Dissecting Cardiometabolic Disease Heterogeneity Using Machine Learning and Multi-Scale Modelling & Simulation

Ambre Bertrand

Reuben College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2025

À Sophie et Samuel,
à qui je dois tout.

À Iris et Loulou,
toujours solidaires.

We deem those happy who from the experience of life have learned to bear its ills without being overcome by them.

— Juvenal, Roman poet (55 AD)

Acknowledgements

I would like to express my deepest gratitude:

To my supervisors: Prof. Blanca Rodriguez, for teaching me "the scientific method" more rigorously than anyone could imagine, for trusting me, and for instilling in me the confidence to trust myself. The countless skills I have learnt through your mentorship are just as valuable if not more than this degree itself, and for that I am extremely thankful to you. Prof. Vicente Grau, for being a calm and steadfast source of expertise, constructive advice and encouragement. Dr. Julia Camps, for your immeasurable commitment to being my supervisor.

To the Computational Cardiovascular Science group, for welcoming me into this field and for providing consistent expertise, technical support and general advice of the highest standard. To Dr. Jakub Tomek, Dr. Andrew Lewis and Dr. Xin Zhou, esteemed colleagues and co-authors, for many excellent discussions and scientific perspectives that have helped me shape a substantial part of this thesis and strengthen its physiological and clinical foundations. To Dr. Albert Dasí, for your friendship and guidance, and for shared values of exciting science and ambition while keeping both feet firmly anchored to the ground.

To past mentors who demystified careers in science and research, turning what was once an unfamiliar path into a passion, and in doing so either lit the spark or fuelled the enthusiasm that ultimately led me to pursue a doctorate: Prof. Daniel Bulte, Prof. Amy Zavatsky, Prof. Paul Goulart, Dr. Mikesh Udani, Dr. Hamza Javed, and my friend and unofficial mentor Stanley Speel. Thank you for your advice, encouragement, and honesty. Thanks also to Mme Jacqueline Forien (Chauvin), my maths teacher at school, for bringing me to Oxford in 2015 for a public event at the Maths Institute; I remember being fascinated by the magic of the city, the talks, and the beautiful Penrose tiling outside the Andrew Wiles Building. Never did it cross my mind at the time that I would one day be writing these lines, finishing a PhD in an office just down the road.

To the institutions, members of staff, and faculty that supported me, financially and administratively: the Department of Computer Science; the Big Data Institute; Reuben College; the Engineering and Physical Sciences Research Council (EPSRC); the Oxford EPSRC Centre for Doctoral Training in Health Data Science; and specifically Jordan Summers, Olga Lappa, Leema Chapman, Kathy Jarvis and my college advisor Prof. Aiden Doherty, for their support and kindness.

To inspiring colleagues turned friends whom I've had the chance to meet (repeatedly) around the world at conferences, training and exchanges over the past few years, and sharing parts of this journey with: Giulia, Carolyna, Ale, Chris, Aron, Giada, Nicolas. I feel fortunate to have found such a diverse and welcoming community of peers, and I hope our paths will cross again in the years to come.

To all the people I have been incredibly lucky to meet at Oxford and who have brought infinite joy and colour to the experience. Particularly, to Kaitlyn, Thea, Jen, Tommy, Aleks, thank you for becoming my reliable, caring, fun, and grounded Oxford family since day one of DPhil at Reuben. I am so grateful you landed in my life and words can hardly do justice to the positive impact you've had over these past four years - I can only hope that this special connection continues to grow beyond our time here together. To Emily, Saadiyah, Bram, Lars, Paul, Nijia, Louise and Sophia; thank you for being kind and consistent friends throughout. A Héctor, por haber compartido conmigo la magia de la electrofisiología cardíaca y mucho más en la vida - gracias por todo.

To Esperanza, Ailsa, Calypso; three very special women and lifelong friends. I am beyond grateful for the deeply joyful, fun, supportive and safe friendships we share, for always being there for each other, and for sharing the ups and downs of life together for many years now. Look forward to making countless new memories with you in the future.

To Dr. Daw and the nurses at UCLH. You saved my life and showed me, first-hand, the value and impact that science can have on people's lives. I am forever grateful to you and hope to keep giving back for as long as I can. To Kasia for your invaluable empathy and support during the most challenging times.

À Marraine Aurélie et Parrain Michel, pour avoir partagé votre sagesse, votre humour et vos excellents goûts littéraires avec moi depuis que je suis toute petite.

Et finalement, à papa et maman, sans qui rien de tout ça n'aurait été possible, et à Iris et Loulou, merci à vous quatre de m'avoir supporté inconditionellement, y compris quand j'étais "dans ma bulle". J'en aurais certainement appris beaucoup sur le coeur à tous ses sens pendant ces dernières années.

Abstract

Metabolic disease is one of the most pressing health and societal challenges of the 21st century. Over half a billion adults are affected by type 2 diabetes, a burden that is expected to grow alongside increasing rates of obesity and older age. Diabetes and obesity are major risk factors for cardiovascular disease, the leading global cause of death, particularly for heart failure with preserved ejection fraction (HFpEF). HFpEF is a complex, multifactorial syndrome affecting over 30 million people worldwide and represents a common "cardiac end-stage" of chronic metabolic dysfunction. Yet, its pathophysiology remains poorly understood and challenging to treat. Understanding these phenotypes and the early cardiac changes preceding overt HFpEF in key drivers such as type 2 diabetes is crucial to improve early identification, personalise care, and prevent disease progression.

Harnessing complementary data-driven and mechanistic computational modelling approaches, this thesis aims to characterise heterogeneity in patient clinical profiles, cardiac structure, function, and electrophysiology, as well as underlying disease mechanisms, to advance clinical identification and mechanistic understanding of adverse cardiac remodelling in HFpEF and type 2 diabetes. I leveraged statistics and machine learning applied to large-scale multi-modal data from the UK Biobank to identify three distinct clinical phenogroups of HFpEF, which differed in cardiometabolic comorbidity burden. Each group exhibited distinct cardiac function and clinical characteristics, highlighting the importance of sex- and comorbidity-specific context in HFpEF characterisation. Focusing on type 2 diabetes as an early stage of the HFpEF spectrum, I demonstrated that despite the absence of diagnosed cardiovascular disease, patients exhibited concurrent changes in ECG and imaging biomarkers suggesting abnormalities in heart rate, repolarisation and impaired left ventricular function; possibly indicative of early-stage HFpEF. Finally, I developed a multi-scale, human-based computational modelling and simulation framework to investigate mechanisms driving these electrophysiological abnormalities in type 2 diabetes. Progressive ionic remodelling characterised by reduced repolarisation currents and increased CaMKII activity underpinned disease severity-dependent changes in action potential duration and calcium handling. The resulting QT prolongation and modest QRS shortening mirror subclinical ECG abnormalities identified at cohort-level.

Overall, this thesis supports the vision of a more personalised and preventive medicine approach in cardiometabolic disease management by offering insights into subclinical cardiac abnormalities in type 2 diabetes and high-resolution phenomapping in overt cardiometabolic HFpEF. This work opens new avenues for in silico investigations to better understand and ultimately prevent cardiac decline in patients with metabolic diseases.

Research Outputs

Journal publications (first author)

1. **Bertrand A.**, Tomek J., Rodriguez B. Cardiac remodelling in type 2 diabetes: pathophysiological mechanisms and opportunities for computational multi-scale modelling and simulation. *Journal of Physiology* 2026 DOI:10.1113/JP287140.

2. **Bertrand A.**, Zhou X., Lewis A., Monfeuga T., Gupta R., Grau V., Rodriguez B. Sex-specific cardiometabolic multimorbidity, metabolic syndrome and left ventricular function in heart failure with preserved ejection fraction in the UK Biobank. *Cardiovascular Diabetology* **24**, 238, 2025. DOI:10.1186/s12933-025-02788-4.

3. **Bertrand A.**, Lewis A., Camps J., Grau V., Rodriguez B. Multi-modal characterisation of early-stage, subclinical cardiac deterioration in patients with type 2 diabetes. *Cardiovascular Diabetology* **23**, 371, 2024. DOI:10.1186/s12933-024-02465-y.

Journal publications (collaborating author)

4. Tomek J., . . . , **Bertrand A.**, . . . , Heijman J. T-World virtual human cardiomyocyte. I. Development, validation, and cell arrhythmogenesis. In press, *Circulation Research*, 2026.
   My contribution: I searched and reviewed extensive literature related to the effects of type 2 diabetes on cardiac ion channels and calcium handling. I had many exchanges with J. Tomek discussing mechanisms and relevant clinical implications of pathophysiological alterations of diabetes in the human heart, which contributed to the development of this new model and its applications to diabetes specifically.

5. Tomek J., . . . , **Bertrand A.**, . . . , Heijman J. T-World virtual human cardiomyocyte. II. Organ-scale simulations and applications. In press, *Circulation Research*, 2026.
   My contribution: as above.

6. Martinez-Navarro H., **Bertrand A.**, et al. ECG analysis of ventricular fibrillation dynamics reflects ischemic progression subject to variability in patient anatomy and electrode location. *Frontiers in Cardiovascular Medicine*, 2024. DOI:10.3389/fcvm.2024.1408822.
   My contribution: data analysis, visualisation, and manuscript drafting and editing for this

modelling and simulation study, assessing ECG metrics to quantify ventricular fibrillation in myocardial ischemia and the effect of electrode positioning on the reliability of these metrics.

7. Xavier R., . . . , **Bertrand A.**, . . . , Lewis A. Characterization of Exercise Responses in HFpEF stratified by Ejection Fraction: Implications for Therapeutic Myosin Modulation. Under review. 2026.
   My contribution: data analysis, discussion, visualisation of CMR biomarkers at rest and under exercise conditions, and other non-imaging biomarkers, to assess left ventricular function and exercise response in HFpEF patients.

8. Manohar S., . . . , **Bertrand A.**, . . . , Ariga R. et al. ECG Deep Learning Dissects Pathophysiology of Hypertrophic Cardiomyopathy. Under review. 2026.
   My contribution: discussions and data support around UK Biobank-specific definition and ascertainment of diseases and outcomes, data availability, and data mining in the UK Biobank.

Conference publications and chapters

7. **Bertrand A.**, . . . , Rodriguez B. A Multi-Scale Computational Framework for Human-Based Modelling and Simulation of Adverse Cardiac Remodelling in Type 2 Diabetes. *Computing in Cardiology*, 2025. DOI:10.22489/CinC.2025.418

8. Pinard A., Wang Z.J., **Bertrand A.**, Rodriguez B., Camps J. Emulation of Cardiac Mechanics using SE(3) Equivariant Graph Neural Networks. *Computing in Cardiology*, 2025. DOI:10.22489/CinC.2025.157

9. **Bertrand A.**, . . . , Rodriguez B. Type 2 diabetes individuals have a shorter QRS duration, prolonged QTc interval, flatter T wave, and reduced left ventricular cycle volumes compared to matched controls in the UK Biobank. *European Society of Cardiology Congress*, 2024. DOI:10.1093/eurheartj/ehae666.2912 (indexed abstract in the European Heart Journal).

10. **Bertrand A.**, . . . , Rodriguez B. Sex-specific, multi-modal assessment of cardiac function in type 2 diabetes using the UK Biobank. *Computing in Cardiology (CinC)*, 2024. DOI:10.22489/CinC.2024.406 (4-page conference paper).

11. **Bertrand A.** et al. Augmentation of Cardiac Ischemic Geometry for Improving Machine Learning Performance in Arrhythmic Risk Stratification. In: McCabe, K.J. (eds) *Computational Physiology. Simula SpringerBriefs on Computing*, vol. 17. Springer, Cham., 2023. DOI:10.1007/978-3-031-53145-3_3.

12. **Bertrand A.**, . . . , Rodriguez B. Deep Learning-Based Emulation of Human Cardiac Activation Sequences. In: Bernard, O., et al. (eds) *Functional Imaging and Modeling of the Heart. FIMH 2023. Lecture Notes in Computer Science*, vol. 13958. Springer, Cham., 2023. DOI:10.1007/978-3-031-35302-4_22 (12-page conference paper).

International conference presentations

- Computing in Cardiology (talk at the closing plenary session). São Paulo, Brazil, September 2025.
- Federated American Societies for Experimental Biology (FASEB) Science Research Conference on Ion Channel Regulation (talk). Southbridge, USA, August 2025.
- Gordon Research Conference on Cardiac Arrhythmia Mechanisms (talk and poster). Lucca, Italy, February 2025.
- Gordon Research Seminar on Cardiac Arrhythmia Mechanisms (talk and poster). Lucca, Italy, February 2025.
- Computing in Cardiology (talk). Karlsruhe, Germany, September 2024.
- European Society of Cardiology (ESC) Congress (e-poster 3 min presentation). London, UK, August 2024.
- Functional Imaging and Modelling of the Heart (FIMH) (poster). Lyon, France, June 2023 (in absentia).

Prizes and awards

- Engineering and Physical Sciences Research Council (EPSRC) Postdoctoral Pathway Fellowship. Oxford, awarded August 2025.
- Runner-up of the Karlsruhe Simulation Award, Computing in Cardiology. São Paulo, Brazil, 2025.
- Travel award, Reuben College, Oxford, 2025.
- Best talk award, Gordon Research Seminar on Cardiac Arrhythmia Mechanisms. Lucca, Italy, 2025.
- Best poster award, Gordon Research Seminar on Cardiac Arrhythmia Mechanisms. Lucca, Italy, 2025.

- Travel award, Gordon Research Seminar on Cardiac Arrhythmia Mechanisms. Lucca, Italy, 2025.
- Best poster award, Health Data Science Centre for Doctoral Training Symposium. Oxford, 2022.
- Winning team, Health Data Science Centre for Doctoral Training's annual data challenge. Oxford, 2022.
- EPSRC Centre for Doctoral Training in Health Data Science Scholarship. Oxford, 2021–2025.

Contents

List of Abbreviations

1D, 2D, 3D . . . One-, two-, three-dimensional.

AngII . . . . . Angiotensin II.

ANOVA . . . . Analysis of variance.

BMI . . . . . Body mass index.

CaMKII . . . . Ca^{2+}/calmodulin-dependent protein kinase II.

CMR . . . . . Cardiac magnetic resonance.

CVD . . . . . Cardiovascular disease.

ECG . . . . . Electrocardiogram.

GPU . . . . . Graphics processing unit.

HFpEF . . . . . Heart failure with preserved ejection fraction.

ICD . . . . . International Classification of Diseases.

LCA . . . . . Latent class analysis.

LVEF . . . . . Left ventricular ejection fraction.

MRI . . . . . Magnetic resonance imaging.

NCX . . . . . Sodium-calcium exchanger membrane protein.

SGLT2 . . . . . Sodium-glucose cotransporter 2.

SHBG . . . . . Sex hormone binding globulin.

T2DM . . . . . Type 2 diabetes mellitus.

UK . . . . . United Kingdom.

Nothing is particularly hard if you divide it into small jobs.

— Henry Ford, American industrialist (1863-1947)

1 Introduction

1.1 Motivation

1.1.1 Clinical background

Metabolic diseases such as type 2 diabetes mellitus (T2DM), obesity, and hypertension, have reached epidemic levels worldwide, affecting over 2.5 billion people and significantly increasing the risk of cardiovascular disease (CVD), the leading global cause of death. In 2023, CVD-related costs exceeded €280 billion in Europe [1], projected to reach $1.5 trillion by 2050 in the US [2], and are estimated at $3.8 trillion in low- to middle-income countries between 2011-2025 [3]. With ageing populations and worsening health trends, the global CVD burden could increase by up to 90%, reaching over 35 million deaths annually by 2050 [4]. Understanding how metabolic diseases drive CVD, clinically and mechanistically, is therefore crucial to reduce this burden.

Chronic inflammation and underlying metabolic disturbances are common to conditions such as diabetes, hypertension and obesity. These conditions are tightly interlinked and frequently co-exist, leading to adverse cardiac and extra-cardiac complications. Indeed, if unmanaged, patients with cardiometabolic comorbidities often develop heart failure with preserved ejection fraction (HFpEF). HFpEF is a highly heterogeneous syndrome which affects over 30 million individuals worldwide and whose prevalence is on the rise due to ageing and more obese populations [5, 6]. It is increasingly being treated

on a phenotype-specific basis, however due to the complexity of the condition, the refinement of HFpEF phenotypes remains an active area of research [7–9]. Developing effective treatments for HFpEF has also been a major clinical challenge; sodium-glucose cotransporter 2 (SGLT2) inhibitors, initially developed to treat diabetes, are the only pharmacological agent with proven benefits in patients with HFpEF, independently of diabetes status [10, 11]. Sex and age also critically influence HFpEF. Women, compared to men, have stiffer ventricles and arteries, smaller vasculature, and higher pulse pressure, which increases their risk of adverse microvascular changes, diastolic dysfunction, and ultimately HFpEF [12–15]. Ageing further modifies these processes, with sex-specific trajectories of cardiometabolic decline and cardiac dysfunction [16, 17]. Thus, it is crucial to better understand the underlying cardiometabolic comorbidity landscape in HFpEF while accounting for age and sex to optimise disease management, slow disease progression, and ultimately improve patient outcomes.

One of the major drivers of HFpEF is T2DM. T2DM is a complex metabolic disease that is characterised by chronically elevated blood glucose levels. It affects over half a billion adults worldwide, and is responsible for a two- to four-fold increase in CVD [18, 19]. It is known to cause adverse cardiac remodelling at cell, tissue, and whole-organ level, driving abnormalities in the heart's electrophysiology, structure and function. These adverse changes are typically underdiagnosed and may progress to HFpEF, cardiac arrhythmias and sudden cardiac death if unidentified and uncontrolled [20–23]. However, large-scale studies of cardiac function in patients with T2DM, especially prior to overt CVD, are scarce, and mechanistic understanding of multi-scale cardiac remodelling in T2DM remains limited. This gap is critical, as early identification and mechanistic understanding of cardiac dysfunction in these high-risk patients could significantly reduce risk and burden of future HFpEF and cardiac arrhythmias.

1.1.2 Advances in medical technology and computational methodologies

The electrocardiogram (ECG) and cardiac imaging are common, non-invasive methods that provide valuable insights into clinical and subclinical changes in cardiac structure

and function. The ECG's cost-effectiveness, widespread availability, and sensitivity to electrophysiological changes make it an ideal candidate for CVD screening in the clinic. Cardiac magnetic resonance (CMR) imaging provides a robust assessment of structural and functional abnormalities in the heart. However, its resource-intensive and expensive nature limit its uptake for routine screening, including in patients with diabetes. Indeed, asymptomatic patients with T2DM are not routinely screened for CVD [24].

Large-scale prospective biobank studies such as the UK Biobank, the world's largest biomedical resource of its kind, provide researchers with standardised, well-curated biological samples, linkage to electronic health records, and a wide range of clinical, socio-demographic and lifestyle information on large cohorts of middle- to older-age individuals, in an easily accessible and reproducible manner [25]. Its richness in breadth and depth of data, including ECG and CMR, and optimal target population age, make the UK Biobank a unique and ideal resource to investigate complex conditions such as HFpEF at high-resolution, and to quantify cardiac abnormalities in high-risk, understudied populations such as T2DM, by harnessing data-driven analysis and statistical modelling methods including machine learning.

However, while these approaches enable the identification of clinical abnormalities through biomarker changes in cohorts of interest, they do not offer mechanistic insight. To this end, biophysical modelling and simulation approaches represent a reliable, virtual testbed to investigate possible pathophysiological mechanisms underlying those changes. Extensively validated, biophysically-detailed, multi-scale computational models of human cardiac electrophysiology have been able to replicate and explain complex biological mechanisms in healthy and diseased populations [26–30]. These models enable us to investigate the mechanisms underlying organ-level abnormalities as identified by clinical biomarkers, by investigating the effect of important factors such as cardiac ionic remodelling and anatomy on those biomarkers.

Together, these advances in medical technology, digitalisation, large biobanks and diverse modelling approaches provide new opportunities to study CVD at an unprecedented scale.

1.2 Aims

Leveraging these innovations, the overarching aim of my DPhil is to characterise adverse cardiac remodelling in patients with cardiometabolic disease, with a focus on T2DM and HFpEF, to improve the identification and mechanistic understanding of underlying early pathophysiological changes in these tightly intertwined conditions. I will achieve this by applying computational statistics, machine learning, and multi-scale human-based cardiac modelling and simulation to biomedical data from selected patient cohorts from the UK Biobank. Specifically, the goals of this thesis are to:

1. Identify distinct clinical subgroups in patients with comorbid cardiometabolic HFpEF, and compare left ventricular function across the different groups identified. In this aim, I consider HFpEF as the cardiac end-stage of diabetes and other cardiometabolic comorbidities. I hypothesise that, despite being considered under the same umbrella of HFpEF, there exist different cardiometabolic phenogroups of HFpEF with varying clinical profiles, and that there are sex-specific differences in cardiometabolic profiles of men and women with HFpEF. I further hypothesise that between those subgroups, and between men and women within the same subgroup, there exist subtle differences in cardiac function, which may impact disease progression and treatment response.

2. Focusing on subclinical cardiac remodelling prior to HFpEF, quantify differences in ECG and cardiac imaging-derived biomarkers in patients with T2DM and no history of CVD, and identify biomarkers associated with the development of CVD in these individuals. I hypothesise that patients with T2DM will exhibit both depolarisation and repolarisation abnormalities visible in the ECG, and early signs of diastolic dysfunction. I will address this by defining a cohort of UK Biobank subjects with T2DM and no CVD, and matched healthy controls, and identifying key changes in ECG and cardiac imaging biomarkers reflecting hypothesised abnormalities through the use of statistical testing and multiple linear regression modelling.

3. Identify pathophysiological mechanisms of adverse cardiac remodelling in T2DM, both at ionic and organ level, that may explain the ECG changes observed in the previous aim. I hypothesise that ECG abnormalities in T2DM are driven by a combination of ionic remodelling, anatomical variability, and organ-level electrophysiological properties; mechanisms which I unravel in this final aim by exploiting multi-scale, human-based computational modelling and simulation of cardiac electrophysiology.

Overall, the intended impact of this work is twofold:

- To refine our understanding of clinical heterogeneity and late-stage metabolism-driven cardiac dysfunction in cardiometabolic HFpEF, contributing to the vision of personalised medicine through a more nuanced understanding of the HFpEF spectrum, and
- To improve the identification and mechanistic understanding of early subclinical abnormalities in T2DM, a major HFpEF risk factor, ultimately aiming to inform future strategies for risk stratification and early prevention of cardiac arrhythmias and HFpEF in this high-risk population.

1.3 Thesis outline

Chapter 1 presents the thesis, its motivation and aims, as well as an outline summarising the contents of each chapter.

Chapter 2 consists of a literature review on the physiological and clinical background that this thesis builds upon. It introduces concepts of cardiac anatomy, function and electrophysiology in the healthy heart, as well as clinical modalities to assess cardiac health including ECG and CMR imaging. It then presents heterogeneities in epidemiology, clinical biomarkers, and multi-scale cardiac remodelling pathways in T2DM and HFpEF.

