org/10.1161/STROKEAHA.121.034777.

Stel, Vianda S. et al. (June 2011). "Survival Analysis I: The Kaplan-Meier Method". In: *Nephron Clinical Practice* 119.1, c83–c88. ISSN: 1660-2110. DOI: 10.1159/000324758. URL: http://dx.doi.org/10.1159/000324758.

Stewart, MH, CJ Lavie, S Shah, et al. (2018). "Prognostic Implications of Left Ventricular Hypertrophy". In: *Progress in Cardiovascular Diseases* 61, pp. 446–455. DOI: 10.1016/j.pcad.2018.11.002.

Stirrat, John and James A. White (Mar. 2013). "The Prognostic Role of Late Gadolinium Enhancement Magnetic Resonance Imaging in Patients With Cardiomyopathy". In: *Canadian Journal of Cardiology* 29.3, 329–336. ISSN: 0828-282X. DOI: 10.1016/j.cjca.2012.11.033. URL: http://dx.doi.org/10.1016/j.cjca.2012.11.033.

Sutton, R. S. and A. G. Barto (2018). *Reinforcement Learning: An Introduction*. MIT Press.

Thygesen, Kristian et al. (2007). "Universal definition of myocardial infarction". In: *Circulation* 116.22, pp. 2634–2653.

Tian, H, J Cui, C Yang, et al. (2018). "Left ventricular remodeling in hypertrophic cardiomyopathy patients with atrial fibrillation". In: *BMC Cardiovascular Disorders* 18, pp. 7–12. DOI: 10.1186/s12872-018-0945-7.

Torrent-Guasp, Francisco et al. (Feb. 2005). "Towards new understanding of the heart structure and function". In: *European Journal of Cardio-Thoracic Surgery* 27.2, 191–201. ISSN: 1010-7940. DOI: 10.1016/j.ejcts.2004.11.026. URL: http://dx.doi.org/10.1016/j.ejcts.2004.11.026.

UK Biobank: Protocol for a large-scale prospective epidemiological resource (2007). UKBB-PROT-09-06 (Main Phase).

Van Rossum, Guido and Fred L. Drake (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace. ISBN: 1441412697.

Varma, S. and R. Simon (2006). "Bias in error estimation when using cross-validation for model selection". In: *BMC Bioinformatics* 7.

Vayena, Effy, Alessandro Blasimme, and I. Glenn Cohen (Nov. 2018). "Machine learning in medicine: Addressing ethical challenges". In: *PLOS Medicine* 15.11, e1002689. ISSN: 1549-1676. DOI: 10.1371/journal.pmed.1002689. URL: http://dx.doi.org/10.1371/journal.pmed.1002689.

Vegte, Y. J. van de et al. (2021). "Atrial fibrillation and left atrial size and function: a Mendelian randomization study". In: *Scientific Reports* 11.

Ververidis, Dimitrios and Constantine Kotropoulos (2005). "Sequential forward feature selection with low computational cost". In: *2005 13th European Signal Processing Conference*, pp. 1–4.

Visseren, FLJ, F Mach, YM Smulders, et al. (2021). "2021 ESC Guidelines on cardiovascular disease prevention in clinical practice". In: *European Heart Journal* 42.

Waning, Jaap I. van et al. (Mar. 2021). "Diagnostic Cardiovascular Magnetic Resonance Imaging Criteria in Noncompaction Cardiomyopathy and the Yield of Genetic Testing". In: *Canadian Journal of Cardiology* 37.3, 433–442. ISSN: 0828-282X. DOI: 10.1016/j.cjca.2020.05.021. URL: http://dx.doi.org/10.1016/j.cjca.2020.05.021.

Weintraub, Robert G, Christopher Semsarian, and Peter Macdonald (July 2017a). "Dilated cardiomyopathy". In: *The Lancet* 390.10092, pp. 400–414. DOI: 10.1016/s0140-6736(16)31713-5. URL: https://doi.org/10.1016/s0140-6736(16)31713-5.

— (2017b). "Dilated cardiomyopathy". In: *The Lancet* 390.10092, pp. 400–414.

Welton, N. J. et al. (2017). "Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis". In: *Health Technology Assessment* 21.

Wong, J. A. et al. (2020). "Modifiable risk factors predict incident atrial fibrillation and heart failure". In: *Open Heart* 7.

Wu, Yan et al. (Aug. 2018). "Survival Prediction in High-grade Osteosarcoma Using Radiomics of Diagnostic Computed Tomography". In: *EBioMedicine* 34, pp. 27–34. DOI: 10.1016/j.ebiom.2018.07.006. URL: https://doi.org/10.1016/j.ebiom.2018.07.006.

Xia, Y. et al. (2018). "Detecting atrial fibrillation by deep convolutional neural networks". In: *Computers in Biology and Medicine* 93, pp. 84–92. DOI: 10.1016/j.compbiomed.2017.12.015.

Xu, Pengcheng et al. (2023). "Feature Selection in Machine Learning for Medical Diagnosis: A Comprehensive Review". In: *Journal of Medical Research* 27.3, pp. 345–362.

Yancy, Clyde W. et al. (Oct. 2013). "2013 ACCF/AHA Guideline for the Management of Heart Failure". In: *Journal of the American College of Cardiology* 62.16, e147–e239. ISSN: 0735-1097. DOI: 10.1016/j.jacc.2013.05.019. URL: http://dx.doi.org/10.1016/j.jacc.2013.05.019.

Yin, L. et al. (2019). "Arrhythmia classification based on multi-domain feature extraction". In: *Journal of Physics: Conference Series*. Vol. 1237. IOP Publishing. DOI: 10.1088/1742-6596/1237/2/022028.

Yip, Stephen S.F. and Hugo J.W.L. Aerts (2016). "Applications and limitations of radiomics". In: *Physics in Medicine and Biology* 61.13, R150.

Yusuf, Salim et al. (2004). "Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study". In: *The Lancet* 364.9438, pp. 937–952.

Zabihi, M. et al. (2017). "Detection of atrial fibrillation in ECG hand-held devices using a random forest classifier". In: *Computing in Cardiology*. Vol. 44.

Zacà, V. et al. (2007). "Left atrial enlargement as a predictor of recurrences in lone paroxysmal atrial fibrillation". In: *Canadian Journal of Cardiology* 23.

Zong, W., R. Mukkamala, and R.G. Mark (n.d.). "A methodology for predicting paroxysmal atrial fibrillation based on ECG arrhythmia feature analysis". In: *Computers in Cardiology 2001. Vol.28 (Cat. No.01CH37287)*. CIC-01. IEEE. DOI: 10.1109/cic.2001.977607. URL: http://dx.doi.org/10.1109/CIC.2001.977607.

Zubair, M., J. Kim, and C. Yoon (2016). "An automated ECG beat classification system using convolutional neural networks". In: *2016 6th International Conference on IT Convergence and Security (ICITCS)*. IEEE. DOI: 10.1109/ICITCS.2016.7740310.

Zwanenburg, Alex et al. (2020). "The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping". In: *Radiology* 295.2, pp. 328–338.