Chapter 3 introduces the UK Biobank, the database this thesis is based on, followed by two key relevant methodological areas. The first half of the methods presented focuses on relevant health data science, statistical modelling and machine learning

approaches for the analysis of data at population level. The second half introduces human-based, biophysically detailed computational modelling and simulation of human cardiac electrophysiology, from ionic to whole organ-level modelling.

Chapter 4 focuses on late-stage cardiometabolic CVD, through a cross-sectional characterisation study of patients with HFpEF in the UK Biobank. I quantify differences in cardiometabolic profiles in HFpEF in men and women and identify distinct phenogroups of cardiometabolic HFpEF using machine learning based on various clinical markers of cardiometabolic health. I then quantify left ventricular function across these phenogroups, examine the relationship between metabolic syndrome components and parameters of left ventricular function, and explore sex-specific differences within each subgroup.

Chapter 5 approaches earlier clinical stages within the broader cardiometabolic disease spectrum, focusing on patients with T2DM as a major risk factor for HFpEF, prior to overt CVD. I present a pair-matched cross-sectional study that considers ECG and CMR image-derived biomarkers in two matched cohorts from the UK Biobank, with and without T2DM, and no diagnosed CVD. I compare cardiac biomarkers in the two cohorts and quantify the association between T2DM and these biomarkers; build multiple linear regression models sequentially adjusted for socio-demographic, lifestyle, and clinical covariates; and carry out subgroup analyses in groups with different cardiovascular outcomes, sex, ethnicity, age, and body mass index.

Chapter 6 builds on the previous chapter by developing a computational framework to enable the investigation of pathophysiological mechanisms underlying electrophysiological disturbances discussed in Chapter 5. I present a multi-scale computational study of the mechanistic effect of ionic remodelling, organ-level electrophysiological properties and anatomical variability on the ECG in representative cases of T2DM patients of different sex, and metabolic condition severity disease, through increasing duration of T2DM and body mass index. 3D patient-specific anatomical meshes are reconstructed from 2D CMR image slices from selected cases in the UK Biobank. Single-cell simulations are carried out using a previously validated model of human ventricular cardiomyocyte electrophysiology, that I adapt to T2DM based on experimental data. This model is used to embed functionality into the reconstructed cardiac meshes, together acting as the

basis of high performance computing-enabled biventricular simulations for 3D cardiac activation sequence and ECG reconstruction in a representative selection of individuals T2DM of different sex and disease severity.

Chapter 7 concludes the thesis by providing a summary of the work and its key contributions, discussing its limitations, and suggesting possible future avenues to explore.

In a world where information is abundant and easy to access, the real advantage is knowing where to focus.

— James Clear, American writer (1986-present)

2

Physiological and Clinical Background

Section 2.5 of this chapter forms part of the following publication:

- **Bertrand A.**, Tomek J., Rodriguez B. Cardiac remodelling in type 2 diabetes: pathophysiological mechanisms and opportunities for computational multi-scale modelling and simulation. 2025 (under review, conditionally accepted at The Journal of Physiology).

2.1 Introduction

This chapter introduces the key physiological and pathological concepts, along with their clinical implications, that motivate the original work presented in Chapters 4, 5 and 6. I begin by describing the healthy human heart and its anatomy, function, tissue-level structure, and electrical activity. Then, I introduce clinical modalities that are used to assess cardiac structure and function, notably the electrocardiogram and cardiac magnetic resonance imaging, which will be considered throughout the rest of the thesis. I continue with an extensive review of the epidemiology, pathophysiological mechanisms, biomarkers, and clinical implications of heart failure with preserved ejection fraction (HFpEF) and type 2 diabetes mellitus (T2DM)-driven cardiac disease, highlighting the diversity and complexity of these tightly interlinked clinical phenotypes to provide the necessary foundations for the research carried out in this thesis.

2.2 The healthy human heart

2.2.1 Organ-level structure and mechanical function

The human heart is an organ that is responsible for the correct circulation of blood to all other organs in the body. It does so in a cyclic manner, acting as a pump. It consists of four chambers: the right atrium and the left atrium, which sit above the right ventricle and the left ventricle, respectively (Figure 2.1, left). The wall of tissue separating the ventricles is called the septum. The top (superior) part of the heart including the atria is referred to as the base of the heart, while the bottom (inferior) part including the ventricles is referred to as the apex. There is also a distinction between the front (anterior) and back (posterior) faces of the heart, which align with the front and back of a person, respectively.

The atria receive blood from other organs in the body, at which point the ventricles are relaxed. This phase of the cardiac cycle is called diastole. The blood is transferred from the atria to the ventricles through the (right) tricuspid valve and (left) mitral valve. The ventricles then pump the blood back out to the lungs (right ventricle) and rest of the body (left ventricle) by contracting; this phase of the cycle is called systole. Even in healthy conditions, the left ventricular wall is significantly thicker than its right counterpart, as it requires more pressure to pump the blood further to more distant organs.

2.2.2 Cell- and tissue-level electromechanical function

The cardiac cycle of mechanical contraction and relaxation, as described above, is driven by the heart's electrical activity. This process is mediated by distinct cardiac cell and tissue types with different properties.

Located in the upper back wall of the right atrium, the sinoatrial node is a group of cells that acts as the primary pacemaker for the heart by spontaneously generating a burst of electrical activity. This electrical signal reaches the atrioventricular node, from where it propagates through the heart's fast conduction system made up of the Bundle of His and the Purkinje fibres, and then to the ventricles through cardiac tissue known as myocardium (Figure 2.1, right). The myocardium is formed of cardiac muscle cells called cardiomyocytes, which are cylindrical and measure $\sim$100μm in length with a diameter of

10-25µm. Cardiomyocytes are arranged end-to-end and are separated by a small gap of ~25nm, which narrows to ~3nm at "gap junctions". These gap junctions are proteins that enable the transfer of ions, and therefore electrical signal, between cells. Consecutive cardiomyocytes form long one-dimensional fibres, which themselves are arranged in adjacent two-dimensional sheets. Together, these sheets make up the different layers of the myocardial wall of the heart: the endocardium (inner layer), midmyocardium (middle layer), and the epicardium (outer layer). Each layer has different electrical properties, and electrical propagation is faster along the fibre sheets rather than across them.

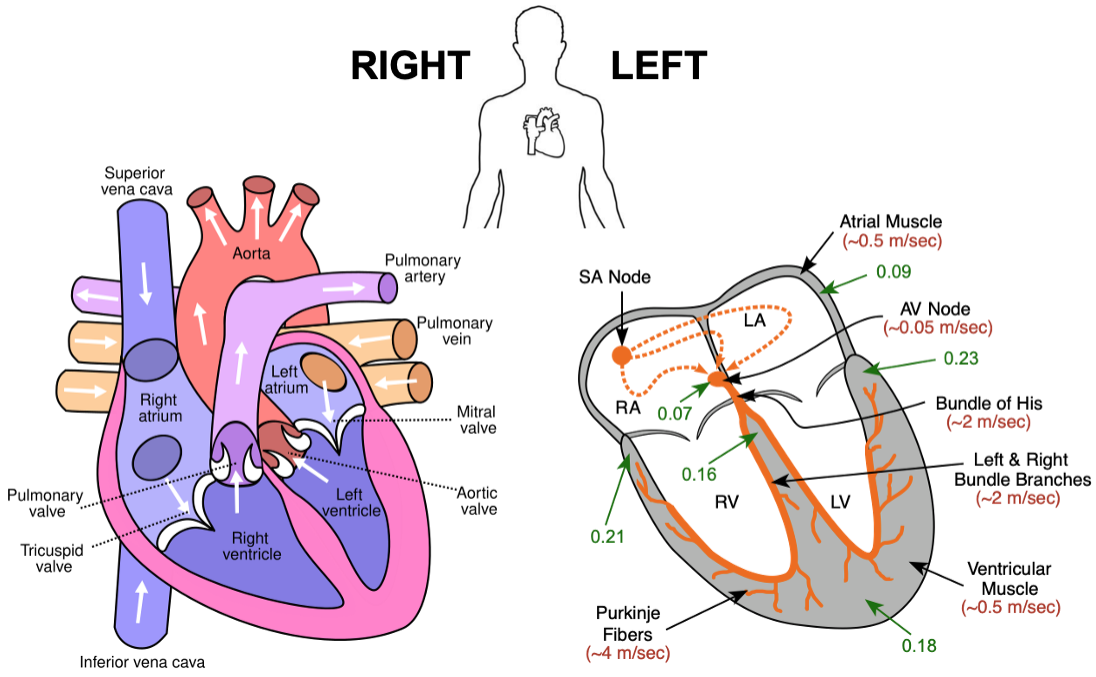

Figure 2.1: Structure and electrical conduction of the human heart. Left: diagram of the human heart, with labels for chambers, valves and blood vessels. Right: components of the heart's electrical activity with typical conduction speeds (red) and activation times (green, in seconds). Adapted from Pierce, E. Wikimedia Commons [31] and the cardiac conduction system [32]. AV: atrioventricular, LA: left atrium, LV: left ventricle, RA: right atrium, RV: right ventricle, SA: sinoatrial.

Unlike sinoatrial node cells, cardiomyocytes cannot self-excite. When an activation stimulus reaches them, however, the ionic concentrations across the cell membrane changes and an extracellular potential field is established. This potential field excites neighbouring cells which results in cell-to-cell electrical propagation within the heart.

When a cell is electrically activated, the sarcomeres, basic contractile units of the myocytes, shorten, making the cell contract along its longitudinal axis. Thus, the orientation of the myocytes and the fibres they form dictates the axis of contraction of the myocardium, in addition to electrical signal propagation direction. Thus, structural alterations of this organisation, such as myocardial fibrosis (an abnormal accumulation of collagen in the heart tissue often resulting from acute or chronic insult), will affect not only the normal mechanical function of the heart, but also its electrical activity [33].

2.2.3 Cellular-level electrophysiology

When it is not stimulated and is in its resting state, a cardiomyocyte's cell membrane maintains a negative potential of around -90mV due to the balance of different ionic concentrations around it. Embedded in the cellular membrane are ion channels, small proteins that enable the passage of ions across the membrane. At rest, most ion channels are closed. An electrical stimulus of sufficient strength reaching the cell can disturb this ionic balance and trigger of an action potential that normally lasts 150-350ms. Ion channels open and close which causes different ionic species to flow in and out of the cell, changing its membrane potential. The cardiomyocyte's action potential can be separated into four key consecutive phases:

- **Phase 0, depolarisation:** a sudden increase in sodium (Na^+) ions into the cell causes an initial upstroke in membrane potential from its resting state, which becomes suddenly much less negative. The sodium current (I_{Na}) is highly voltage-dependent and is the key initiator of the action potential.

- **Phase 1, early repolarisation:** once the maximum depolarisation has been reached, I_{Na} becomes inactive. The transient outward potassium current (I_{to}) activates and potassium (K^+) ions flow out of the cell. This phase appears as a notch in the action potential trace.

- **Phase 2, plateau:** Calcium (Ca^{2+}) ions enter the cell as the L-type calcium current (I_{CaL}) increases, while slow delayed rectifying potassium current (I_{Ks}) decreases and K^+ ions leave the cell. The inward rectifying potassium current (I_{K1}),

sodium calcium exchanger (I_{NaCa}) and sodium potassium pump (I_{NaK}) play a role in balancing Ca^{2+}, N^{+} and K^{+} ions here. The late sodium current (I_{NaL}) also helps to maintain a stable voltage throughout this phase.

- **Phase 3, repolarisation:** K^{+} ions leave the cell as I_{K1} and the fast and slow delayed rectifying potassium currents (I_{Kr}, I_{Ks}) decrease. The cell repolarises completely.

- **Phase 4, resting state:** the cell returns to its resting state, K^{+} exiting the cell through I_{K1}.

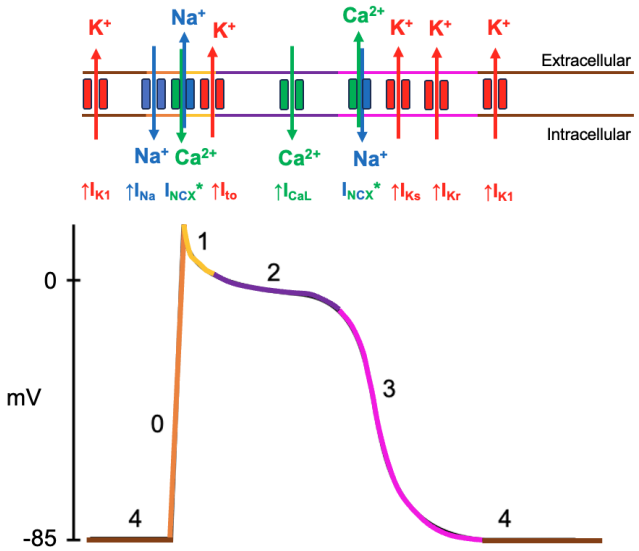

Figure 2.2: The four main phases of the cardiac myocyte action potential and its key ionic current contributors. An increase in each current reflects the activation and opening of its corresponding ion channel, allowing ions to flow into or out of the cell. The asterisk next to I_{NCX} highlights the complexity of the sodium-calcium exchanger's behaviour: it first operates in "reverse mode" during depolarisation (phase 0), whereby Ca^{2+} flows into the cell and Na^{+} flows out of the cell, followed by "forward mode" during repolarisation (phases 1 to 3), allowing Na^{+} into the cell and extruding Ca^{2+}.

2.3 Clinical modalities

In this section, I introduce two key clinical modalities that are used for structural and functional assessment of the heart: the electrocardiogram (ECG) and cardiac imaging techniques.

2.3.1 The electrocardiogram

Because the thorax acts as a conductive medium surrounding the heart, the potential field generated by the heart's electrical activity propagates to the surface of the body. This field can be recorded as an electrical signal, varying in time with the heartbeat, by attaching a pair of surface electrodes to a human subject and connecting these electrodes to a high-input impedance differential amplifier. This signal is known as the ECG and has a peak amplitude of ~1mV before amplification.

The ECG represents a valuable clinical tool to assess the electrical activity of the heart. By analysing the shape of the ECG, it is possible to identify abnormalities that may reflect possible underlying structural and electrophysiological pathologies. The ECG has many benefits in clinical applications as it is non-invasive, fast, low-power, low-cost and easily deployable. The ECG can have 12, 2, or a single lead, each configuration serving different purposes depending on clinical need, setting, and level of detail required. A summary is given in Table 2.1.

Table 2.1: Summary of pros and cons of single-lead, 2-lead and 12-lead ECG setups.

Feature	**Single-lead**	**2-lead**	**12-lead**
Rhythm detection	Good	Good	Excellent
Pathology detection	Poor	Poor	Good
Continuous monitoring	Good	Very good	Limited
Portability	Excellent	Very good	Poor
Diagnostic depth	Poor	Poor	Excellent

Electrode placement and signal reconstruction

The modern ECG was developed by Willem Einthoven in the early 1900s who defined the positioning of the electrodes on the limbs and measured the first ECG signals from electrode pairs. Placement of these pairs include the right arm (RA) and left arm (LA); RA and left leg (LL); LA and LL. An additional electrode, usually attached to the right leg (RL), connects the patient to the common ground of the instrumentation, completing the electrical circuit and minimising noise.

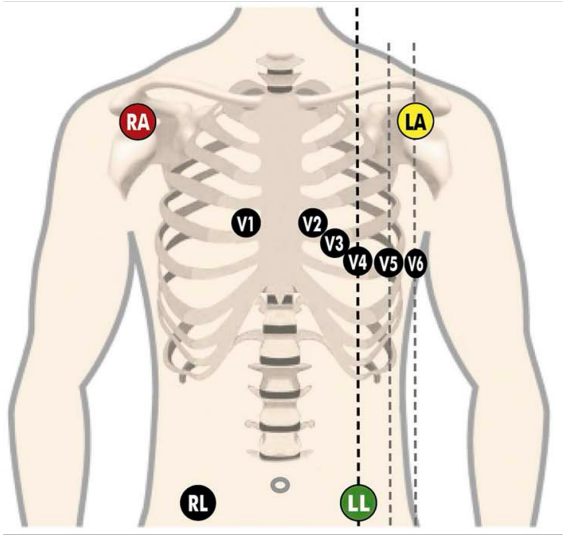

Figure 2.3: Classic Mason-Likar electrode placement for the 12-lead ECG. Alternatively, RA (right arm), LA (left arm), RL (right leg), and LL (left leg) may be positioned at the extremities of the limbs.

Today, the classic Mason-Likar set-up for the 12-lead ECG includes 10 electrodes placed on the skin: four on the corners of the chest, and six further electrodes (V1 to V6) positioned on the upper left area of the chest surrounding the heart (Figure 2.3). The 12 ECG leads arise from the difference between voltages recorded at distinct electrode positions (RA, LA, RL, LL). These different leads offer extensive coverage of the heart's electrical activity, as measured from a range of different complementary angles.

Limb leads:

$$\text{Lead I} = \text{LA} - \text{RA},$$

$$\text{Lead II} = \text{LL} - \text{RA},$$

$$\text{Lead III} = \text{LL} - \text{LA}.$$

Augmented limb leads, with Wilson's reference terminal $V_W = \frac{1}{3}(\text{RA} + \text{LA} + \text{LL})$:

$$\text{Lead aVR} = \tfrac{3}{2}(\text{RA} - V_W),$$

$$\text{Lead aVL} = \tfrac{3}{2}(\text{LA} - V_W),$$

$$\text{Lead aVF} = \tfrac{3}{2}(\text{LL} - V_W).$$

Precordial (chest) leads V_1–V_6: The difference between the voltage at each chest electrode (V_1–V_6) and the Wilson reference terminal V_W.

ECG signal shape

If a depolarising electrical wave in the heart travels toward a specific electrode, or a repolarising wave travels away from it, this will be reflected as a positive deflection on the ECG. A depolarising wave traveling away from the electrode or a repolarising wave travels toward it will result in a negative deflection on the ECG. The voltage measured at each electrode is directly proportional to the source current and inversely proportional to the distance between the source and the electrode, meaning that the amplitude of the ECG depends on both the strength of the electrical signal generated by the heart and the distance between the electrode and the myocardium.

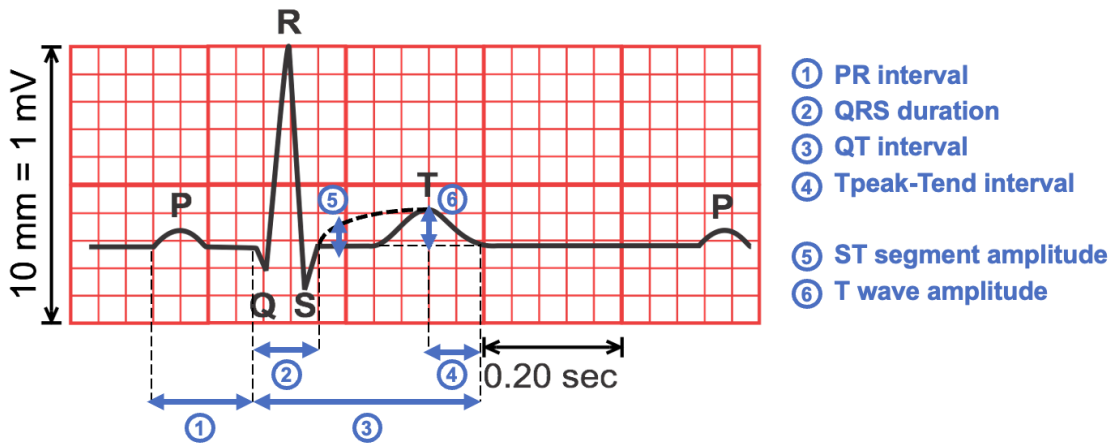

Figure 2.4: Typical ECG morphology and key biomarkers. ST segment amplitude (biomarker 5) tends to be zero in healthy cases; under certain pathological conditions e.g. myocardial ischemia or infarction, the ST segment may be depressed or elevated, as represented here in the ECG by a dotted line (adapted, original ECG trace from www.cvphysiology.com).

The complete ECG signal represents consecutive heartbeats. In the ECG, a single heartbeat consists of three main stages: the P wave, the QRS complex, and the T wave (Figure 2.4). The P wave represents the depolarisation of the atria, which marks the start of a heartbeat. It is usually positive in all leads but aVR, and tends to be <120ms long and <0.3mV tall. Next comes the QRS complex representing the depolarisation of the ventricles, which in healthy patients tends to last between 80ms and 120ms [34]. Typically, Q waves are negative, R waves are tall and positive and S waves are negative. This complex morphology arises from the electrical wave travelling from the top of the basal region of the heart down toward the apex through the septum, and back up through

the ventricular walls. Q, R, and S wave amplitude and polarity differ across ECG leads depending on their position relative to the ventricles. Under normal conditions, Q, R, and S wave amplitudes tend to follow a progression in the chest leads, from a short R wave and deep S wave in lead V1, to the tallest R wave in lead V5, and a slight reduction in R wave amplitude in V6. Finally, the T wave represents the repolarisation of the ventricles. It is normally wide, shallow, and upright. T wave inversion may be observed in certain leads in healthy patients, but can also be associated with pathological conditions [34].

2.3.2 Imaging modalities

There are three main types of imaging modalities that are used clinically to gather information on a patient's cardiac anatomy, structure and function in a non-invasive manner. Each relies on a different physical phenomenon to produce images of the heart. This, along with their cost, accessibility, and image resolution, dictates their scope of application in clinical practice.

Echocardiography

Echocardiography relies on ultrasound and signal reflection. A probe is pressed against the chest and emits high-frequency sound waves, which travel through the chest, are reflected at the surfaces between different tissues and organs, and travel back to reach the probe which receives the returned signal. Depending on the distance travelled by the original signal before being reflected back, the received signal will be more or less strong, creating contrast between different tissues and organs in the final image. This approach is fast, cheap, non-ionising and has good temporal resolution but poor spatial resolution, making it an ideal candidate for routine cardiac functional assessments but not suitable for precise characterisation of tissue or blood vessels.

Computed tomography

Computed tomography relies on X-ray transmission and attenuation. In the scanner, the patient is surrounded by X-ray emitters and detectors. Photons emitted by the machine interact with different tissues in the body. This leads to a reduction in the received signal (photoelectric attenuation), which creates contrast in the image, and a small amount of

energy release and X-ray deflection (Compton scattering), which causes noise in the image. Different tissues attenuate the rays to different extents to create the contrast, which can be further achieved by contrast agents. Therefore, computed tomography can produce images of high spatial resolution, including for blood vessels, however it is an ionising modality so its use is limited.

Magnetic resonance imaging

Magnetic resonance imaging (MRI), or cardiac magnetic resonance (CMR) for the heart, relies on magnetic fields and intrinsic tissue properties. MRI effectively measures the signal emitted by hydrogen H+ protons in the body. Protons have a physical property called "spin" that is receptive to magnetic fields. Under normal conditions, protons all precess, i.e. rotate at an angle from an axis, out of phase and with individual spin axes pointing in different directions. In an MRI scanner, a large external magnetic field B0 is applied; spin axes start aligning with that of B0, and protons start precessing in phase. A smaller magnetic field B1 is then applied perpendicular to B0 and "pushes" the spin axes in the direction of B1. Once B1 is removed, the spins return to their original state. The rate at which they do so is characterised by two distinct time constants, T1 (recovery time) and T2 (relaxation time), that describe the time taken for orthogonal components of the spin to return to their original state. The signal emitted by the recovering spins will differ depending on the chemical environment the protons are in, e.g. blood, muscle, bone, fat, etc., creating excellent contrast in the reconstructed image. The addition of gadolinium as a contrast agent can further enhance this contrast, as its uptake in certain tissues, e.g. myocardial scar, is higher, appearing as bright areas in the final images; this is referred to as late gadolinium enhanced MRI. T1 mapping, both pre- (native) and post-contrast, measures tissue T1 recovery time and can enable non-invasive quantification of the myocardium's characteristics, including myocardial inflammation and extracellular volume, that can be used as a proxy marker of diffuse myocardial fibrosis [35]. Overall, due to its non-ionising and high-resolution properties, MRI is an ideal option to characterise tissue and blood vessels and diagnose certain cardiac conditions. However, its cost, energy use (the magnets

used to produce the fields must be constantly stored at very cold temperatures) and long procedure times limit its widespread, routine use in clinical settings.

2.4 Literature review: heart failure with preserved ejection fraction

In this section, I review the key epidemiological, clinical and pathophysiological aspects of HFpEF, challenges in its clinical heterogeneities and pharmacological treatment, and novel avenues to optimise patient management and treatment.

2.4.1 Epidemiology, diagnosis, and clinical challenges of a heterogenous syndrome

HFpEF is a complex clinical syndrome that affects over 30 million individuals worldwide, with a rising prevalence due to older and more comorbid populations [6, 36]. Common symptoms of heart failure include dizziness, shortness of breath, fatigue, and chest pain. Patients who present to the clinic with those symptoms, and who are identified as being at risk of heart failure, undergo a three-stage screening process that consists of an ECG recording, blood test, and echocardiography. If the ECG is deemed abnormal, a serum sample is collected to measure levels of circulating NT-proBNP (B-type natriuretic peptide). This circulating biomarker originates in the myocardial cells and is secreted in response to wall stress. It is a very sensitive marker of heart failure, such that heart failure is almost certain if NT-proBNP concentration is above a certain threshold. A diagnosis is confirmed by abnormal findings following echocardiographic assessment [37]. Patients with heart failure are further classified into different categories based on their left ventricular ejection fraction (LVEF), a measure of cardiac function based on image-derived biomarkers of cardiac chamber volume, namely end-diastolic volume (EDV) and end-systolic volume (ESV) of the left ventricle (LVEF (%) = ((EDV - ESV) / EDV) x 100). A distinction is thus made between heart failure with reduced ejection fraction (HFrEF) with LVEF<35-40% or preserved ejection fraction (HFpEF) with LVEF>40-50%. Heart failure with mildly reduced ejection fraction (HFm(r)EF) is sometimes used to describe patients with LVEF in the range 41-49% [37].

However, the diagnosis and clinical definition of heart failure pose challenges in many respects. In the 2021 European Society of Cardiology guidelines, an NT-proBNP level above 125 pg/mL in non-acute settings is considered strongly indicative of heart failure [37]. However, accounting for more recent trial results, this threshold may not be met in certain patient subgroups, particularly those with obesity or HFpEF, suggesting that normal NT-proBNP concentrations do not reliably exclude heart failure in all cases [38]. Furthermore, heart failure is a highly heterogeneous condition; its etiology and pathology vary greatly from patient to patient, as it is typically the cardiac "end stage" of a number of different diseases. Each condition has its own distinct pathophysiological mechanism of action that can cause direct or indirect adverse changes in the heart. These different mechanisms explain most of the phenotypical variations within the heart failure population, with HFrEF typically driven by co-existing cardiac conditions such as myocardial infarction, while HFpEF tends to be linked to non-cardiac systemic conditions such as T2DM and obesity [5, 39]. Indeed, patients with obesity or HFpEF exhibit similar molecular and clinical features, while patients with both conditions concurrently tend to have an exacerbated symptom burden, be more prone to exercise intolerance, and have higher rates of hospitalisation for heart failure [40–42]. Backed by a growing body of experimental and clinical evidence, a recent unifying hypothesis postulates that dysfunctional visceral adipose tissue and secreted adipokines play a central role in the pathogenesis of HFpEF, strongly supporting the notion of an obesity-inflammatory HFpEF phenotype [43]. Throughout this section and indeed thesis, I will focus primarily on HFpEF rather than HFrEF.

2.4.2 Sex and age

Women are at typically at higher risk of HFpEF compared to men, while men are at higher risk of HFrEF [13, 15]. This trend can be attributed to multiple factors, from molecular- to organ-level scale. Pre-menopausal women benefit from the well-established cardioprotective effect of estrogens. Estrogens activate the receptors ER-alpha and ER-beta that modulate vasodilation, angiogenesis, and protecting against atherosclerosis and apoptosis in cardiac myocytes [44]. Estrogens also reduce collagen production in female cardiac fibroblasts, whilst it stimulates it in males. Thus, estrogen deficiency is

believed to be one of the major causal factors responsible for the development of HFpEF in post-menopausal women due to increased fibrosis formation, impaired angiogenesis, and capillary rarefaction [45]. Additionally, young men are more likely to develop inflammation and fibrosis than women, however cardiac aging and estrogen deficiency leads to more microvascular changes in older women [46]. The activation of these inflammatory mechanisms may be responsible for left ventricular concentric remodelling and diastolic dysfunction leading to HFpEF in elderly women. Differences in sex hormone binding globulin (SHBG), which controls testosterone levels, may also impact risk of heart failure. High SHBG levels typically reduce testosterone availability. One study showed that higher levels of SHBG correlated with heart failure severity and was associated with higher risk of death [47]. Meanwhile, low levels of testosterone have been associated with higher levels of NT-proBNP, myocardial damage, more severe heart failure symptoms and higher risk of death [48, 49].

Inherent differences in cardiac structure both at tissue and whole-organ level may also play a defining role in heart failure phenotypes both in men and women. Women typically have smaller and stiffer ventricles than men, smaller vasculature, increased ventricular wall thickness, increased myocardial blood flow, greater arterial stiffness and pulse pressure, contributing to higher risk of diastolic dysfunction and HFpEF [9, 50]. Men are more likely to develop macrovascular cardiovascular problems such as coronary artery disease and myocardial infarction, making them more at risk of HFrEF compared to microvascular HFpEF-causing dysfunctions in women [14].

2.4.3 Structural and ionic cardiac remodelling

Cardiac conditions such as myocardial infarction, atrial fibrillation, and hypertrophic cardiomyopathy are strongly associated with heart failure. Structural and functional changes in cardiac tissue arise following myocardial injury in myocardial infarction. Initially, remodelling at the cellular level is driven by loss of cardiomyocytes and influx of inflammatory cells, which destroy collagen. This leads to wall thinning and dilation in the infarct region, and subsequent scar formation due to increased levels of fibroblasts and collagen deposition. Following myocardial infarction, healthy tissue must compensate

for the lack of function of the impaired infarcted area. As time progresses, this causes an increase in left ventricular wall stress, dilation, and size, which can ultimately lead to increased end-diastolic and end-systolic volumes, and a reduction of the LVEF leading to HFrEF [51]. Left ventricular remodelling in hypertrophic cardiomyopathy is driven by multiple interacting mechanisms including mutations due to genetic factors, altered energetics and calcium homeostasis, and coronary microvascular dysfunction causing ischemia and fibrosis. The main remodelling phenotypes in hypertrophic cardiomyopathy include left ventricular wall thickening and cavity enlargement and wall thinning with fibrotic replacement and loss of left ventricular systolic and/or diastolic function [52]. The latter typically causes heart failure as a result of reduced blood output due to stiffer, weaker ventricles. In atrial fibrillation, atrial remodelling reduces the refractory period and slows conduction velocity at cellular level, and causes atrial dilation at tissue level [53].

Systemic metabolic conditions such as T2DM, chronic kidney disease, dyslipidemia, and obesity, cause microvascular endothelial inflammation, i.e. inflammation of the cells lining small blood vessels. This has a negative impact on cardiovascular function and may lead to HFpEF [5, 8, 39, 54]. The remodelling cascade in most systemic pro-inflammatory conditions is driven by the production of reactive oxygen species in the endothelium which limits nitric oxide availability, which in turn reduces PKG activity, causing an increase in cardiomyocyte resting tension and hypertrophy, and fibrosis. This leads to concentric left ventricular remodelling and diastolic left ventricular dysfunction, characterised by stiffness and impaired relaxation, hallmarks of heart failure [55]. Additionally, the newly-proposed unifying hypothesis of HFpEF suggests that an imbalance between cardioprotective and pro-inflammatory circulating adipokines, released from dysfunctional visceral adipose tissue in obesity, plays a central role in structural and functional cardiac remodelling [43]. This includes the development of hypertrophy and myocardial fibrosis, mediated through endocrine and paracrine signalling pathways.

Hypertension and pressure overload cause vascular dysfunction, leading to hypertrophy, fibrosis, and ultimately diastolic dysfunction both in the ventricles and atria [56]. The presence of hypertension has also been shown to be independently associated with diastolic dysfunction in diabetic patients [57]. Patients with coexistent diabetes and hypertension

have a higher incidence of coronary artery disease that may worsen myocardial dysfunction, while undiagnosed or "silent" myocardial ischemia may also complicate the diagnosis of diabetic myocardial disorder. The interplay of these conditions is clearly important to consider from a clinical perspective when diagnosing and treating patients who are likely to have two or more of these diseases in conjunction [19, 58].

At ionic level, in heart failure, there is a decrease in repolarising K^+ currents (I_{K1}, I_{Ks}, and I_{to}) causing impaired repolarisation and action potential prolongation. Increased depolarising currents (I_{Na}, I_{NaL}) and early afterdepolarisations are caused by an increase of inward currents and reduction of outward currents [59]. In healthy myocytes, Ca^{2+} entering through the L-type Ca^{2+} channel (I_{CaL}) triggers the release of Ca^{2+} from the sarcoplasmic reticulum through the Ca^{2+} release channel. During repolarisation, Ca^{2+} that has accumulated in the cytoplasm via I_{CaL} entry and sarcoplasmic reticulum release is removed by two main processes: Ca^{2+} extrusion via NCX and Ca^{2+} uptake into the sarcoplasmic reticulum by the SERCA pump. In heart failure, the Ca^{2+} release channel is hyperphosphorylated by beta-adrenergic-controlled protein kinase A and/or by CaMKII. This causes abnormal diastolic Ca^{2+} leak from the sarcoplasmic reticulum. NCX activity is enhanced, promoting the formation of delayed afterdepolarisation and reducing intracellular Ca^{2+} stores by extruding more Ca^{2+} out of the cell. SERCA function is reduced, which further impairs sarcoplasmic reticulum Ca^{2+} reserves and reduces myocardial contractility [59].

Many of the mechanisms described above are common to T2DM; cardiac remodelling driven specifically by T2DM will be explored in more detail in the next section of this chapter.

2.4.4 Imaging biomarkers

The most commonly used imaging marker of heart failure is the LVEF (see Section 2.4.1). It can be obtained from structural images of various modalities including echocardiography and magnetic resonance, using manual or automated methods to identify end-diastolic and end-systolic volumes. As explained earlier in Section 2.4.1, LVEF can confirm or refine an existing diagnosis by determining which type of heart failure is present [37].

Alternative imaging metrics have also been evaluated based on their prognostic value of heart failure. For instance, left ventricular volume to mass ratio is a common marker of left ventricular hypertrophy. In echocardiography, the ratio of early mitral inflow velocity to mitral annular early diastolic velocity (E/e' ratio) is a measure of diastolic dysfunction, suggestive of potential heart failure. Global longitudinal strain is a relatively novel measure of myocardium contractility, that has been shown to be independently associated with NT-proBNP, the main diagnostic blood-based biochemical marker for heart failure; severity of heart failure; and cardiac death [60, 61]. This alternative marker may help improve the prognostic value of imaging biomarkers of heart failure, either in conjunction with or instead of LVEF.

2.4.5 Pharmacological treatment

There exists a range of approved therapeutic agents for heart failure, with varying effects in different subgroups of disease. The current European Society of Cardiology guidelines for heart failure therapy recommend the administration of specific medication based on a patient's symptom severity and diagnosis [37, 38]. Until recently, the European Society of Cardiology recommendations concerned primarily patients with HFrEF, based on more significant changes in outcomes in clinical trials of the drugs above in patients with HFrEF rather than HFpEF. HFpEF therapy mostly relied on targeting the underlying etiological signalling cascade rather than heart failure itself. Given the phenotypic diversity of HFpEF, this resulted in a plethora of treatment strategies based on different patient phenotypes and predispositions [7, 9]. I will explore this further in the next section.

Despite the clinical heterogeneity of HFpEF, a number of pharmacological treatments across different drug classes have proven effective in reducing adverse outcomes in patients with HFpEF. As demonstrated by the DELIVER and EMPEROR-Preserved trials, sodium-glucose cotransporter 2 (SGLT2) inhibitors dapagliflozin and empagliflozin reduced combined risk of worsening heart failure or cardiovascular death, and combined risk of cardiovascular death or hospitalisation for heart failure, respectively, in patients with HFpEF [10, 11]. The significant results of these two trials have led to a modification of the most recent European Society of Cardiology guidelines. Now, dapagliflozin

and empagliflozin are included as Class I, i.e. most highly recommended, drugs for all HFmrEF and HFpEF patients [38]. SGLT2 inhibitors were originally developed to treat patients with diabetes, their main purpose and mechanism being to reduce blood glucose levels. However, they also seem to exert a direct and beneficial effect on the heart, regardless of the presence or absence of diabetes. SGLT2 inhibitors have been shown to affect myocardial contractility, which may potentially be mediated by inhibitory effects of the drug on the sodium-hydrogen exchanger and subsequent changes in Ca^{2+} homeostasis [62, 63]. In a study on zebrafish ion channels, empagliflozin caused a significant increase in I_{Kr} and I_{Ks}, thus shortening the action potential duration [64]. This could help counteract cases of delayed repolarisation in diabetes and heart failure, however the exact mechanisms of action in human are yet to be elucidated.

Another medication class of emerging importance in the management of HFpEF is glucagon-like peptide-1 (GLP1) receptor agonists. Semaglutide, a GLP1 agonist, was shown to improve symptoms and physical limitations of patients with obesity and HFpEF in the STEP-HFpEF trial [65, 66], with an additional link established between body weight reduction and magnitude of treatment effect [67]. While these results point toward the relevance of identifying and targeting modifiable risk factors such as obesity to mitigate disease severity and improve outcomes in these patients, follow-up times were short and observed reductions in composite endpoints appear largely driven by fewer cardiovascular deaths, with minimal effect on heart failure hospitalisations. Results from the more recent SUMMIT trial assessed the effects of tirzepatide, another GLP1 agonist, on a composite endpoint consisting of death from cardiovascular causes or worsening heart failure, in patients with HFpEF and obesity [68]. Definitive conclusions about whether GLP1 agonists directly modify the progression of heart failure itself remain to be established, prompting the need for further trials of GLP1 agonists in well-phenotyped heart failure cohorts [69].

While mineralocorticoid receptor (MRA) antagonists are one of the main pillars of HFrEF pharmacological therapy [38], their benefits in HFpEF remained unclear until recently. The FINEARTS-HF trial showed that non-steroidal MRA finerenone significantly reduced the risk of cardiovascular death and hospitalisation in patients with HFpEF or

HFmrEF [70], with the added benefit of conferring a lower risk of hormone-related side effects compared to steroidal MRA counterparts such as spironolactone and eplerenone. Given these results, it is likely that upcoming guidelines for the management of HFpEF will be updated to reflect the use of finerenone as a recommended treatment option.

In addition to etiology-specific differences, there exists many sex-specific variations in treatment response which reinforce the need for a personalised approach to heart failure therapy. A significant difference in reduction of heart failure hospitalisation risk was found between women and men in the MERIT-HF beta blockers trial, where risk in women was found to be reduced 32% further than in men [71]. Women with hypertension taking beta blockers, however, are typically more at risk of developing heart failure [72]. In the angiotensin II receptor blocker CHARM trial, women benefitted more significantly from candesartan therapy than men in terms of reduction in heart failure hospitalisation and cardiovascular mortality [73]. Sacubitril/valsartan, the only FDA-approved angiotensin receptor/neprilysin inhibitor combination drug, is typically used to reduce risk of death and hospitalisation for HFrEF patients who do not respond to angiotensin-converting enzyme inhibitors, beta blockers, or mineralocorticoid receptor antagonists. In the PARADIGM-HF angiotensin receptor/neprilysin inhibitor trial, prognostic for cardiovascular-related death or first heart failure hospitalisation was the same in both sexes. When considering cardiovascular death alone, however, the treatment significantly improved prognosis in men but not in women [74]. As for mineralocorticoid receptor antagonists, results of the RALES and EPHESUS trials did not suggest any sex differences in prognosis, however women represented less than 30% of the patient cohorts in both trials [75, 76]. The underrepresentation of women in clinical trials of heart failure is common, and represents an urgent need to address, in order to ensure that statistically powerful and thus significant sex-specific results can be obtained.

2.4.6 Data-driven clustering and phenomapping studies of HFpEF

The diagnostic and prognostic utility of biomarkers in HFpEF, whether derived from imaging or blood biochemistry such as NT-proBNP, is highly dependent on the underlying etiology and stage of disease. Single biomarkers in isolation often fail to capture the

complex, heterogeneous nature of HFpEF, limiting their effectiveness in guiding clinical decision-making. Furthermore, as seen in the previous section, pharmacological treatment of HFpEF is a major clinical challenge.

Integrating multiple biomarker modalities enables the identification of pathophysiologically distinct patient subgroups, offering a more precise and individualised characterisation of a given condition. In systems medicine, this data-driven approach known as phenomapping[1] leverages high-dimensional multi-modal datasets to disentangle the heterogeneity of heart failure presentations. It allows for the stratification of patients beyond single diagnoses and into groups of similar clinical profiles, potentially underpinned by common pathophysiological mechanisms, which may lead to divergent disease trajectories and treatment responses.

In this section, I review methods and findings from existing phenogrouping studies of HFpEF that have used large patient cohorts and powerful statistical methods to leverage the wealth of information contained across different types of relevant clinical data, optimising classification and disease management in HFpEF.

Clustering

The first major phenotyping study of heart failure considered demographics, physical characteristics, ECG, and echocardiography data from 397 patients with HFpEF [77]. Shah et al. used hierarchical clustering of phenotypical variables followed by a support vector machine, a supervised machine learning algorithm that finds the optimal "hyperplane" to separate data into distinct classes or clusters by maximising the margin between them. The three main clusters extracted contained younger patients with low BNP levels, with the least extent of myocardial remodelling and dysfunction; patients with obesity and diabetes and high levels of fasting glucose, with poor left ventricular relaxation and high pulmonary vascular resistance; older patients with high serum creatinine and low glomerular filtration rate indicative of kidney disease, high BNP levels, with most severe remodelling including

[1] Broadly, phenomapping or phenogrouping refers to a data-driven exploratory approach aiming to discover new groupings based on patterns emerging from multi-modal data, while phenotyping is more hypothesis-driven and tends to categorise based on predefined features and grounded clinical knowledge. In this chapter, I use them interchangeably, but my own work presented in Chapter 4 should be considered explorative phenomapping rather than phenotyping.

longer QRS duration, highest cardiac wall thickness and left ventricular mass, highest E/e' ratio, high pressure to stroke volume ratio. Significant variations in ECG characteristics, cardiac structure and function, and hemodynamic data were observed across the three groups. This work formed the basis of another study proposing personalised therapeutic strategies for HFpEF, targeting distinct signalling pathways and phenotypes rather using than a population-wide treatment approach [7]. Another study used over 300 clinical features including medical history, physical examinations, echocardiography, circulating blood biomarkers to identify subgroups of HFpEF patients amongst a cohort of >1700 patients [78]. They used principal component analysis to reduce feature dimensions, hierarchical clustering to cluster biomarkers, and associated these clusters with heart failure hospitalisation and all-cause mortality. They found four major clusters which included primarily patients with diabetes; elderly patients with age-related comorbidities; younger patients with few comorbidities aside from obesity and with lowest NT-proBNP levels and risk of adverse outcomes; and ischemic patients who typically smoke, have lung disease, and high levels of NT-proBNP and troponin. Instead of using clinical biomarkers, a similar grouping approach was applied using natural language processing and k-means clustering on text-based electronic health records of >29,000 heart failure patients [79]. The study found 15 distinct patient phenotypes, each correlating to different extents with different co-morbidities, and certain clusters additionally being sex-specific.

Risk prediction

Risk prediction models for heart failure have been developed using large patient cohorts and multi-modal data. The three groups established in the study by Shah et al. described above correlated with lowest, medium and highest risk of adverse heart failure-related outcomes, respectively [77]. Unsupervised machine learning methods have also been applied successfully to identify phenogroups of HFpEF patients and establish risk profiles based on subgroup characteristics [80]. Finite mixture model-based clustering was applied to the TOPCAT clinical trial data, to extract three subgroups based on various metabolic and cardiovascular features. The most at-risk group consisted of patients with multiple cardiometabolic risk factors and comorbidities, including left ventricular abnormalities

and elevated natriuretic peptide levels. Another study showed the effectiveness of using clinical phenogroups instead of LVEF to stratify a population of >2400 at-risk heart failure patients based on demographic information, circulating blood biomarkers, comorbidities, and medication, using unsupervised machine learning-based clustering [81]. They found that better risk stratification was achieved by using six different clinical subpopulations (no comorbidities, coronary artery disease, atrial fibrillation, sleep apnoea, severe valvular heart disease, chronic obstructive pulmonary disease) rather than using HFrEF, HFmEF and HFpEF. Another study considered >28,000 patients across four cohorts, of which one was used to validate the models, with longitudinal follow-up data tracking incident rates of heart failure over a period of 12 years [82]. They built predictive risk models for different types of heart failure; the HFpEF model included age, sex, systolic blood pressure, body mass index (BMI), anti-hypertensive treatment, and previous myocardial infarction, while the HFrEF model included the same parameters with the addition of smoking, left ventricular hypertrophy, left bundle branch block and diabetes. They found that age, obesity, hypertension, and previous myocardial infarction were the main risk factors leading to both types of heart failure. Myocardial infarction, a major risk factor for heart failure, had been used as a clinical indicator to stratify heart failure patients in another phenotyping study [83–85]. This study used a combination of circulating blood biomarkers, comorbidities, and medication to predict incidence of heart failure and observe the differences in disease development in patients with and without prior myocardial infarction. Without infarction, it was found that metabolic- and hypertension-related factors were associated with onset heart failure, whereas kidney dysfunction and smoking status were stronger predictors of disease onset in patients with prior myocardial infarction. They hypothesised that neurohormonal activation and increased inflammation linked to adverse myocardial effects, and volume overload contribute towards progressive heart failure in post-infarction patients.

Treatment outcomes

Other phenotyping studies have linked patient subgroups to different responses to heart failure treatment. One study performed phenomapping in patients undergoing

HFpEF treatment, with angiotensin receptor blockers irbesartan and candesartan in the I-PRESERVE and CHARM trials, respectively [86]. The main clinical features in this work included basic patient demographics and markers of comorbidities, both diagnosis- and blood biochemistry-based. Using latent class analysis (see Chapter 3, Section 3.3.4), six main clinical subgroups were extracted. Age, gender, atrial fibrillation prevalence, BMI, diabetes, and renal dysfunction were determining factors in distinguishing between groups. Each group was clearly linked with different levels of adverse outcomes following treatment. Two other studies also focused on HFpEF patients, treated with the mineralocorticoid receptor antagonist spironolactone in the TOPCAT study [87, 88]. One of the studies used a combination of patient characteristics, comorbidities, blood biomarkers, LVEF, and QRS duration, and extracted three main phenogroups using latent class analysis to explain confounding between related variables. The three subgroups included: younger patients with few comorbidities, older patients with prevalent AF and markers of impaired immunity function, and relatively young patients with many comorbidities including diabetes. Each group exhibited different responses to treatment, with the most beneficial impact observed in the first group and no significant benefit in the other two [87]. The second study based on the same trial found three similar phenogroups, and additionally suggested that each group exhibited different extents of cardiac tissue remodelling. This study differed in its results of treatment response, finding that spironolactone had beneficial effects in diabetic and obese patients but not in young patients with few comorbidities [88].

Salient clinical phenotypes and outlook

The studies described above highlight the crucial yet complex role of age and sex, as well as cardiometabolic diseases including obesity, diabetes, and hypertension, in the clinical landscape of HFpEF. The importance of diabetes as a key subgroup of HFpEF is further supported by previous cohort studies and ancillary findings in the RELAX trial of phosphodiesterase 5 inhibitor sildenafil (Viagra) for HFpEF, which showed that patients with diabetes tend to have a more severe comorbidity burden and distinctive pathophysiology, and typically worse outcomes before and after treatment [89, 90]. Obese patients with HFpEF were described more precisely in another study which revealed

higher levels of diabetes, distinct cardiovascular remodelling leading to structural and functional alterations, and more impaired haemodynamic behaviour in obese patients with HFpEF compared to the non-obese patients [91]. These subgroups are clearly connected and align with existing mechanistic hypotheses behind systemic inflammation and oxidative stress as drivers of HFpEF.

Overall, the studies presented in this section highlight the complexity of HFpEF, challenge its status as a single clinical condition, and motivate the need for further exploration of clinical subgroups of HFpEF and sex-specific differences as a strategic priority to optimise the management of the condition.

2.5 Literature review: type 2 diabetes-driven cardiac disorders

In this final section, I review adverse cardiac adaptations in T2DM, a major yet comparatively understudied clinical risk factor for HFpEF. I present the key pathophysiological mechanisms of electrophysiological, structural and nervous cardiac remodelling in T2DM through the review of experimental and clinical studies, along with the cardiac implications of pharmacological therapies commonly taken by patients with T2DM. Notably, I emphasise the relevance of these factors in the development of HFpEF, cardiac arrhythmias and sudden cardiac death.

2.5.1 Epidemiology and cardiovascular risk of type 2 diabetes

T2DM is a complex metabolic disease that affects over half a billion adults globally [18]. Its development is influenced by diet, body weight, and age; its prevalence increasing steadily due to obesity and aging populations. T2DM is characterised by chronically high levels of blood glucose linked to impaired insulin secretion and insulin resistance, which causes abnormal glucose absorption from the blood into different organs. Insulin being the only hormone that can lower blood glucose concentration, these combined factors contribute to chronic hyperglycaemia, the central trigger responsible for a cascade of adverse mechanisms affecting the heart.

Individuals with T2DM have a two- to four-fold increased risk of cardiovascular disease [19], notably a higher risk of HFpEF (Section 2.4.1), cardiac arrhythmias including atrial fibrillation [22, 92–94] and ventricular arrhythmias [22, 95–97], which are linked to sudden cardiac death. Arrhythmic risk is further increased in patients with concomitant HFpEF and T2DM, which make up 20-45% of HFpEF cases [98–100]. Silent myocardial infarction and ischemia are also more common in T2DM, further increasing the risk of heart failure, coronary artery disease, and death [101–103]. However, recent European Society of Cardiology clinical guidelines suggest that sudden cardiac death in diabetes is usually related to underlying structural cardiac abnormalities. Currently, there exists no diabetes-specific protocol for sudden cardiac death screening, which is centred instead on the presence of other co-existing conditions [19, 104].

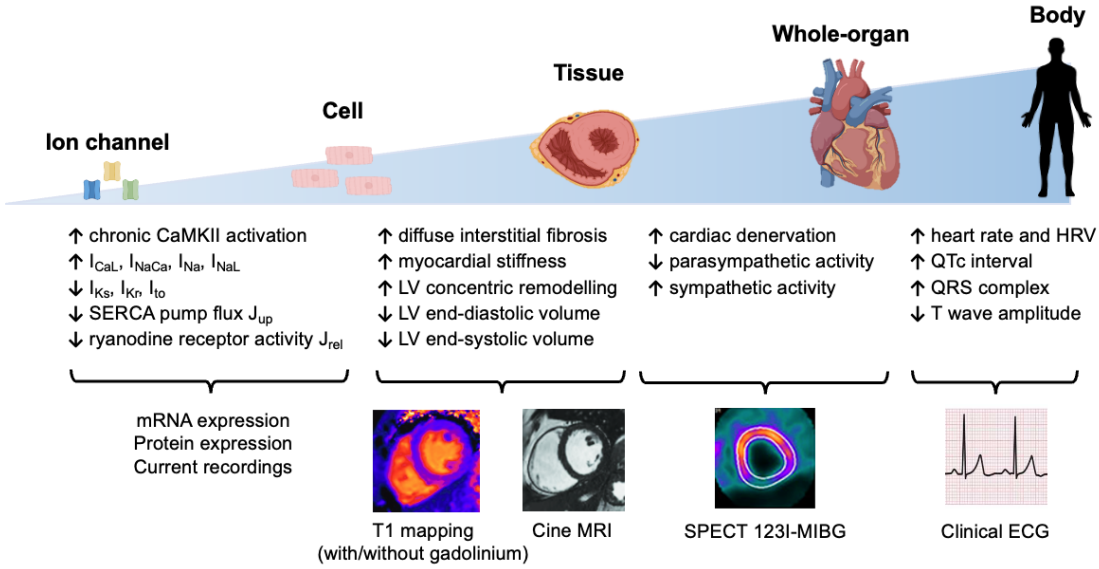

Figure 2.5: Overview of multi-scale cardiac remodelling in type 2 diabetes and how to measure its effects. Remodelling occurs at ionic channel, protein, cellular, tissue and whole-organ level, affecting the electromechanical function and structure of the diabetic heart. Experimental and clinical data can capture some of these changes. HR: heart rate, HRV: heart rate variability, LV: left ventricular, MIBG: iodine meta-iodobenzylguanidine, MRI: magnetic resonance imaging, SPECT: single photon emission computed tomography. This figure was made with icons from Biorender.

2.5.2 Structural cardiac remodelling

The structural and functional effects of diabetes-induced cardiac remodelling at ventricular level, which closely resemble heart failure, have been formally referred to as diabetic

cardiomyopathy since the 1970s [105]. This condition was defined as the occurrence of myocardial structural and/or functional abnormalities associated with T2DM, in the absence of coronary heart disease, hypertension and/or obesity. Given that the majority of T2DM patients nowadays have one or more of these contributing comorbidities, this definition was recently relaxed. The condition is now referred to as diabetic myocardial disorder and is considered a T2DM-specific pre-heart failure stage that warrants careful clinical attention [106]. In this section, we discuss some of the hallmarks of diabetic myocardial disorder at tissue and whole-organ level.

Anatomical remodelling

Patients with T2DM tend to have a higher left ventricular mass [107, 108] and wall thickness [107]. These changes worsen as the disease progresses [109, 110]. Overt left ventricular hypertrophy is frequently observed in patients with T2DM [108], characterised by an increased left ventricular mass to left ventricular end-diastolic volume ratio, but a normal left ventricular mass index [111]. This is thought to be caused in part by cardiac steatosis, the excess accumulation of triglyceride in myocytes. Cardiac steatosis causes enhanced hypertrophic signalling and concentric remodelling, hyperactivation of the insulin signalling cascade in patients with obesity and T2DM patients, and increased levels of circulating pro-inflammatory cytokines [112, 113].

Fibrosis

At tissue level, cardiac remodelling in T2DM is mainly characterised by the presence of diffuse, interstitial fibrosis. As opposed to dense replacement fibrosis, which occurs due to cardiomyocyte death typically after an infarction and is mostly permanent, diffuse fibrosis is more widespread, less dense, and generally reversible [114]. In hyperglycemia, the increased production of reactive oxygen species and advanced glycation end products cause cross-linkage of extracellular matrix proteins in the myocardium [115, 116]. Collagen, the key structural component of the extracellular matrix, is non-conductive, hindering electrical wave propagation. Additional sources of fibrosis in T2DM include the activation of cardiac fibroblasts into myofibroblasts, impaired autophagy of apoptotic

or necrotic cells, and elevated sympathetic nervous system activity, which enhances beta-1 adrenergic receptor signalling [117–119].

A post-mortem study identified diffuse fibrosis throughout the myocardium of four middle- to old-age diabetic patients with a history of diabetes of 5 or more years [105], while another study examining human myocardial biopsies confirmed that interstitial fibrosis was elevated in patients with T2DM compared to non-diabetic patients [120]. In healthy tissue, diffuse fibrosis can be present in 6% of the myocardium [121, 122]. This proportion can increase to up to 20-35% in cardiac disorders with (micro)structural abnormalities [121, 123] including T2DM [124, 125], obesity [126] and HFpEF [127]. Clinical and experimental evidence on the pathological changes in diabetes and obesity suggest that the extent of fibrosis correlates with severity of disease, both in late-stage vs. early-stage diabetes [105, 118, 128, 129] and in obesity [126], where the amount of epicardial adipose tissue and therefore fibrosis is directly related to general body adiposity [130]. Using H(1) magnetic resonance spectroscopy, Levelt et al. showed that cardiac steatosis is increased in T2DM patients [111]. One CMR study showed that epicardial adipose tissue, which increases with BMI, is positively associated with myocardial fibrosis and cardiac steatosis [131]. Another CMR study looking at late gadolinium enhancement lesions and extracellular volume reported heterogeneities in the distribution of fibrosis in T2DM, suggesting that there is a higher proportion of fibrotic tissue in the basal and mid-ventricular septum, inferior, and inferolateral segments of the heart [125]. Fibrosis and epicardial adipose tissue both contribute to contractile dysfunction and an increased likelihood of arrhythmias via delayed conduction and the formation of re-entry circuits [121, 132].

Diastolic dysfunction

Fibrosis causes tissue to be stiffer, impairing normal contraction and relaxation of the myocardium in T2DM. Initially, this manifests as substantial changes in diastolic function, and may lead to systolic dysfunction at later, more severe stages of T2DM [116]. Impaired myocardial relaxation can be identified at early stages by CMR by a decrease in initial and peak filling rates, or by echocardiography by abnormal wall motion and longer

isovolumic relaxation times [119]. At later stages, coronary microcirculation becomes further impaired, and diastolic and systolic dysfunction become more severe. Indeed, in advanced T2DM, diastolic dysfunction and other symptoms start resembling the HFpEF phenotype more closely [133]. In addition to structural causes of diastolic dysfunction, abnormal Ca^{2+} handling in T2DM with higher $[Ca^{2+}]_i$ levels and longer $[Ca^{2+}]_i$ decay leads to prolonged relaxation times (see Section 2.5.4).

2.5.3 Electrophysiological remodelling at tissue- and whole-organ level

As discussed in the previous section, diffuse fibrosis represents an important substrate for cardiac arrhythmias in T2DM. Remodelling of connexin 43 proteins and damage to the systemic and intrinsic cardiac nervous system further exacerbate the arrhythmic substrate in T2DM.

Connexin 43 and impaired conduction

Along with their role in the formation of fibrosis, myofibroblasts express gap junction connexin proteins Cx43 and Cx45 which enable them to directly modulate cardiomyocytes' electrophysiological behaviour, including a reduction in cardiac impulse conduction [134]. Cx43 appears to be lateralised in T2DM, i.e. redistributed along the lateral sides of cardiomyocytes rather than the cell-cell junctions, and associated with non-functional gap junctions [135]. Furthermore, Cx43 expression levels are reduced overall in diabetes [136]; this reduction is spatially heterogeneous in the ventricles [137]. Analogous findings on the spatially heterogenous reduction of Cx43 have been reported in studies of related myocardial disorders such as heart failure in humans [138] and hypertrophy in rats [139]. These Cx43 alterations are associated with heterogeneities in resting membrane potentials, action potential duration and refractory period, leading to spatially dispersed current conduction in the myocardium and thus to an enhanced vulnerability to ventricular arrhythmias, even in the absence of fibrosis [140, 141]. These changes may explain the QT dispersion observed clinically in the ECG.

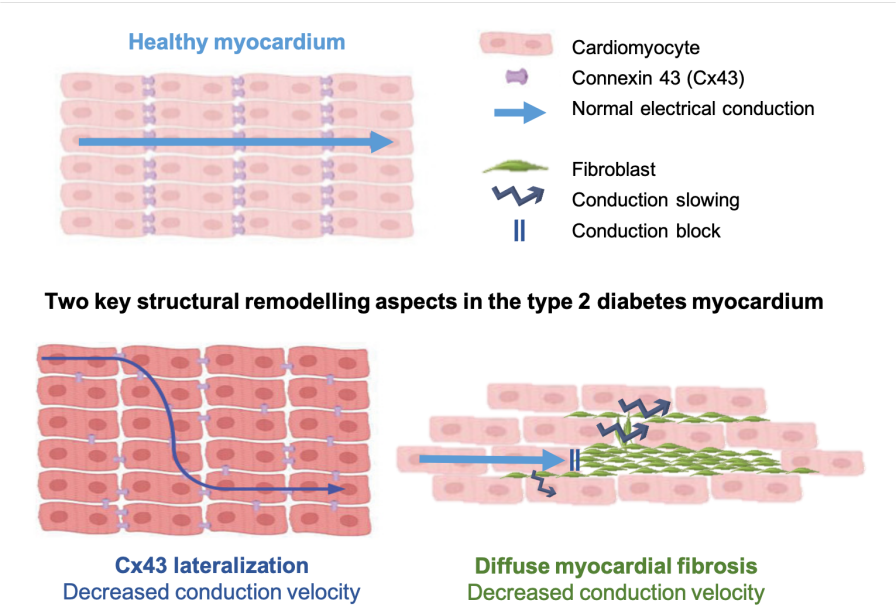

Figure 2.6: Effect of connexin 43 lateralisation on myocardial conduction velocity. Figure adapted from [142].

Cardiac autonomic dysfunction

Chronic hyperglycemia is responsible for the progression of cardiac autonomic neuropathy, another common cardiac complication of T2DM which affects the intrinsic cardiac nervous system [143, 144]. The intrinsic cardiac nervous system is made up of post-stellate ganglionic, sympathetic nerve fibres and parasympathetic nerve fibres descending from the vagus nerve [145]. Neurons emerging from the intrinsic cardiac nervous system cluster in distinct areas of the heart to form various ganglionated intracardiac plexi [146, 147]. In T2DM, intracardiac ganglia undergo structural and functional remodelling, while the cardiac nerves connected to the parasympathetic fibres at the atrioventricular and sinoatrial nodes are damaged progressively from the apex of the ventricles to the base [148, 149].

In the early stages of T2DM-induced cardiac autonomic neuropathy, parasympathetic function decreases relative to sympathetic function which increases, causing heart rate variability to decrease [150, 151]. Heightened sympathetic activity stimulates the renin-angiotensin-aldosterone system and increases heart rate, stroke volume and vascular resistance, leading to high blood pressure and ventricular dysfunction [152, 153]. At later stages, resting tachycardia is common, as well as both severely impaired heart rate

variability and blunted cardiac output response following exercise, indicating almost complete cardiac denervation [151]. Cardiac denervation in T2DM is likely to blunt patients' perception of myocardial pain, explaining the higher risk of silent myocardial ischemia and further subclinical cardiac decline [150]. Cardiac denervation may also explain the higher rates of atrioventricular block and bundle-branch block in T2DM [154–156]. This could be explained by damaged atrioventricular and sinoatrial nodes, which may have a downstream impact on electrical signal propagation through the Bundle of His and downstream Purkinje fibres responsible for the control and rapid distribution of electrical activity through the ventricles.

Cardiac sympathetic denervation has been shown to correlate with diffuse fibrosis and modulate arrhythmic risk in non-ischemic cardiomyopathies, which includes diabetic cardiomyopathy [157]. Epicardial adipose tissue is also significantly correlated with increased cardiac denervation [158], while adipocytes themselves are known to directly release catecholamines, notably norepinephrine and epinephrine, which further modulates sympathetic activity in the heart [159]. The imbalance of sympathetic and parasympathetic activity causes an increase in dispersion of electrical activation and repolarisation in the ventricles from the endocardium to epicardium, creating a more pro-arrhythmic setting [144]. Even in the absence of overt nerve fibre degeneration, sympathetic dysfunction is associated with an increased susceptibility of ventricular arrhythmias in mice with T2DM [96].

Clinically, characteristic autonomic effects may be reflected in the ECG, notably the QT-RR slope and QTc interval prolongation [96, 160], while cardiac denervation and the sympathetic nervous system can be assessed using single-photon emission computed tomography imaging with an iodine-123 metaiodobenzylguanidine radioactive tracer [157].

2.5.4 Cellular electrophysiological remodelling and calcium handling

While certain electrophysiological disturbances in T2DM may be observed clinically in the ECG, experimental studies remain crucial to unravel the underlying mechanisms responsible for those changes and the trigger of cardiac arrhythmias.

Ashrafi et al. collected ventricular tissue mRNA data from patients with T2DM and translated these results to relative current changes [161]. They found an increase in L-type calcium current (I_{CaL}), sodium-calcium exchanger current (I_{NaCa}), sodium current (I_{Na}), late sodium current (I_{NaL}); a decrease in slow and rapid delayed rectifier potassium currents (I_{Ks}, I_{Kr}), transient outward potassium current (I_{to}); and a decrease in SERCA pump flux (J_{up}) and ryanodine receptor activity (J_{rel}). Multiple animal studies on ionic current changes in T2DM have been conducted, however results are heterogenous [162].

In this section, I review important mechanisms underlying subcellular electrophysiological changes in T2DM. Key changes in ionic current and calcium handling protein remodelling; CaMKII activity; and action potential, calcium transient and contractility biomarkers recorded in experimental studies of T2DM are recapitulated in Tables 2.2, 2.3, 2.4, 2.5, 2.6 and summarised visually in Figure 2.7.

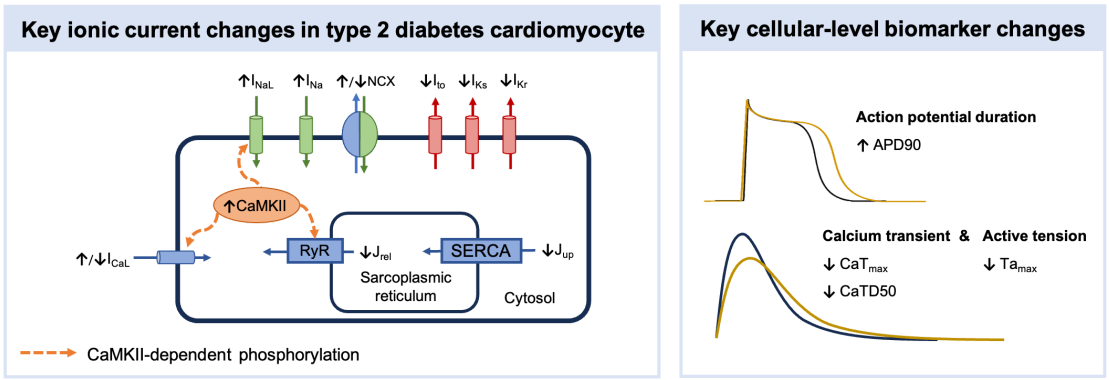

Figure 2.7: Increase and decrease in ionic currents and CaMKII are indicated by up and down arrows, respectively. Factors most likely to drive the APD prolongation are predominantly the increase in I_{NaL}, I_{NCX} and I_{CaL} when present, as well as the reduction in I_{Kr} and I_{Ks}. The reduced calcium transient amplitude is multifactorial, with likely key factors being the increased NCX activity, reduced SERCA pumps and increased SR leak, although these are partly offset by increased I_{NaL} and longer calcium influx via I_{CaL}, a consequence of the prolonged APD. APD, action potential duration; CaMKII, Ca^{2+}/calmodulin-dependent protein kinase II; CaT, intracellular calcium transient; D50, duration at half-maximum amplitude; SR, sarcoplasmic reticulum; Ta, active tension.

Table 2.2: Experimental data of ionic channel remodelling observed in cardiac ventricular myocytes of humans and/or animals with type 2 diabetes. Units of change from baseline are indicated in each row depending on the type of data; percent change for protein expression data and pA/pF for ionic currents. HP: holding potential, LTCC: L-type calcium channel, mo: months, PCR: polymerase chain reaction, SD: standard deviation, STZ: streptozotocin, yrs: years.

Ionic current	Reference	Change from baseline (approx.)	HbA1c (mmol mol^{-1}, mean ± SD)	Blood glucose (mg/dL, mean ± SD)	Experimental model and mean age	Experimental protocol and conditions
I_{Na}	Ashrafi et al., 2017 (Table 6) [161]	+21.08%	57.9 ± 14.57	–	Human (n=7) 74.9 yrs	Nav1.5 mRNA expression by quantitative PCR.
I_{to}	Ashrafi et al., 2017 (Table 6) [161]	+26.04%	57.9 ± 14.57	–	Human (n=7) 74.9 yrs	Kv1.4, 4.2 and 4.3 mRNA expression by quantitative PCR.
I_{CaL}	Ashrafi et al., 2017 (Table 6) [161]	+14.18%	57.9 ± 14.57	–	Human (n=7) 74.9 yrs	Cav1.2 and 1.3 mRNA expression by quantitative PCR.
I_{CaL}	Pereira et al., 2006 (Fig. 2) [163]	–4 pA/pF (–31.2%)	–	–	Db/db mice (n=20) 15 wks	Patch-clamp; HP = 100 mV; voltage steps to 20 mV; >180 sweeps; recorded at 21–23 °C.
I_{CaL}	Pereira et al., 2006 (Fig. 3) [163]	–38%	–	–	Db/db mice (n=5) 15 wks	Expression of pore-forming 1Cα subunit of LTCC by Western blotting.
I_{Kr}	Ashrafi et al., 2017 (Table 6) [161]	–65.24%	57.9 ± 14.57	–	Human (n=7) 74.9 yrs	ERG mRNA expression by quantitative PCR.
I_{Ks}	Ashrafi et al., 2017 (Table 6) [161]	–5.38%	57.9 ± 14.57	–	Human (n=7) 74.9 yrs	KvLQT1 mRNA expression by quantitative PCR.
I_{NaL}	Lu et al., 2013 (Fig. 2) [164]	+0.1 pA/pF (+100%)	–	>500	Db/db mice (n=7) 2.5 mo	Current elicited by 750 ms depolarising voltage steps from HP = –80 to 0 mV (10 mV increments); recorded at room temperature.

Table 2.3: Experimental data of ionic current/flux remodelling observed in cardiac ventricular myocytes of humans and/or animals with type 2 diabetes. Units of change from baseline are indicated in each row depending on the type of data; percent change for protein expression data and pA/pF or sparks/s for ionic fluxes. HP: holding potential, LTCC: L-type calcium channel, mo: months, PCR: polymerase chain reaction, SD: standard deviation, STZ: streptozotocin, wks: weeks, yrs: years.

Ionic current/flux	Reference	Change from baseline (approx.)	HbA1c (mmol mol^{-1}, mean ± SD)	Blood glucose (mg/dL, mean ± SD)	Experimental model and mean age	Experimental protocol and conditions
I_{NaCa} (NCX)	Ashrafi et al., 2017 (Table 6) [161]	+243.83%	57.9 ± 14.57	–	Human (n=7) 74.9 yrs	NCX1 mRNA expression by quantitative PCR.
I_{NaCa} (NCX)	Hattori et al., 2000 [165]	45%	–	605 ± 9	STZ rat (n=26) 12–14 wks, 8 wks old at STZ injection	Whole-cell voltage clamp, ramp pulses from HP = 60 mV to +50 mV, then 100 mV to 60 mV, at steady speed of 0.67 V s^{-1}. Recorded at 36 ± 1°C.
J_{rel} (RyR)	Ashrafi et al., 2017 (Table 6) [161]	10.07%	57.9 ± 14.57	–	Human (n=7) 74.9 yrs	RYR2 mRNA expression by quantitative PCR.
J_{rel} (RyR)	Teshima et al., 2000 (Fig. 2) [166]	38%	–	773 ± 17	STZ rat (n=4) "mature", data collected 12 wks post-STZ injection	RYR mRNA expression by Northern blotting.
J_{rel} (RyR)	Netticadan et al., 2001 (Fig. 1) [167]	45%	–	322 ± 24	STZ rat (n=4) undisclosed age, data collected 6 wks post-STZ injection	RYR mRNA expression by Western blotting.
J_{up} (SERCA2)	Pereira et al., 2006 [163]	1.3 Ca^{2+} sparks/s/100 µm (34%)	–	–	Db/db mice (n=14) 15 wks	Line-scan recording in saponin-permeabilised myocytes. Recorded at room temperature (21–23°C).
J_{up} (SERCA2)	Belke et al., 2004 (Fig. 5) [168]	13%	–	489 ± 44	Db/db mice (n=9) 12 wks	SERCA 2a expression by Western blotting.
J_{up} (SERCA2)	L. Zhang et al., 2008 (Fig. 4) [169]	50%	–	553.2 ± 16.2	STZ Wistar rat (n=8) "adult"	SERCA 2a expression by Western blotting.

Table 2.4: Experimental data of CaMKII activity observed in cardiac ventricular myocytes of humans and/or animals with type 2 diabetes. Units of change from baseline are indicated in each row depending on the type of data. PCR: polymerase chain reaction, HP: holding potential, HIP: Human Islet Amyloid Polypeptide, STZ: streptozotocin, LTCC: L-type calcium channel, CFP: cyan fluorescent protein, YFP: yellow fluorescent protein, ZDF: Zucker Diabetic Fatty.

CaMKII activity (type)	Reference	Change from baseline (approx.)	Experimental model	Experimental protocol and conditions
Relative CaMKII activity	Erickson et al., 2013 (ext. Fig. 6c)	+110%	HIP rat (n=3)	See Despa et al., 2012 for detailed experimental methods.
Autonomous CaMKII activation	Erickson et al., 2013 (Fig. 1b)	+40%	HIP rat (n=3) glucose 240 mg/dL	Autonomous CaMKII activation assay under moderate hyperglycemia.
Autonomous CaMKII activation	Erickson et al., 2013 (Fig. 1b)	+70%	HIP rat (n=3) glucose 500 mg/dL	Autonomous CaMKII activation assay under severe hyperglycemia.
Ratio of fluorescence of CFP to YFP (CaMKII activation state following increased $[Ca^{2+}]_i$)	Erickson et al., 2013 (Fig. 1e)	+10%	Human (n=6) glucose 100 mg/dL	0.5 Hz pacing.
Ratio of fluorescence of CFP to YFP (CaMKII activation state following increased $[Ca^{2+}]_i$)	Erickson et al., 2013 (Fig. 1e)	+20%	Human (n=6) glucose 200 mg/dL	0.5 Hz pacing.
Ratio of fluorescence of CFP to YFP (CaMKII activation state following isoprenaline stimulation)	Erickson et al., 2013 (Fig. 1e)	+25%	Human (n=9) glucose 100 mg/dL	Isoprenaline (100 nM for 20 min).
Ratio of fluorescence of CFP to YFP (CaMKII activation state following isoprenaline stimulation)	Erickson et al., 2013 (Fig. 1e)	+30%	Human (n=9) glucose 200 mg/dL	Isoprenaline (100 nM for 20 min).
CaMKII phosphorylation	Daniels et al., 2018	+82%	ZDF rat (n=8)	CaMKII band intensities normalised to GAPDH; ratio of phosphorylated CaMKII relative to total CaMKII measured by Western blot.

Table 2.5: APD_{90} biomarkers in cardiac ventricular myocytes of humans and/or animals with type 2 diabetes. STZ: streptozotocin, APD: action potential duration.

Biomarker	Reference	Change from baseline (approx.)	Experimental model	Experimental protocol and conditions
APD_{90}	Lu et al., 2013 (Fig. 1)	+130 ms	Db/db mice (n=20) glucose >500 mg/dL	–
APD_{90}	Lu et al., 2013 (Fig. 1)	±20 ms	Db/db mice (n=10) glucose >500 mg/dL	PI3K messenger, PIP3.
APD_{90}	Lu et al., 2013 (Fig. 1)	±20 ms	Db/db mice (n=7) glucose >500 mg/dL	I_{NaL} blocker mexiletine (4 mg/mL for 2 h); pulse amplitude = 120 pA, t = 10 ms, cycle length = 1 s; recorded at room temperature.
APD_{90}	Hegyi et al., 2021 (Figs. 1, 5)	+15 ms	Mouse ventricle glucose 30 mM, 6 min	–
APD_{90}	Hegyi et al., 2021 (Figs. 1, 5)	±0 ms	Rabbit ventricle glucose 30 mM, 6 min	–
APD_{90}	Hegyi et al., 2021 (Figs. 1, 5)	60 ms	Rabbit ventricle glucose 30 mM, 6 min	I_{NaL} inhibitor GS-967 (1 µM).
APD_{90}	Hegyi et al., 2021 (Figs. 1, 5)	±0 ms	Mouse ventricle glucose 30 mM, 6 min	CaMKII inhibitor AIP (1 µM); 1 Hz steady-state pacing.
APD_{90}	L. Zhang et al., 2008	59 ms (30%)	STZ Wistar rat (n=7)	0.1–7 Hz stimulation frequencies; 37°C.
APD_{90}	L. Zhang et al., 2008	36 ms (37%)	STZ Wistar rat (n=7)	0.1–7 Hz stimulation frequencies; 37°C.

Table 2.6: CaT and active tension biomarkers in cardiac ventricular myocytes of humans and/or animals with type 2 diabetes. STZ: streptozotocin, CaT: intracellular calcium [Ca2+]i transient, Ta: active tension, D50: duration at half-maximum amplitude, TTP: time to peak, RT50: relaxation time as peak time to 50% decay, RT90: relaxation time as peak time to 90% decay. 1 mN/mm2 = 1 kPa.

Biomarker	**Reference**	**Change from baseline (approx.)**	**Experimental model**	**Experimental protocol and conditions**
CaT_{max}	Pereira et al., 2006 (Fig. 1)	0.7 F/F_0	Db/db mice (77 cells from n=4 hearts)	Fluorescence recorded with MetaZeiss LSM510 confocal microscope; water immersion 60× objective (NA 1.2); line-scan mode (1.5 ms/line).
CaT_{max}	Belke et al., 2004	50%	Db/db mice (n=9)	Solomere Technologies dual emission system; Nikon Diaphot epifluorescence microscope; recorded at 100 Hz, room temperature.
CaT, D_{50}	Pereira et al., 2006	+72.8 ms	Db/db mice (n=4)	Fluorescence recorded with MetaZeiss LSM510; water immersion 60× objective (NA 1.2); line-scan mode (1.5 ms/line).
CaT, D_{50}	Belke et al., 2004	+33%	Db/db mice (n=9)	Solomere Technologies dual emission system; Nikon Diaphot epifluorescence microscope; recorded at 100 Hz, room temperature.
T_a^{max}	L. Zhang et al., 2008	7 mN/mm^2 (41%)	STZ Wistar rat (n=8)	1.5 mmol/L $[Ca^{2+}]_o$; 37°C; 5 Hz pacing.
T_a, TTP	L. Zhang et al., 2008	+10 ms (+15%)	STZ Wistar rat (n=9)	1.5 mmol/L $[Ca^{2+}]_o$; 37°C; 5 Hz pacing.
T_a, RT_{50}	L. Zhang et al., 2008	+10 ms (+31%)	STZ Wistar rat (n=9)	1.5 mmol/L $[Ca^{2+}]_o$; 37°C; 5 Hz pacing.
T_a, RT_{90}	L. Zhang et al., 2008	+20 ms (+36%)	STZ Wistar rat (n=9)	1.5 mmol/L $[Ca^{2+}]_o$; 37°C; 5 Hz pacing.

CaMKII and dysregulation of sodium and calcium

Hyperglycemia activates the hexosamine biosynthetic pathway, further leading to post-translational modification, specifically O-GlcNAcylation, of Ca^{2+}/calmodulin-dependent protein kinase II (CaMKII). This modification causes CaMKII to exist in a chronically upregulated state in diabetes [170–172]. One major consequence is the CaMKII-dependent phosphorylation of the Na^+ channel, increasing I_{NaL} considerably. This leads to a prolongation of the action potential [173] and increases influx of sodium ions into the cell. This, in turn, can promote calcium overload of the cell through two distinct mechanisms. First, prolonged action potential duration leads to longer opening of L-type calcium channels, the main source of calcium influx. Second, higher sodium loading increases influx and reduces calcium efflux via I_{NaCa}, increasing $[Ca^{2+}]_i$, further reinforcing CaMKII hyperactivity [174].

Hyperactive CaMKII can promote early afterdepolarisations through increased phosphorylation of L-type calcium channels, which increases the amplitude of I_{CaL} [175, 176]. Early afterdepolarisations are further promoted by action potential duration prolongation following I_{NaL} hyperactivity. Hyperactive CaMKII and calcium overload further promote the formation of delayed afterdepolarisations. CaMKII-dependent phosphorylation sensitises ryanodine receptors to Ca^{2+}, increasing leak from the sarcoplasmic reticulum, which is supported by findings from mouse studies that noted an increase in calcium leakage from the sarcoplasmic reticulum in diabetic myocytes [168, 177]. The threshold of sarcoplasmic reticulum Ca^{2+} required to generate DADs is also lower in T2DM due to enhanced ryanodine receptor activation [178]. As well as Ca^{2+} leak from the sarcoplasmic reticulum, decreased SERCA2 expression and Ca^{2+} reuptake contribute to decreased Ca^{2+} transport back into the sarcoplasmic reticulum, further increasing diastolic $[Ca^{2+}]_i$ [161].

These disruptions to Ca^{2+} and Na^+ handling resemble those present in heart failure, where the sodium/calcium exchanger membrane protein (NCX) is hyperactive [174, 179, 180]. However, this is further exacerbated in T2DM by the increased activity of the sodium/glucose cotransporter, which provides another substantial source of Na^+ influx.

Data on the expression and activity levels of NCX itself in T2DM are controversial. Ashrafi et al. showed an overexpression of NCX1 (an isoform in the NCX family) mRNA

expression in T2DM human cardiomyocytes [161]. Increased NCX activity was also observed in a mouse study of T2DM [**Stlen2009IntervalCardiomyopathy**]. However, another study in diabetic rat hearts showed a decrease in protein expression and mRNA levels of NCX1, accompanied by a decrease in I_{NaCa} current density [165]. Further investigations on NCX function in T2DM are warranted to understand these differences.

PI3K/AKT signalling and ICaL

In contrast to the CaMKII-induced increase in I_{CaL}, other studies have reported a reduction in I_{CaL} in T2DM, possibly because of a reduced expression of CaV1.2, the key subunit involved in I_{CaL} [163], which may be due to impaired PI3K/AKT signalling [181]. Attenuated PI3K/AKT signalling is also responsible for increased I_{NaL} and action potential prolongation in diabetes [164].

Potassium currents and repolarisation abnormalities

Studies reported a reduction of I_{to} in rodent models of T2DM, a likely contributor of the increased action potential duration and delayed repolarisation noted in T2DM [182–186]. Regional differences in I_{to} reduction have also been proposed, with a greater attenuation in epicardial cells than in endocardial cells in T2DM [183, 187]. Hyperglycemia is thought to increase angiotensin II (AngII) in tissue, including the myocardium [188]. AngII inhibits I_{to} via downregulation of Kv4.3 and the type 2 AngII receptor [189, 190]. AngII receptor blocker valsartan reversed the reduced I_{to} observed in a mouse model of T2DM [191], supporting the involvement of AngII in I_{to} reduction in T2DM. In addition, interleukin-1-beta and tumour necrosis factor-alpha are two pro-inflammatory cytokines that are elevated in T2DM and known to inhibit I_{to} [192].

It is important to note that while I_{to} is a key repolarising current in small rodents, it plays a comparatively much smaller role in human-like species, which rely predominantly on I_{Kr} and to a lesser extent on I_{Ks} for repolarisation [193]. Thus, when changes in action potential duration are reported in small rodents, these may be driven by changes in I_{to}, and may not directly translate into similar changes in human as they are mechanistically distinct. Much less evidence is available for potassium currents most relevant for human action potential duration control. The human gene expression study by Ashrafi et al.

suggested that the ERG gene is downregulated, which causes a reduction in I_{Kr} density. This is one of the factors that may explain the QT prolongation and increased risk of early afterdepolarisation formation observed in T2DM.

2.5.5 Cardiac effects of pharmacological therapies in type 2 diabetes

As demonstrated so far, the combination of ionic, fibrotic, and nerve-related cardiac remodelling driven by T2DM makes the diabetic myocardium a high-risk substrate for cardiac arrhythmias. Pro-arrhythmic pharmacotherapy may contribute to the increased incidence of sudden cardiac death in T2DM, while other medications can also influence the heart's contractile function [194]. Here, I summarise the implications of glucose-lowering drugs and other common medication in T2DM, on arrhythmic risk and myocardial contractility.

Metformin

Metformin is the most common drug administered to T2DM patients. It helps lower blood glucose levels through the reduction of glucose production in the liver, and through the increase of glucagon-like peptide-1 secretion and glucose utilisation in the gut [195]. Experimental and clinical observational findings suggest that metformin may exert anti-arrhythmic effects, notably in the reduction of atrial fibrillation risk, while evidence on its association with ventricular arrhythmias is limited [196]. Possible mechanisms proposed include the attenuation of inflammation and oxidative stress, improvement of impaired calcium handling, and reversal of structural myocardial remodelling [196]. One study in rats reported QT interval prolongation following high doses of metformin and QT shortening at low doses [197], while a mouse study reported a shortening of the QT interval [198]. Results from Wang et al. suggest that this change occurred because of a decrease in I_{CaL} due to a decrease in expression of Cav1.2 channels and its associated CACNA1C mRNA [198]. Meanwhile, another rat study showed no change in I_{CaL} after metformin treatment, however a reduction in peak I_{to} density was found along with a prolongation of the QTc interval [199]. These results may be specific to small rodents, which rely on I_{to} for action potential duration. Furthermore, the diabetic status of the

animal models was not consistent across these studies, which may contribute to the discrepancies in these results. Overall, these experimental and clinical findings motivate the need for further studies to investigate possible anti-arrhythmic properties of metformin.

Insulin

Although patients with type 1 diabetes are fully dependent on insulin treatment, some patients with T2DM may also be prescribed insulin therapy if other medications are insufficient to control blood glucose. Patients treated with insulin were shown to be at higher risk of developing atrial fibrillation [200, 201]. Insulin can induce hypoglycemia which can in turn trigger arrhythmias, likely due to hypokalaemia that tends to accompany hypoglycaemia [202]. One study demonstrated a positive effect of insulin on NCX, with increased activity and expression, facilitating calcium removal and promoting better cardiac relaxation [203]. In the context of decreased NCX in T2DM, this effect may have beneficial effects on calcium handling in T2DM.

Sodium-glucose cotransporter-2 inhibitors

SGLT2i are commonly administered to T2DM patients to help control blood sugar levels by preventing glucose from being reabsorbed into the kidneys [204]. More recently, SGLT2i have been shown to be the only drug with significant proven benefit in patients with HFpEF, which is notoriously challenging to treat due to its aetiological heterogeneity [10, 11]. Consistent findings suggest that SGLT2i reduce the risk of ventricular arrhythmias, atrial fibrillation and sudden cardiac death [97, 205–207]. This may be in part due to the inhibition of Na^+/H^+ exchanger by SGLT2i empagliflozin, canagliflozin, and dapagliflozin, resulting in reduced cytoplasmic Na^+ and Ca^{2+} and increased mitochondrial Ca^{2+} which, in T2DM, may mitigate some of the pro-arrhythmic consequences of increased $[Na^+]i$ [208]. Myocardial expression of the SGLT1 protein is up-regulated in T2DM and HFpEF; this up-regulation increases $[Na^+]i$ which in turn increases cytosolic Ca^{2+} via NCX [209]. One study showed that the treatment of human induced pluripotent stem cells under high-glucose conditions with empagliflozin restored SGLT1 expression to control levels, thus abolishing its adverse effects on increased $[Na^+]i$ [210]. Another study showed that empagliflozin inhibits the late sodium current I_{NaL} by binding directly to

the Nav1.5 sodium channel [211], while Hegyi et al. suggested that I_{NaL} regulation by empagliflozin may occur via an indirect mechanism involving CaMKII inhibition [212]. The possible involvement of CaMKII in the mechanism of action of SGLT2i is supported by another study showing that, in a mouse model of T2DM, empagliflozin inhibited CaMKII-dependent ryanodine receptor phosphorylation and reduced O-GlcNAc levels, two key contributing factors to the pro-arrhythmic setting in T2DM [213]. In addition to their effects on the heart's cellular electrophysiology, SGLT2i have also been shown to reduce adverse structural remodelling such as fibrosis, as well as stabilise impaired myocardial energetics, which may further explain their direct benefits on the heart in terms of preventing arrhythmias and improving contraction [214–216].

Sulfonylureas

Sulfonylureas are typically the second-line T2DM treatment in cases where metformin fails [217], however they have been associated with a higher risk of atrial fibrillation, ventricular arrhythmias and sudden cardiac death compared to metformin [218–220]. Sulfonylureas stimulate insulin secretion by binding to the KATP channel in pancreatic beta cells. Their KATP channel-blocking action also applies to cardiomyocytes, leading to membrane depolarisation, subsequent influx of Ca^{2+}, action potential duration prolongation and increased QT interval in the ECG [221]. Glibenclamide, a type of sulfonylurea, was shown to block I_{Kr} and I_{Ks}, the latter to a lesser extent; this may further contribute to QT prolongation observed in patients treated with this drug [222].

Statins

Statins are commonly prescribed to patients with dyslipidemia, which is common in T2DM, to help reduce cholesterol and triglyceride levels. Studies have shown that certain statins reduce the QTc interval, including in heart failure patients treated with atorvastatin [223, 224] and in heart transplant patients treated with either pravastatin or a combination of statins [225]. One study found a significant decrease in QT dispersion in patients with T2DM treated with simvastatin, suggesting positive effects of the drug on cardiac repolarisation heterogeneity [226]. Another study based on experimental and clinical data found an association between rosuvastatin and QT prolongation [227]. These findings

align with a previous study showing the hERG-blocking effect of rosuvastatin, thus reducing I_{Kr} and prolonging cardiac repolarisation [228]. These heterogeneous results warrant careful consideration when modelling the cardiac effect of statins in T2DM, as its effects seem to vary across different compounds within this drug class.

The knowledge of certain principles easily compensates the lack of knowledge of certain facts.

— Claude Adrien Helvétius, French philosopher (18th century)

3

Methodology: Multi-Scale Modelling Approaches in Cardiovascular Medicine from Big Data to Mechanisms of Disease

This chapter lays out the core methodological concepts relevant to the work undertaken for this thesis. First, I introduce the UK Biobank, the world's largest biomedical database, which provided the data for the entire thesis. Then, I explain principles of epidemiological studies and core concepts of statistical analysis and machine learning for medical data that are later applied in Chapters 4 and 5. Finally, I present the different mathematical and computational components involved in the development of a framework for multi-scale modelling and simulation of human cardiac electrophysiology, which forms the basis of the work described in Chapter 6.

3.1 The UK Biobank

3.1.1 Study design and aims

The UK Biobank is a multi-centre, prospective cohort study of over 500,000 adults living in England, Scotland and Wales who were recruited between the ages of 40 and 69 years, from 2006 onwards. The study, of unprecedented scale, aims to enable scientific discoveries

that contribute to improving the prevention, diagnosis, and treatment of disease, notably common non-communicable diseases such as cardiovascular diseases and cancer [229].

The UK Biobank is a longitudinal study by design, its objective being to follow its participants over multiple decades to study how and why certain diseases develop. Data was collected at dedicated UK Biobank assessment centres across England, Scotland and Wales at four main visits. All participants underwent a baseline assessment during their initial visit (visit 1) at a UK Biobank centre, with approx. 20,000 participants returning for a follow-up assessment (visit 2) a few years later. The imaging sub-study described below formed a third visit (visit 3) for approx. 100,000 participants selected at random from the baseline cohort, with a repeat visit (visit 4) for approx. 5000 of the original imaging cohort, forming the fourth and final visit to UK Biobank assessment centres to-date.

3.1.2 Data

The UK Biobank is rich in breadth and depth of data, containing information on socio-demographic and lifestyle factors, anthropometric measurements, genetics, whole-organ imaging, and a variety of blood-based samples including circulating blood biomarkers, proteomics, metabolomics, and more. In this section, I describe the data components of the UK Biobank that are most relevant to my thesis. Figure 3.1 provides a visual summary of the different data and visits I considered in each of my results chapters.

Linkage to medical records

The medical history of each UK Biobank participant is accessible thanks to data linkage of the UK Biobank database with databases from the National Health Service containing participants' health records. Hospital episode statistics, containing information on hospital visits, are available for most of the cohort, while primary care records which include routine and non-emergency visits at general practice clinics are linked for about 45% of the entire study cohort (as of the time of writing of this thesis) [25]. This information is crucial as it allows to ascertain prevalent conditions that have been coded at or before a given point in time (i.e. a selected UK Biobank visit), as well as identifying cases of incident conditions (i.e. outcomes) coded after said timepoint of interest.

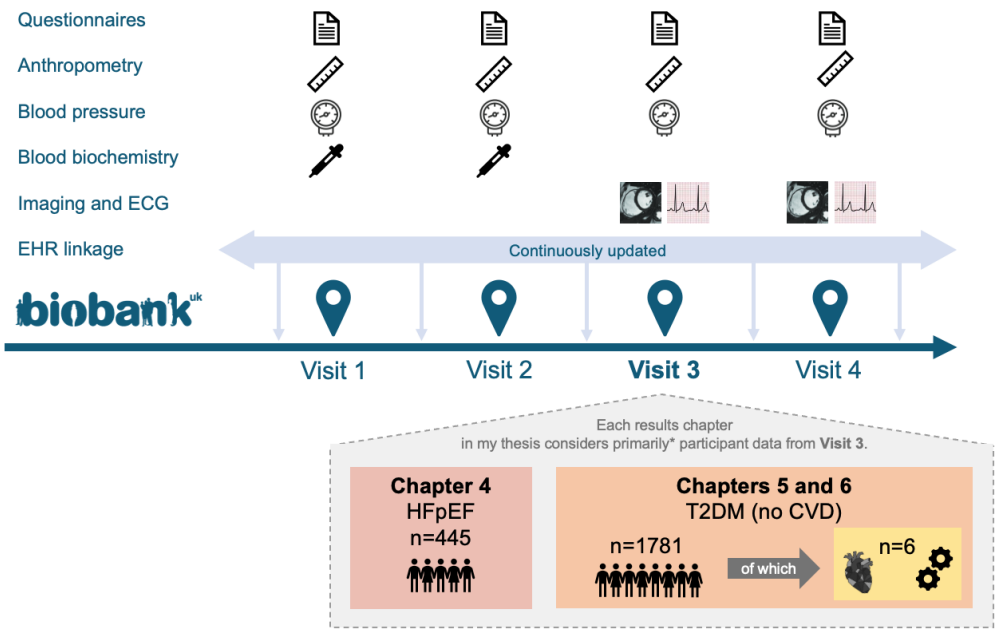

Figure 3.1: Overview of UK Biobank data types most relevant to my thesis, UK Biobank timeline and data availability, and the composition of the cohorts I derived from the UK Biobank for my thesis (primarily* from Visit 3). This asterisk refers to individuals' medical data that was originally recorded externally to UK Biobank i.e. during hospital admissions and primary care settings in the UK and subsequently linked to each participant's UK Biobank record, rather than recorded at a UK Biobank assessment centre visit. Additionally, certain blood-based biomarkers of interest were not available at Visit 3. In this case, the latest available measurement was selected from a prior visit as a proxy measurement. For Chapters where this applies, details will be specified in the corresponding Methods section. CVD: cardiovascular disease, EHR: electronic health records, HFpEF: heart failure with preserved ejection fraction, T2DM: type 2 diabetes mellitus.

Cardiac magnetic resonance imaging

The UK Biobank imaging sub-study recalled about 20% of the original study cohort for a multi-organ imaging assessment, which included a cardiac magnetic resonance (CMR) scan performed on a 1.5 Tesla scanner (MAGNETOM Aera, Syngo Platform VD13A, Siemens Healthcare). Full details of the imaging protocol are available in [230]. The raw CMR images formed the basis of a number of studies aiming to characterise cardiac and aortic structure and function. Quantitative imaging-derived biomarkers were computed and returned to the UK Biobank for future use by the UK Biobank research community. Available biomarkers are grouped into categories based on the algorithms used to extract them; these include UK Biobank category 133 (automatic extraction of 8 key biomarkers using built-in Siemens InlineVF algorithm [231]), category 157 (82

quantitative imaging phenotypes extracted using a machine learning-based image analysis pipeline developed by the authors [232]), and category 162 (72 quantitative imaging phenotypes derived from a smaller subset of participants using a collection of different algorithms for distinct biomarker families (e.g. LV, RV, aorta, strain, pericardial fat, etc.) [233–237]). More recently, 3D cardiac meshes reconstructed from the 2D CMR images of approx. 55,000 UK Biobank participants were also returned to the UK Biobank, and made accessible to the wider public in an open-source repository using sex, age and BMI ranges to avoid re-identification [238]. At the time of starting analyses for this thesis, fewer than 50,000 cases were available with a full imaging dataset, however the UK Biobank has recently reached the 100,000-scan milestone they originally set out to achieve [239], making it the world's largest multi-organ imaging study.

Electrocardiograms

A 12-lead resting ECG (CardioSoft ECG system, GE) was recorded on the same day as the CMR scan (used to gate the images) for participants in the imaging sub-study. This data provides valuable information on the heart's electrophysiological behaviour, complementing structural and functional metrics from the imaging. Key ECG biomarkers including ventricular rate, interval durations, wave amplitudes, and more, were computed automatically by the CardioSoft system and are included in different fields of the raw xml files themselves. A selection of biomarkers is also directly available in the UK Biobank externally to the raw ECG datasets, bypassing the need to access bulk data to extract these. 4-lead exercise ECGs (CAM-USB 6.5, Cardiosoft v6.51, GE) were also recorded at baseline assessment in a subset of approx. 75,000 participants, and again during the first follow-up visit in approx. 20,000 participants. While I did not consider the exercise ECG data in my thesis, I acknowledge its potential value in related future work.

Blood-based biomarkers

Blood samples were collected at all four visits and were used to characterise circulating molecules using different techniques. 30 key blood biochemistry markers (including cholesterol, C-reactive protein, different hormone markers, glucose, HbA1c, albumin, and more) were extracted for the majority of the cohort at baseline and later at the first

follow-up visit using a Beckman Coulter AU5800 automated, high-throughput clinical chemistry analyser. In my analyses, I included a selection of these biomarkers as a representation of standard clinical measurements that provide a broad characterisation of participants' molecular profiles. Additional, more detailed blood-based molecular data are available in the UK Biobank, including genetic, proteomic and metabolomic measurements. Genome-wide genetic data was obtained for every participant in the cohort using purpose-designed assays from Affymetrix (ThermoFisher Scientific). The Olink Explore 3072 platform was used to measure almost 3000 individual proteins in the plasma of over 50,000 participants. Proteins result from the transcription of DNA onto RNA which itself is used to synthesise proteins. Olink uses an affinity-based, high-throughput approach that allows for the targeted identification of a large number of known proteins including low-abundance ones using selected antibody binders, compared to mass spectrometry methods which are usually preferred for untargeted discovery analyses. Finally, nuclear magnetic resonance (NMR) spectroscopy metabolomic analysis (Nightingale Health) was performed on plasma samples of the majority of the cohort at baseline and participants from the first follow-up visit to identify levels of 251 metabolites; very light molecules derived from proteins. While I did not use these two molecular "omic" data types in my thesis, I recognise their complementary value to quantify one of the most fundamental aspects of disease progression across time: molecular function.

Additional clinical data

In my analyses, I also accounted for other types of relevant clinical data that do not pertain to the categories listed above, for example blood pressure and anthropometry i.e. body shape measurements. These can add further information complementing the blood-, imaging- and ECG-based characterisation of an individual. Together, these observable, measurable traits that characterise a person form the "phenome".

External exposures

Beyond organ-level imaging, ECG, and blood-based biomarkers, external environmental factors provide yet another layer of characterisation of a person's health or disease state. Indeed, external factors such as physical activity, nutrition, sleep, socio-economic status,

medication, smoking, air pollution, and more, can influence the transcription process, which may have downstream effects on protein and metabolite activity in the body and, subsequently, disease formation and progression. These characteristics are often referred to as collectively as the "exposome" i.e. the collection of (external) factors that an individual is exposed to. A recent UK Biobank study highlighted its importance showing that, compared to genetic-related risk, the exposome explained a greater proportion of variation in risk of certain diseases, including those of the heart [240]. In the UK Biobank, the majority of this information was self-reported and recorded through the use of questionnaires, either in person with a nurse or via a touchscreen, during the assessment centre visits.

3.1.3 Contextual relevance

Overall, the scale and prospective nature of the UK Biobank along with its breath and depth of multi-modal data described above make it an ideal resource to carry out the work proposed in this thesis. It offers a novel and rare opportunity to combine two fundamentally different yet complementary methodological approaches to investigate my proposed research questions. Indeed, it enables data-driven, machine-learning enabled, population-level studies to uncover biomarker changes in patient cohorts of interest, integrating a variety of data types as listed above to provide a comprehensive characterisation, as well as augmenting and complementing these analyses by enabling computational modelling and simulation investigations based on its rich and unique offering of cardiac imaging and ECG data. The latter makes it possible to uncover underlying pathophysiological mechanisms of disease, which in this scientific context is not possible through solely data-driven, machine learning-based approaches.

3.1.4 Ethics

General ethical approval was granted for UK Biobank studies by the United Kingdom's National Health Service Research Ethics Service (11/NW/0382). Participants provided written informed consent for their data to be stored and used for research purposes. The study presented in Chapter 4 was conducted under UK Biobank Application

Number 116292, while those in Chapters 5 and 6 were conducted under UK Biobank Application Number 40161.

3.2 Health data science: quantitative methods for Big Data in medicine

3.2.1 Evolution and aims of epidemiology

Epidemiology is considered a key quantitative discipline underpinning public health; organised community efforts aimed at the prevention of disease and promotion of health. It is defined by Miquel Porta, a pioneering epidemiologist, as "the study of the occurrence and distribution of health-related states or events, states, and processes in specified populations, including the study of the determinants influencing such processes, and the application of this knowledge to control relevant health problems" [241].

The discipline's roots stems from efforts, many centuries ago, to understand and address health patterns in rapidly growing populations facing significant public health challenges. In the United Kingdom, records of epidemiological work date back to the 1600s, with John Graunt's analysis of data from the "Bills of Mortality". This was among the first works quantifying patterns of birth, death, and disease occurrence, and included findings on disparities between males and females, high infant mortality, urban vs. rural differences, and seasonal variations in health. In the 1800s, the classical epidemiological sequence from descriptive (hypothesis generation) to analytical epidemiology (hypothesis testing) to application, as well as other surveillance and vital statistics practices still used today, were established by the pioneering work of William Farr through the systematic collection and analysis of census data and deaths by cause, and of John Snow through the comparison of cholera mortality in London between populations with different water supplies. During the early 20th century, non-infectious diseases started gathering more attention, including the work of Bradford Hill and Richard Doll linking lung cancer to smoking leading to the development of rigorous study designs and statistical methods still used today for multivariable analysis. The return of infectious diseases in the late 20th and 21st centuries such as the Acquired Immunodeficiency Syndrome (AIDS), severe

acute respiratory syndrome (SARS), avian flu and COVID-19, rekindled the importance of cohort studies, large scale randomised trials, preventive interventions and therapies.

Concurrent technological developments in the collection, storage and processing methods of large amounts of data have propelled epidemiological studies especially since the start of the 21^{th} century, including those carried out using the UK Biobank. Today, data-driven approaches play a central role in understanding disease patterns, identifying risk factors, and informing clinical care and public health decisions at scale and with unprecedented efficiency.

3.2.2 Epidemiological study design

A distinction should be made between two main types of epidemiological studies: observational, subdivided into descriptive and analytical, and interventional. This thesis considers observational epidemiology only. The aim of descriptive observational epidemiology is to generate hypotheses through the description of individuals, geographical location, timeframe of a specific disease, while analytical epidemiology aims to test these hypotheses and quantitatively explain differences in health and related events between groups by searching for causes and other contributing factors including e.g. demographic characteristics, genetic make-up, environmental exposures, and personal lifestyle.

Common analytical epidemiology study designs include the following: cohort studies, case-controlled studies, and cross-sectional studies [242]. Cohort studies considers a group of individuals over time to assess the incidence of an outcome, or multiple outcomes, based on a single exposure. Their longitudinal component, typically large sample size, and relatively small set of exclusion criteria make them a powerful and highly generalisable approach to measure risk, and establish temporality and more specifically, causality. Case-control studies compare individuals who already have the outcome of interest (cases) to those without the outcome (controls), retrospectively assessing exposure history. They do not require follow-up so tend to be rapid and cost-effective. This makes them ideal for investigating rare diseases and outbreaks, as well as diseases that take a long time to develop. Finally, cross-sectional studies can examine exposures and other known population characteristics at a single point in time,

providing a snapshot of associations within a population. This type of study design also has the advantage of not requiring long follow-up times, however it is limited in its ability to infer causality due to the lack of temporality.

3.2.3 Key epidemiological measures

There are several key measures in epidemiology. The first is prevalence to describe the current cases of a condition or disease. It is often described as a proportion of cases among the population at any given time, i.e. number of people with a disease at a given time point divided by the total number of people in the population. Incidence describes the occurrence of new cases of a disease over time. Two measures related to incidence are risk or cumulative incidence, the probability of new onset of disease among those at risk of developing the illness, as well as rate or incidence density, i.e. rate of new onset of disease observed in an at-risk population within a given period. A final measure, related to risk, is odds, defined as the ratio of number of outcome cases to no outcome, i.e. risk/(1-risk). High prevalence may reflect high incidence and/or prolonged survival without cure, while low prevalence may reflect low incidence but also possibly a rapidly fatal process or rapid recovery.

Two variables are considered to be associated if each varies by levels of the other, more than could be expected by random chance alone. That is, association explains how a particular outcome (e.g. health-related event or death) varies by a given exposure (e.g. a possible risk factor).

3.2.4 Error and bias

To avoid making false claims and overstating conclusions in epidemiological studies, before asking whether an expected observed association is causal, it is important to consider whether this association might be the result of chance, bias and/or confounding.

Random errors are a result of chance; errors are random if the average of an increasing number of observations converges to the truth. Such errors are more common in smaller studies, which tend to be imprecise, as extreme results are more probable if there are few observations but improbable if there are many. Systematic errors result in a biased

association; even with an infinite number of observations, the result will not converge to the truth. This forms the basis of bias, which is defined as any systematic error, where subjects are selected or information is measured, analysed or reported, that results in an incorrect or invalid estimate of the measure of association. There exist different forms of bias, common ones including:

- **Selection bias:** a systematic difference in characteristics between the people included in a study and those who are not and can be subdivided into two common types: participants may not be representative of the target population, or there may be a differential loss to follow-up in study participants.

- **Misclassification:** arising from an error in the classification of an exposure or the disease. Differential misclassification refers to errors on exposure or disease that are related to the other variable. Non-differential misclassification refers to errors in one variable that are unrelated to the other.

- **Information bias:** a distortion in the way that exposures are obtained, measured or reported and can occur for example due to interviewer bias, recall bias or reporting bias.

Potential sources of bias and error should be properly considered before, and acknowledged after, carrying out a study. This will allow results to be better interpreted and placed in a meaningful, realistic context. Appropriate measures can be taken during the study design and analysis steps to mitigate these biases and errors; I discuss some examples in the following section.

3.3 Statistics and machine learning

In this section, I present core theoretical concepts in statistics and machine learning that can be applied to quantitatively describe and analyse biomedical data. These form the basis of the analyses presented in Chapters 4 and 5.

3.3.1 Parametric and non-parametric statistical tests

In a study, to quantitatively compare two groups of individuals based on a specific variable, statistical tests can be used to assess differences in variable distributions across different groups. The choice of test differs depending on the type of variable, i.e. continuous or discrete/categorical, and whether the variables follow a normal distribution or not.

Continuous variables

The normality of continuous variable distributions may be assessed using the Shapiro-Wilk test. In cases where the sample size is very large (threshold depending on context and software used to run the tests but usually in the 100s/1000s), the Shapiro-Wilk test may reject the null hypothesis, i.e. suggest that the data is not normally distributed, when it is. Thus, in large samples, a significant p-value may indicate that the sample deviates slightly from perfect normality, rather than meaning it is insufficiently normal for parametric tests. Further tests, such as visual plots like Q-Q plots, and skewness and kurtosis analysis, respectively, can help assess data distributions better, beyond the p-value obtained from the Shapiro-Wilk test. A Q-Q plot measures how closely a dataset's distribution matches a theoretical distribution (most commonly the normal distribution) by plotting quantiles against each other. If data points form a straight line, the fit is good, while deviations from the line suggest non-normality or differences from the chosen distribution. Skewness and kurtosis analysis measure asymmetry and tailedness, respectively.

For comparisons of variables across two groups only, the independent samples t-test is used to compare normal distributions. The t-test is parametric in the sense that it assumes an underlying distribution. The t-test can only be performed if the sample size is $\geq$30, to preserve the validity of the Central Limit Theorem assumption which states that the sample size must be sufficiently large (and that samples must be randomly selected and independent). The non-parametric Mann-Whitney U-test is used for non-normal distributions and holds regardless of sample size. It also remains valid for normally distributed data and will not produce biased results if applied to compare such data. For comparison across three groups, normally distributed variables can be compared

using one-way analysis of variance (ANOVA) and non-normally distributed variables can be compared using the Kruksal-Wallis test.

Multiple comparison correction tends to be necessary when testing multiple variables as this increases the risk of false positives (type I error). Methods such as the Bonferroni correction control the error rate by adjusting the significance threshold according to the number of comparisons performed. It is particularly appropriate when several related pairwise tests are carried out within the same analysis framework, such as comparisons between multiple study groups for the same variable. I used Bonferroni correction in Chapter 4 for this exact purpose, i.e. to control the family-wise error rate across multiple pairwise subgroup comparisons for the same variable, ensuring that statistically significant differences remained robust despite the increased number of tests performed within this defined family of hypotheses. In contrast, Bonferroni correction can be overly conservative when applied to exploratory analyses involving a large number of distinct, and often correlated, variables, as it substantially increases the risk of false negatives (type II error). Alternative approaches, such as false discovery rate (FDR) correction (e.g. the Benjamini-Hochberg procedure), can be used to control type I error among significant results. In Chapter 5, however, analyses involved a limited set of comparisons across individual pre-specified ECG and CMR biomarkers, each representing distinct physiological parameters, and were primarily descriptive and exploratory rather than testing a single family of related hypotheses; therefore, formal multiple-comparison correction was not applied. Results were therefore interpreted cautiously as evidence of significant differences rather than definitive hypothesis tests. While controlling type II error and preserving signal was prioritised here, with cautious interpretation over formal FDR correction, I acknowledge that Benjamini-Hochberg correction could be a reasonable alternative for future work in large datasets.

Categorical variables

To compare two or more categorial variable distributions, the Chi-Squared test must be used. For comparison across three categorical groups using the Chi-Squared test, multiple comparison corrections must be applied.

3.3.2 Correlation

Correlation measures the strength and direction of a linear relationship between two variables. The Pearson correlation coefficient is given by:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}. \tag{3.1}$$

where x_i is the i-th observation of variable x, y_i is the i-th observation of variable y, $\bar{x}$ is the mean of x, $\bar{y}$ is the mean of y, and n is the number of paired observations. r varies between -1 and 1, where 1/-1 indicates a perfect positive/negative relationship whereby as one variable increases, the other increases/decreases in a perfectly linear manner. This model assumes that the relation between the two variables in question is approximately linear and that both variables are normally distributed.

3.3.3 Association and regression modelling

Linear regression

Linear regression models are a group of statistical models that are used to quantify the strength, direction, and significance of associations in data beyond random chance. Before fitting the models, key linear regression assumptions should be assessed using diagnostic tests to ensure validity. These assumptions include homoscedasticity of residuals (i.e. constant variance; tested through residuals vs. fitted values plots), normality of residuals (tested through Q-Q plots or histograms), independence of errors and absence of endogeneity (i.e. when independent variables are uncorrelated with the error term).

In simple linear regression, the relationship between one independent variable X_1 (e.g. a risk factor) and a dependent variable Y (e.g. an outcome) is modelled by:

$$Y = \beta_0 + \beta_1 X_1 + \epsilon \tag{3.2}$$

where β_0 and β_1 are the intercept and slope, respectively, indicating the change in Y per unit change in X_1. ϵ is the error of estimation, i.e. the difference between the actual Y and Y predicted by the regression model. On average, across all observations, $\epsilon = 0$. In practice, we use the linear regression equation:

$$\hat{Y} = \beta_0 + \beta_1 X_1 \tag{3.3}$$

where $\hat{Y}$ is the estimated or predicted value. The distance between the observed values and the regression line is the estimator of ϵ.

This model can be extended to incorporate multiple predictor variables X_i, to account for more complex relationships exploring the effect of several different factors on a given outcome. Thus, multiple regression extends the original model as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon \tag{3.4}$$

where $X_1, X_2, \ldots, X_n$ represent multiple predictors with respective coefficients $\beta_1, \beta_2, \ldots, \beta_n$ quantifying the impact of each predictor on Y while accounting for the effects of other variables. Indeed, this approach can be used to "adjust" a simple regression model investigating the effect of X_1 on Y, to account for other variables that may be related to Y and sometimes also X_1. These variables, referred to as confounders, can distort the true relationship between X_1 and Y. Without mathematically adjusting for these confounders, the estimated association between X_1 and Y may be biased, leading to inflated or incorrect conclusions.

In clinical studies, this model adjustment is done in progressive model strata, such that different potential confounders are accounted for sequentially. For example, if we want to model the effect of body mass index (BMI) on blood pressure (BP), we may first run an initial, baseline, non-adjusted model:

$$\text{BP} = \beta_0 + \beta_1 \cdot \text{BMI} \tag{3.5}$$

which we then adjust for sex and age, which we know are important patient characteristics that influence BP:

$$BP = \beta_0 + \beta_1\, BMI + \beta_2\, sex + \beta_3\, age \tag{3.6}$$

Finally, adjusting for average daily number of cigarettes smoked and grams of salt consumed, we obtain a "fully adjusted" model:

$$\text{BP} = \beta_0 + \beta_1 \cdot \text{BMI} + \beta_2 \cdot \text{sex} + \beta_3 \cdot \text{age} + \beta_4 \cdot \text{cigarettes} + \beta_5 \cdot \text{salt} \tag{3.7}$$

Confounders

Adjusting for confounders is essential to improve the validity of conclusions in biomedical research, ensuring that any associations observed reflect true relationships rather than spurious correlations due to irrelevant factors. Naturally the choice of confounder variables to include is an inherent study design decision, which needs to be carefully considered when setting up the study. The final choice of variables to include will depend on the research question being investigated and the data that is available.

An important distinction should be made between confounders and mediators. Confounders are associated with both exposure and outcome and require adjustment to avoid biased estimates, whereas mediators lie on the causal pathway between exposure and outcome, and adjusting for them may obscure part of the true effect of the exposure on the outcome. In my thesis, likely confounders were included in adjusted models based on covariates identified from related literature, while variables potentially on the causal pathway (e.g. BMI or BP) were examined through sequential adjustment to illustrate attenuation of associations rather than to derive fully causal estimates. Additionally, some variables may act as effect modifiers rather than confounders; context matters to determine a variable's status in a given analysis. In Chapters 4 and 5 of my thesis, I considered sex explicitly using stratified analyses to explore biologically meaningful heterogeneity, rather than just adjusting for it as a covariate.

Multicollinearity, over-adjustment and collider bias

While more complete models may be more robust, the selected predictors should not be too strongly correlated between themselves. This is known as multicollinearity and is an issue in regression modelling as it can lead to unstable, highly sensitive coefficients that are difficult to interpret; inflated errors in the coefficient estimates; and reduced predictive power of the model, despite individual predictors being seemingly correlated with the outcome as the model. To avoid this, it is sometimes necessary to remove some

of the correlated (potential) predictor variables, use regularisation methods that penalise large coefficients, or transform correlated predictors into a smaller set of uncorrelated components using dimensionality reduction techniques.

Care must also be taken to avoid adjusting for variables that may act as colliders, i.e. affected by both exposure and outcome, as this can induce spurious associations. This risk must be considered when selecting model covariates.

Finally, multivariate regression involves multiple dependent variables, enabling the simultaneous analysis of multiple outcomes and their associations with a set of predictors. The model can be represented as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{3.8}$$

where:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1n} \\ 1 & X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{m1} & X_{m2} & \dots & X_{mn} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{bmatrix}.$$

Instead of using such a multivariate framework, it is also possible to run distinct multiple linear regression models to analyse each dependent outcome independently. This is the approach I took in my thesis analyses, as the aim was to estimate the specific, independent contribution of given predictors to each individual outcome.

Linear regression models are used to predict continuous outcomes. If the dependent outcome variable is binary, logistic regression can be used instead. It uses the logistic function to map any linear combination of predictor variables to a probability value between 0 and 1 that represents the outcome. The logistic regression model is represented as:

$$\log\left[\frac{P(Y=1|X)}{1-P(Y=1|X)}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{3.9}$$

3.3.4 Cluster analysis

A regression model represents the relationship between variables to predict a specific outcome given a set of variables. In contrast, clustering is a form of unsupervised learning that models relationships within the data by grouping similar observations together, without the need for predefined outcome labels. In cases where there is no underlying "true" cluster label, it is a useful way to describe the data quantitatively.

Non-probabilistic clustering: K-means

One of the most commonly used clustering methods is K-means clustering. In a sample of n observation points $x_1, x_2, \ldots, x_n$, consider there exists K subgroups. The aim of K-means clustering is to assign each $x_1, x_2, \ldots, x_n$ to one of K clusters such that points assigned to a specific cluster tend to be more similar to each other than to those in different clusters. This similarity is quantified using the squared Euclidean distance between points:

$$d(x_i, x_{i'}) = \|x_i - x_{i'}\|_2^2 \tag{3.10}$$

This expression is used to define within-clutter scatter, which we wish to minimise, and between-cluster scatter which we wish to maximise. This is challenging as there are many possible ways to assign points to clusters. The K-means algorithm does this iteratively by picking initial cluster assignments randomly, calculating the "mean" of each cluster, and reassigning each point to the cluster whose centre it is nearest to. This process is repeated until the cluster assignment stops changing, i.e. the algorithm has converged. Thus, K-means is a deterministic i.e. non-probabilistic model; every point is assigned exactly one cluster without considering the probability of belonging to said cluster, or a possible spread across other clusters. In the context of clinical studies, this "hard assignment" nature represents an important limitation of K-means. It does not reflect the complexity of real-world settings where patients often have multiple overlapping characteristics, which here would be modelled as different clusters.

Probabilistic clustering: latent class analysis

In contrast to deterministic models such as K-means, probabilistic models such as Latent Class Analysis (LCA) can model the uncertainty in patient classifications, allowing for probabilistic assignments where each data point has probability of belonging to multiple "latent" i.e. unobserved subgroups. LCA estimates unobserved groups in a population given a set of observed indicators, assuming that the data being treated comes from an underlying model which contains hidden subgroups, to which each data point is probabilistically assigned to. In LCA, the likelihood of observing data Y_i for the ith patient is modelled as:

$$P(Y_i = y_i \mid \theta) = \sum_{k=1}^{K} P(C_i = k \mid \pi) \cdot P(Y_i = y_i \mid C_i = k, \theta_k) \tag{3.11}$$

Where $P(C_i = k \mid \pi)$ is the probability that patient i belongs to latent class k, and $P(Y_i = y_i \mid C_i = k, \theta_k)$ is the probability of observing data y_i given class k as parametrised by θ_k. LCA estimates the class membership probabilities π_k and the class-specific parameters θ_k, typically estimated using an Maximum Likelihood Estimation or Expectation-Maximisation algorithm. This approach offers a more accurate representation of clinical settings in the real world, especially for the study of conditions where patients exhibit very varied and overlapping disease phenotypes.

3.4 Multi-scale computational modelling and simulation of the human heart

While valuable for generating population-level insights, cohort studies based on data do not offer mechanistic insight into pathophysiological abnormalities and mechanisms underlying the clinical changes observed. To uncover these mechanisms of disease, experimental studies are essential. While human experimental studies are ideal for this purpose, they come with ethical challenges, particularly in balancing the risks and benefits for the patients involved as well as broader medical and societal benefits. Animal studies, often conducted in rodents, are more commonly used to explore disease mechanisms. However, the heterogeneity of animal models varying in factors such as age, animal

strain, disease induction methods, disease duration, and emergent comorbidities, represent significant limitations. Furthermore, the fact that these models often use rats and mice, which differ substantially from humans, further restricts the applicability of findings from animal experiments to human pathophysiology. Recent technological advances now allow us to augment experimental and clinical data through computational modelling and simulation of human (patho)physiology. These methods can be considered complementary to pre-clinical experimental studies, representing a reliable virtual testbed to generate and verify new hypotheses, allowing us to replicate and explain key mechanisms in healthy and diseased populations at different scales, from subcellular to whole-organ level.

Such methods have significantly progressed in cardiac electrophysiology and mechanics, with biophysically-detailed, multi-scale, human-based models enabling computational simulations of the heart from subcellular to whole organ behaviour, and its manifestation in the electrocardiogram and pressure-volume loops [26, 28, 243–245]. These novel approaches have been widely and successfully applied to different cardiac diseases such as atrial fibrillation [246, 247], myocardial ischemia and infarction [28, 248–250], hypertrophic cardiomyopathy [29, 30] as well as therapeutic approaches such as pharmacological treatment and ablation [251, 252] of atrial fibrillation, stem cell injection [253] and cardiac resynchronisation therapy [254].

3.4.1 Ion channel modelling

Electrophysiology of the cell can be assessed experimentally using single cell imaging and microelectrodes. As presented in Chapter 2, the main current fluxes in and out of a cardiac myocyte cell are driven by the movement of sodium, potassium, calcium, and other ions across cellular membranes via ion channels. Ion flux is driven by diffusional and electric field forces, thanks to the electrochemical gradient present across the channels. Each ionic species x has a characteristic equilibrium potential, which is given by the Nernst potential E_x:

$$E_x = \frac{RT}{zF} \ln\left(\frac{x_o}{x_i}\right) \tag{3.12}$$

where R is the universal gas constant (J K^{-1} mol^{-1}), T is the temperature (K), F is the Faraday constant (C/mol), z is the ion valence (dimensionless), x_o is the extracellular ionic concentration (mol/L), and x_i is the intracellular ionic concentration (mol/L). Each ionic species has its own E_x, and its respective current i_x is proportional to the difference between the transmembrane voltage V_m and E_x, that is:

$$i_x = g_x \cdot (V_m - E_x) \tag{3.13}$$

The proportionality constant is equal to x's conductance g_x, the ability of x to flow through the channel. Channel kinetic equations can represent macroscopic changes at the cellular level, where voltage-gated channels open and close following changes in V_m. The average percent of open channels, x_1, is given by:

$$\frac{dx_1}{dt} = \frac{x_{1,\infty} - x_1}{T_{x_1}} \tag{3.14}$$

where $x_{1,\infty}$ is the steady-state value of x_1, and T_{x_1} its time constant. In addition to open and closed states, a channel can also be in an activated or inactivated state. Activation and inactivation time constants can be determined from experimental data, making it possible to model different ion channel behaviours.

3.4.2 Single-cell modelling

Ion channel models derived from distinct experimental sources and distinct mathematical methods can be combined further into single-cell models, encapsulating the behaviour of multiple ionic channels at cellular scale. The Nobel prize-winning Hodgkin-Huxley model blazed the trail for cellular action potential modelling from the 1950s until today. Hodgkin and Huxley described the formation and propagation of an action potential in a squid nerve membrane, based on the ionic mechanisms of excitability and inhibition of voltage-dependent sodium and potassium channels, and fixed-conductance leak channels [255]. This work was expanded upon by Denis Noble, who developed the first mathematical model of cardiac action potentials based on ionic currents recorded using current clamp experiments in sheep and dog hearts [256]. This was a major step for mathematical modelling and cellular biology and fuelled the development of

models of many cell types including ventricular and atrial myocytes, both in animal and human. Notable human ventricular myocyte models include the O'Hara-Rudy model (ORd) [257], Grandi-Bers model [258], Ten Tusscher model [259] and ToR-ORd model [26] (Figure 3.2). The latter was developed as an extension of the ORd model with a thorough calibration and validation strategy, using multiple and distinct sources of data, including human, for both respective stages. The ToR-ORd model can simulate healthy and diseased myocytes, reproduce certain arrhythmia mechanisms such as early delayed depolarisations and alternans. It also improves the modelled response of the action potential to different drugs compared to previous similar models. Different models represent different ionic channels and mechanisms in the myocyte with varying levels of biophysical detail; the complexity of each ion channel model included within the cellular model strongly determines its context of use.

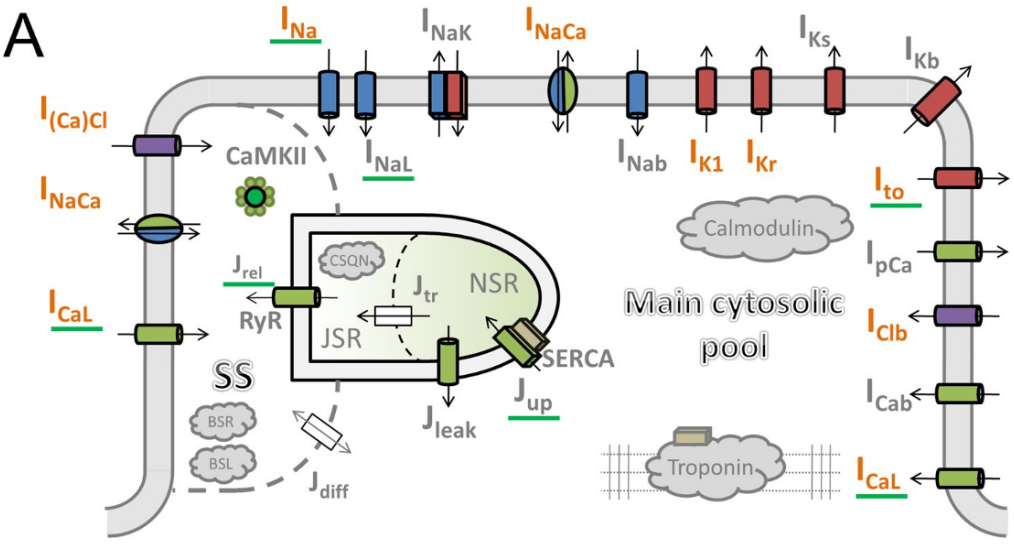

Figure 3.2: Schematic of the ToR-ORd human ventricular myocyte model for electrophysiology and calcium handling. Components underlined in green are modulated by CaMKII signalling. SS: junctional subspace compartment, SR: sarcoplasmic reticulum, JSR: junctional SR compartment, NSR: network SR compartment. Main cytosolic pool refers to the remaining intracellular space. Reproduced from [26]

The electromechanical behaviour of healthy human ventricular cardiomyocytes was simulated using the ToR-ORd model for human ventricular electrophysiology [26] coupled with the Land model for human contractility [260], as developed originally in [261]. An updated formulation of the coupled ToR-Ord-Land model was recently presented

in [262], in the context of a computational study of failing cardiomyocytes. This new model includes a revised equation for calcium buffering, allowing for a more accurate representation of calcium transient kinetics, diastolic calcium levels, and active tension under SERCA inhibition. This new model also includes two new components to enable the reproduction of enhanced calcium-activated potassium current (I_{KCa}) and junctional sarcoplasmic reticulum calcium leak (J_{leak_JSR}) in heart failure remodelling.

3.4.3 Tissue-level electrical propagation modelling

Cable theory and derived propagation models

Tissue-level cardiac electrophysiology is simulated by representing ionic flow across cells and intracellular gap junctions as a 1D chain (Figure 3.3). This process can be represented mathematically by modelling current in electric cables via a series of capacitor–resistor modules in parallel, separated by a single resistor. In this analogy, each capacitor–resistor module represents a cell membrane, while the separating resistors represent the gap junctions between individual cells.

The cable equation for this system is:

$$\frac{1}{r_l}\frac{\partial^2 V_m}{\partial x^2} = c_m\frac{\partial V_m}{\partial t} + \frac{V_m}{r_m} \tag{3.15}$$

where:

- V_m is the transmembrane potential,
- c_m is the membrane capacitance per unit length,
- r_m is the membrane resistance per unit length,
- r_l is the longitudinal resistance per unit length,
- x is the spatial coordinate along the tissue.

This representation gives rise to the bidomain model, an anatomically accurate, biophysically detailed set of equations used as a functional reaction–diffusion model of electrical conduction in cardiac tissue [264]. This concept of two coupled homogenised

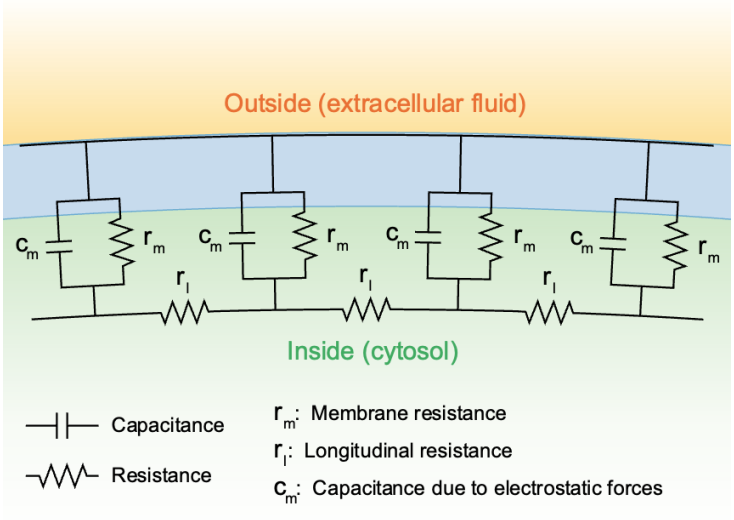

Figure 3.3: Cable theory representation of cell-to-cell electrical propagation (image adapted from [263])

domains, with one to describe intracellular space and the other to describe extracellular and interstitial space, is revised again 20 years later [265].

Building from Ohm's law assuming no external field, we can define field current densities J_i inside the cell and J_e outside the cell in terms of conductivities σ and potentials θ as follows:

$$J_i = -\sigma_i \nabla \theta_i \tag{3.16}$$

$$J_e = -\sigma_e \nabla \theta_e \tag{3.17}$$

Assuming an enclosed cardiac system, the currents in and out are equal and opposite, and equal to the product of the cellular membrane surface-to-volume ratio β and the ionic current density per unit area I_m:

$$-\nabla \cdot J_i = \nabla \cdot J_e = \beta I_m \tag{3.18}$$

Accounting for an added stimulation current I_{stim}, we can represent the current across the membrane with a cable equation as:

$$\beta I_m = \beta \left(c_m \frac{\partial V_m}{\partial t} + I_{\text{ion}} \right) - I_{\text{stim}} \tag{3.19}$$

where:

- β is the membrane surface-to-volume ratio (cm^{-1}),
- c_m is the membrane capacitance per unit area ($\mathrm{F\ cm}^{-2}$),
- $V_m = \theta_i - \theta_e$ is the transmembrane potential (V),
- I_{ion} is the ionic current density per unit area ($\mathrm{A\ cm}^{-2}$),
- I_{stim} is the externally applied stimulation current density ($\mathrm{A\ cm}^{-3}$),
- I_m is the total membrane current density ($\mathrm{A\ cm}^{-3}$).

Thus, rearranging and equating the above and replacing V_m by $\theta_i - \theta_e$, we obtain:

$$\nabla\left(\sigma_i \nabla\theta_e\right) + \nabla\left(\sigma_i \nabla V_m\right) = \beta\left(C_m \frac{\partial V_m}{\partial t} + I_{\text{ion}}\right) - I_{\text{stim}} \tag{3.20}$$

$$\nabla\left(\left(\sigma_i + \sigma_e\right) \nabla\theta_e\right) = -\nabla\left(\sigma_i \nabla V_m\right) \tag{3.21}$$

Which we can then solve for V_m and θ_e, which is feasible but highly complex and therefore computationally expensive. To circumvent this limitation, a simplification of the bidomain model, referred to as the monodomain model, operates under the assumption of an anisotropic but linear relationship between the intracellular and extracellular domains in terms of tissue conductivity. This assumption enables the relaxed model to be described by a single reaction-diffusion equation in the transmembrane potential, as opposed to one equation per space. The monodomain model was shown to be much less computationally expensive than the bidomain model whilst producing negligible differences in accuracy [266, 267].

$$\frac{\lambda}{1+\lambda}\,\mathrm{div}\left(\sigma_i \nabla V_m\right) = \beta\left(C_m \frac{\partial V_m}{\partial t} + I_{\text{ion}}\right) - I_{\text{stim}} \tag{3.22}$$

In order to solve the bi- or monodomain equations computationally, the domain of interest must be discretised in time and space. Coupled numerical solver methods have been developed specifically for the purpose of achieving this task in the context of cardiac simulations. These systems of partial differential equations can be solved using the finite

element method or finite volume method to discretise the tissue in space, while forward and backward Euler differentiation schemes enable time discretisation of the problem [268, 269]. The extent of spatial and temporal resolution of the solution depends on the size of each discrete element in space and the time step of each iteration, respectively. For example, the MonoAlg3D solver uses parallel computing with multicore and graphics processing units (GPUs) to solve the monodomain equation [270]. The solver uses the finite volume method with an implicit first-order Euler scheme and an adaptive time step, and a non-uniform mesh with finer space discretisation at the wavefront. This method accelerates the computation time and reduces memory requirements while maintaining high accuracy, compared to other algorithms [271–273].

Graph-based propagation algorithms

The expensive numerical and computational costs of these traditional solvers drive the need for more efficient yet accurate methods. To this end, graph-based approaches have been developed to speed up the computation of excitation patterns in the myocardium. Franzone et al. introduced the Eikonal equation as an efficient way of computing arrival times of depolarising wavefronts [274–276]. Derived from wave optics, the Eikonal equation computes the time taken to travel the shortest path over a given domain, given a starting location and wave velocity. The core formulation of the Eikonal equation is given below, where d is a distance vector, with the solution constrained to zero at the origin of coordinates in 3D space :

$$\begin{cases} \sqrt{\nabla^{\mathrm{T}} d \, \nabla d} = 1, \quad \Omega \subset \mathbb{R}^3, \\ d(0,0,0) = 0 \end{cases} \tag{3.23}$$

We can adapt this to describe an electrical wavefront propagating in an anisotropic cardiac mesh, given:

$$\nabla d = v \nabla t \tag{3.24}$$

Where:

$$\begin{cases} v(f,s,n) = \text{speed in fibre, sheet, normal directions} \\ t(x,y,z) = \text{time travelling through point (x,y,z)} \end{cases} \tag{3.25}$$

Thus obtaining the new formulation:

$$\begin{cases} \sqrt{\nabla t^{\mathrm{T}}.V.\nabla t} = 1, \quad \Omega \subset \mathbb{R}^3, \\ d(0,0,0) = 0 \end{cases} \tag{3.26}$$

Where:

$$V = \begin{pmatrix} x_l & x_t & x_n \\ y_l & y_t & y_n \\ z_l & z_t & z_n \end{pmatrix} \begin{pmatrix} v_f^2 & 0 & 0 \\ 0 & v_s^2 & 0 \\ 0 & 0 & v_n^2 \end{pmatrix} \begin{pmatrix} x_l & x_t & x_n \\ y_l & y_t & y_n \\ z_l & z_t & z_n \end{pmatrix}^T \tag{3.27}$$

The general solution to the Eikonal equation is the Euclidean distance:

$$d(x,y,z) = \sqrt{x^2 + y^2 + z^2} \tag{3.28}$$

As current can propagate from node a to b through edge (a,b) with Euclidean distance $\left\| (\vec{a,b}) \right\|_2$, the time cost between two adjacent nodes becomes:

$$c_{(a,b)} = \sqrt{(\vec{a,b})^T V (\vec{a,b})} \tag{3.29}$$

Finding the cardiac activation sequence can thus be represented as a shortest path-finding problem on a manifold, mesh or graph that represents the cardiac anatomy, and using the location of the electrical stimulus as a starting point for the wave to propagate. The Eikonal activation times can be solved using a multi-source, multi-destination extension of Dijkstra's algorithm; other typical methods to compute the shortest path on a discrete domain include the Fast Marching Method, Fast Sweeping Method and Fast Iterative Method. These methods rely on iterative updates whereby all nodes must first be searched before being updated, resulting in an algorithm that is difficult to parallelise, thus inefficient. The Eikonal equation can also be solved using either of these methods with varying degrees of efficiency, and has been successfully applied to rapid cardiac simulation studies [277–281]. The Eikonal method estimates cardiac activation sequences by representing the anatomy as a connected graph consisting of nodes and edges. Spatial

nodes represent points in the myocardium, while edge weights correspond to the wavefront travel time along that edge. The connected graph represents both surfaces and intra-wall myocardial tissue, and the activation time at each node is obtained using the discrete A* shortest-path algorithm. This method circumvents the need for all nodes to be searched at each iteration, providing a more efficient solution yielding a prediction 10^4 to 10^5 faster than the bidomain method using a tessellated or tetrahedral mesh, and 10-100 times faster than the Fast Marching Method using a graph of the same resolution. Plank et al. used the Eikonal equation to adapt the original bidomain model by replacing its diffusion component by an Eikonal solver to speed up the simulation [282]. Here, instead of using a differential equation to represent diffusion, the propagation of the depolarising wave in cardiac tissue is represented using the Eikonal operating on a discrete domain with relaxed spatial constraints, yielding rapid high-fidelity activation patterns that determine the behaviour of depolarising wavefronts.

In addition to improvements in performance, fast methods for activation sequence prediction also contribute significantly to cardiac parameter inference by producing hundreds and thousands of simulations in a relatively short period of time, thus enabling rapid model personalisation and uncertainty quantification [279, 280, 283]. Other advantages of graph-based approaches also include the ability to study tissue behaviour at different scales, as shown in [284]. Here, the authors use graph-based homogenisation to compute activation times in fibrotic tissue, by modifying the coarse anatomical mesh in order to capture fine electrophysiological dynamics using the Eikonal, thus representing micro-scale tissue properties across a macro-scale mesh.

Electrocardiogram simulation and the pseudo-ECG

To link the outputs of the models described above with clinical data, it is crucial for these models to be able to generate realistic simulated electrocardiograms (ECG), considering its clinical significance and widespread use in the assessment of cardiac function.

The ECG can be computed by approximating the extracellular potential at a single point in space using the transmembrane potential from the monodomain model as the bioelectric source [285], and by using the infinite volume conductor assumption that

considers the heart as immersed in a spatially infinite medium with of homogeneous conductivity [286]. However, this approach is of limited accuracy. The bidomain model, which considers current flow explicitly in both intra- and extracellular domains, yields a more accurate reconstructed ECG, but its computational cost is prohibitively high [287]. To simulate the bidomain, two key equations need solving, namely a parabolic equation for the transmembrane voltage and an elliptic equation for the extracellular potential. The latter is particularly computationally expensive as it involves a large sparse system of linear equations across the entire domain that need to be solved at every timestep. To overcome this, Bishop and Plank developed a computationally efficient pseudo-bidomain method that combines an augmented monodomain simulation with "infrequent" solves (i.e. not at every timestep) of the elliptic component of the bidomain equations. This approach achieves a close match to the full bidomain model in terms of activation patterns, extracellular potential distributions, and ECG traces, while offering a substantial reduction in computational cost [288]. This pseudo-bidomain ECG reconstruction method is the one implemented in the MonoAlg3D solver, which we will refer to throughout the rest of this thesis. For a 1D fibre source, the voltage at the sensor point (x',y',z') is given by:

$$\phi(x', y', z') = \frac{\alpha^2 \sigma_i}{4\sigma_e} \int \left[(-\nabla V_m) \cdot \nabla \left(\frac{1}{r} \right) \right] dx\, dy\, dz \tag{3.30}$$

Where α is a constant that depends on fibre radius, integration occurs across ventricular activation period, ∇V_m is the spatial gradient of transmembrane potential V_m, $\nabla \left(\frac{1}{r} \right)$ is the Euclidean distance from point (x, y, z) to the electrode location. In the Eikonal method, V_m is either 1 or 0 depending on whether that node has been activated.

3.4.4 Organ-level modelling

Whole-organ reconstruction

Based on whole-organ images acquired using high-resolution clinical imaging modalities such as cardiac magnetic resonance or computed tomography, it is possible to create a virtual and physiologically accurate representation of the patient's cardiac anatomy as a three-dimensional mesh. These reconstruction pipelines typically rely on convolutional neural networks and/or statistical shape models to segment the images, delineate the

myocardium, and obtain a smooth surface representing the heart in 3D [289–294]. These surfaces can be transformed further into tetrahedral meshes, including rule-based cardiac fibre orientation, recapitulating experimental findings [295, 296]. These patient-specific virtual cardiac geometries form the anatomical graph or manifold basis of whole-organ 3D cardiac simulations, as described in Section 3.4.3.

Fibrosis

Advanced imaging techniques like late gadolinium enhanced CMR and T1 mapping are essential for characterising myocardial fibrosis non-invasively in type 2 diabetes mellitus (T2DM) [297, 298], with T1 mapping-derived extracellular volume fraction serving as a proxy marker of diffuse fibrosis [299], the most common type of fibrosis found in T2DM (see Chapter 2, Section 2.5.2). O'Hara et al. used fused late gadolinium enhanced T1 CMR data from human hypertrophic cardiomyopathy patients and threshold-based rules to obtain a personalised reconstruction of fibrosis in a virtual heart built from the same imaging data [300]. Another computational study investigated the pro-arrhythmic effects of adipose and fibrotic tissue in the heart in 2D using the TP06 model of a ventricular cardiomyocyte [301]. Further mathematical concepts such as Perlin noise have been applied to model fibrosis in 3D, generating physiologically realistic, virtual representations of fibrosis at tissue with automatic calibration based on histological image slices [302]. Coleman et al. incorporated the effects of fibrosis on electrical conduction in a virtual whole-organ study by modelling delayed electrical wave propagation by stochastically assigning certain elements in the cardiac mesh as unexcitable using a probability density function [30].

It is far more important to know what person the disease has than what disease the person has.

— Hippocrates, the Father of Medicine (500BC)

4

Elucidating Cardiometabolic Profiles and Variation in Left Ventricular Function in Heart Failure with Preserved Ejection Fraction

Building on the clinical context outlined in Chapter 2, this chapter investigates heterogeneity within cardiometabolic heart failure with preserved ejection fraction (HFpEF) in the UK Biobank by harnessing statistical methods described in Chapter 3. Accounting for sex-specific differences, I characterise cardiometabolic profiles and ventricular function in an elderly UK Biobank cohort using latent class analysis, allowing for a more granular and continuous grouping of disease. This analysis revealed three distinct, clinically relevant phenogroups, namely a mostly male and multimorbid group; a group with a high prevalence of severe obesity and the highest relative proportion of females; and a group with an apparently lower comorbidity burden aside from hypertension. I found a loss of expected sex differences in myocardial contractility (through quantification of global longitudinal strain) in the highly comorbid vs. most hypertensive subgroups; worse contractility in women in the highly comorbid vs. obese subgroup; and an elevated cardiac output in the obese subgroup consistent with "high output HFpEF". Contractility and cardiac output reflect sex- and comorbidity-dependent pathophysiological context,

motivating the need for context-aware interpretation of cardiac functional metrics in HFpEF in future personalised management strategies.

This chapter is based on the following publication:

- **Bertrand A.**, Zhou X., Lewis A., Monfeuga T., Gupta R., Grau V., Rodriguez B. Sex-specific cardiometabolic multimorbidity, metabolic syndrome and left ventricular function in heart failure with preserved ejection fraction in the UK Biobank. Cardiovasc Diabetol 24, 238, 2025. DOI:10.1186/s12933-025-02788-4

4.1 Introduction

As introduced in Chapter 2 (Section 2.4), chronic inflammation and metabolic disturbances underlying hypertension, diabetes and obesity are overlapping drivers of the pathogenesis of HFpEF [303–305]. Sex and age also play an important role in cardiometabolic HFpEF. Compared to men, women have stiffer ventricles and arteries, smaller vasculature, and greater pulse pressure, placing them at a higher risk of developing adverse microvascular changes, diastolic dysfunction and eventually HFpEF [12–15]. In addition, ageing processes differ between sexes, making the progression of cardiometabolic decline and cardiac dysfunction highly age- and sex-dependent [16, 17]. Metabolic syndrome also worsens in women during the perimenopausal stage [306, 307]. However, the limited number of sex-specific studies of HFpEF and poor representation of women in clinical trials motivate the need for further sex-specific investigations of HFpEF, especially including post-menopausal women [308–310].

Considering the multi-factorial nature of the disease, HFpEF treatment is typically adapted on an etiology-specific basis to complement first-line treatment with SGLT2 inhibitors [9, 38]. As seen earlier, following seminal work by Shah et al. [77], HFpEF phenotyping studies have elucidated common clinical traits in different clusters of HFpEF [78, 86, 88, 311–318]. These studies, their findings and shortcomings, and persisting clinical challenges were reviewed in Chapter 2, Section 2.4.6. Despite their relevance in HFpEF pathophysiology, cardiometabolic diseases and components of metabolic syndrome are not always consistently and extensively covered in these studies, and

results are not always stratified by sex despite its important role in HFpEF. Cardiac function and hemodynamic parameters beyond left ventricular ejection fraction (LVEF), notably global longitudinal strain, which has been associated with poorer outcomes in HFpEF [319, 320], are not always consistently accounted for. There is limited data on the variation of cardiac function across different HFpEF phenotypes including across sexes [321, 322]. Overall, the need to improve the identification and refinement of clinically distinct phenogroups of HFpEF is well-acknowledged, especially subgroups pertaining to the cardio-metabolic-renal axis, in order to improve the classification and management of HFpEF [8, 19, 323, 324].

To address these gaps, this study presents a high-resolution mapping of the heterogeneities in cardiometabolic HFpEF in an elderly cohort of UK Biobank participants, ultimately aiming to help inform disease management strategies in this high-risk population. I hypothesise that cardiometabolic profiles differ in men and women with HFpEF, and that patients belonging to distinct subgroups may exhibit different extents of ventricular dysfunction. I elucidate these questions by examining the cardiometabolic comorbidity burden and metabolic syndrome in the selected cohort; extracting probabilistic phenogroups from the cohort using latent class analysis based on cardiometabolic diseases and metabolic syndrome components; and comparing whole-organ imaging-derived parameters of left ventricular function across the different cardiometabolic phenogroups identified. These analyses are stratified by sex to reflect underlying differences in men and women.

4.2 Methods

4.2.1 Study design and population

The study sample was drawn from the UK Biobank, which I presented in Chapter 3. Here, I performed a retrospective cohort study with hybrid prevalence/incidence analysis, with a modified cross-sectional component considering prevalent and incident cases at the time of the imaging assessment. I examined the relations between cardiometabolic comorbidity burden, metabolic syndrome components and left ventricular function in patients with

HFpEF, stratified by sex. As a subgroup analysis, I compared the cardiometabolic profile of patients with prevalent vs. incident HFpEF at time of imaging, to support the design rationale and mitigate any selection bias that could arise from the inclusion of both subgroups in this study. Results are reported in line with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement. Ethical approval for the study was obtained as outlined in Chapter 3, Section 3.1.4.

4.2.2 Cohort definition

The study sample I defined for this study consists of UK Biobank participants with HFpEF, including cases of prevalent HFpEF diagnosed before the time of imaging (mean time since diagnosis: 5.7 years) and incident HFpEF diagnosed after the imaging assessment (mean follow-up time: 3.3 years). The rationale behind this hybrid approach was to capture HFpEF exhaustively considering that the age of participants in the UK Biobank is typically lower at assessment compared to the typical onset of HFpEF as evidenced by low baseline prevalence, coupled with the fact that HFpEF is commonly underdiagnosed, especially when co-existing with cardiometabolic conditions [325, 326].

To ascertain HFpEF, I first identified all UK Biobank cases with a heart failure diagnosis by retaining participants with a non-null date of first report of an ICD-10 code I50, which codes for a heart failure diagnosis (UK Biobank field number 131354). This field number 131354 refers to "Date I50 first reported (heart failure)" and forms part of UK Biobank Category 1712 "First occurrences", a collection of fields compiled internally by UK Biobank teams. "First occurrences" fields contain the earliest known diagnosis date of a report of a given ICD-10 code (or its mapped equivalent) across combined data sources from hospital episode statistics, primary care records, death registers, and self-reported medical conditions. To further classify heart failure based on LVEF, I retained participants with CMR image-derived LVEF values greater or equal to 50% (fields 31060, 22420, 24103). The selection process and cohort sizes are summarised in Figure 4.1.

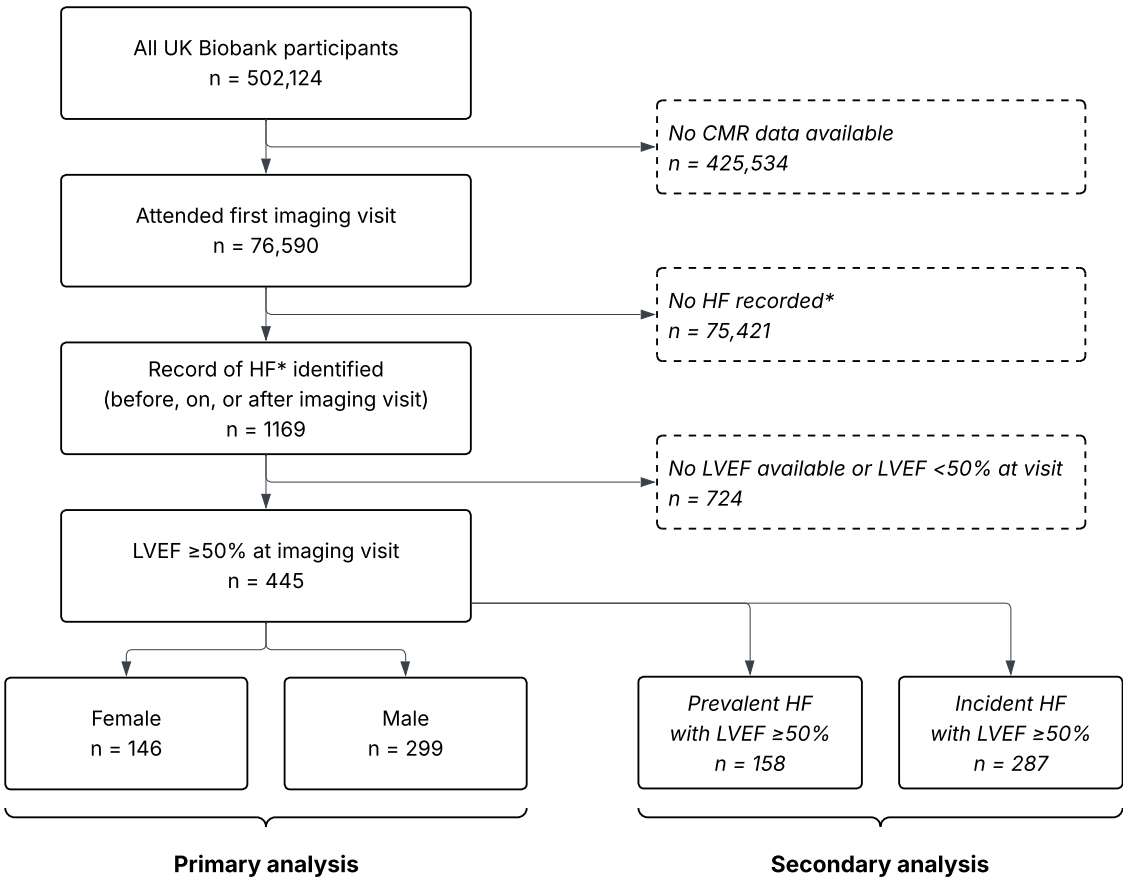

Figure 4.1: Cohort selection flowchart. The asterisk refers to the ascertainment method for HF, which I carried out using UK Biobank field number 131354 "Date I50 first reported (heart failure)", representing the earliest known diagnosis date of a report of an ICD-10 code I50 or equivalent, combining data from hospital episode statistics, primary care records, death registers, and self-reported medical conditions (see UK Biobank Category 1712). CMR: cardiac magnetic resonance, HF: heart failure, LVEF: left ventricular ejection fraction.

4.2.3 Baseline covariates

I considered the following demographic covariates: sex (UK Biobank field 31), ethnic background (21000), age as calculated between date of birth (33) and date at which the imaging data was recorded (53), as well as the following socio-economic variables: qualifications (6138), employment status (6142) and access to private healthcare (4674). The latter was selected as an individual-level proxy for healthcare access and health-seeking behaviour. Area-level indices such as the Townsend Deprivation Index may capture broader socioeconomic deprivation and could be considered in future studies. I also considered the following lifestyle and anthropomorphic covariates: body mass index (BMI) calculated using height (50) and weight (21002), waist-to-hip ratio calculated

using waist circumference (48) and hip circumference (49), smoking status (20116), and alcohol consumption (1558). I included the following clinical covariates: diastolic blood pressure (DBP) (4079), systolic blood pressure (SBP) (4080), triglycerides (30870), high-density lipoprotein (HDL) cholesterol (30760), blood glucose (30740) and cholesterol-lowering medication, blood pressure medication and insulin (6153, 6177). Certain blood biochemistry measurements were not obtained concurrently with the imaging data; in these cases, I used the most recent available measurement from a prior visit as a proxy. To account for possible existing cardiac structural and functional abnormalities at the time of imaging, I included known cardiomyopathies (131338, 131340) and history of cancer, including lung and breast cancer (20001), as detailed information on cardiotoxic treatments or chest radiotherapy was not available. For each covariate, I used the most recent measurement obtained prior to or on the date of the imaging assessment.

4.2.4 Outcomes

I considered two key study outcomes: cardiometabolic burden, consisting of cardiometabolic disease diagnoses and metabolic syndrome components, and CMR imaging-derived parameters of left ventricular function, in the selected cohort subjects.

Definition of cardiometabolic disease

I defined cardiometabolic diseases as ischemic heart disease, hypertension, diabetes, obesity, chronic kidney disease, liver disease (primarily metabolic dysfunction-associated fatty liver disease (MASLD)) and stroke. I used ICD-10 codes to identify each disease category (see Appendix A, Table A.1). I further ascertained hypertension and obesity based on clinical biomarker thresholds, as established by recent clinical guidelines, using data collected at the assessment centre. This step helps minimise the number of cases that are missed due to incomplete medical records or a lack of linkage of primary care records to the UK Biobank. Individuals with a systolic SBP ≥140 mmHg and/or DBP ≥90 mmHg were classified as hypertensive [58]. Those with BMI ≥30 kg/m^2, or BMI ≥25 kg/m^2 coupled with waist-to-height ratio (WtHR) ≥0.5, were classified as obese [327]. BMI thresholds were applied uniformly across the cohort. While optimal BMI cut-offs

may vary by ethnicity, the application of uniform thresholds in the context of the UK Biobank is acceptable given that participants, including the cohort defined for this study, is predominantly of white ethnic background. I acknowledge that while clinical symptoms of obesity are becoming increasingly important in its diagnosis [327], symptoms were not considered in the ascertainment of obesity for the purposes of this study.

Definition of metabolic syndrome

I identified abnormal metabolic syndrome components based on the following criteria: waist circumference ≥102cm in men or ≥89cm in women, blood pressure ≥130/85 mmHg, fasting triglyceride level ≥150 mg/dl, fasting HDL cholesterol level ≤40 mg/dl in men or ≤50 mg/dl in women, and fasting blood glucose ≥100 mg/dl [328]. In addition to reporting abnormalities in individual components, I reported cases of clinically defined metabolic syndrome using the National Cholesterol Education Program's Adult Treatment Panel (NCEP ATP) III definition, defined as the presence of three or more of these criteria [328].

Imaging-derived parameters of left ventricular function

I included the following imaging parameters, measured at rest: LVEF (UK Biobank fields 22420, 24103, 31060), LV end-diastolic volume (22421, 24100, 31061), LV stroke volume (22423, 24102, 31064), cardiac output (22424, 24104), and LV global longitudinal strain (22181). These parameters were computed from the raw imaging data using automated software and in-house algorithms that were developed in previous studies [232, 233, 235]. Details of the verification, validation, and quality control measures are available in respective study publications. Parameters with multiple UK Biobank field numbers indicate the measurements acquired in each of the referenced studies. For cases with multiple reported measurements of the same parameter, I calculated the average value of all the available parameter measurements.

4.2.5 Statistical analysis

All statistical analyses were performed using Python version 3.8 and R version 4.1.1.

Statistical tests

Tests were applied following the criteria described in Chapter 3, Section 3.3.1. For comparisons of n>2 groups, post-hoc pairwise p-values between groups were calculated using the Tukey Honest Significant Difference test after analysis of variance (ANOVA), or Dunn's test for the Kruksal-Wallis test, and further adjusted using Bonferroni correction to adjust those p-values by multiplying each pairwise p-value by the number of comparisons. I applied Bonferroni correction here to control the family-wise error rate across the multiple pairwise tests performed for the same variable, ensuring that any significant differences observed among subgroups remained robust to the increased number of comparisons within this single analysis framework.

All statistical tests were two-sided. Statistical significance was defined as $p \leq 0.05$. Median and interquartile range are reported for continuous variables, while count and percentage are reported for categorical variables.

Missing data

Missing count and percentage of the full sample are reported for all baseline variables as part of the cohort description. For the outcome analysis of imaging-derived parameters, I removed outliers with a left ventricular end-diastolic volume ≥500 ml to exclude extreme values that likely represent measurement error. I also acknowledge that very low ventricular volumes, potentially arising from incomplete short-axis coverage in the CMR scanning, were not systematically excluded and may affect measurement validity.

For imaging parameters that were missing completely at random, I used median imputation. I acknowledge that this reduces variance and may bias associations towards the null. More advanced alternative methods were not considered here due to the relatively low proportion of missingness and the scope of the study. To mitigate bias and ensure the robustness of these results, I repeated the same outcome analysis after discarding cases that did not have a full set of imaging parameters.

Clustering analysis

HFpEF subgroups were established using latent class analysis (LCA), which was previously described in Chapter 3, Section 3.3.4. Here, I selected LCA rather than heuristic clustering, as probabilistic clustering approaches tend to reflect reality more closely both in terms of probabilistic class belonging and the existence of an underlying distribution, therefore making results more applicable to other cohorts.

The indicator variables considered in my analysis include binary variables representing the presence or absence of individual cardiometabolic diseases as defined earlier; the presence or absence of abnormal individual metabolic syndrome components; and male or female sex. I chose to perform LCA on the full cohort with sex as an indicator variable, instead of performing two distinct LCAs in the male and female subgroups, to avoid further subgrouping into smaller clusters that may be less representative. I calculated pairwise Phi coefficients to measure the association between all possible pairs of these binary variables, to avoid including highly associated variables in the model. To select the optimal number of classes, consecutive models were fit with an increasing number of classes. The following metrics were computed: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the bootstrap likelihood ratio test. AIC favours model fit and complexity, capturing more complex patterns, while BIC penalises complexity and is less prone to overfitting. The aim is to minimise both, and select a model that fits the data well while being neither too simple nor too complex. The LR test compares the fit of two models with a different number of classes by evaluating whether adding an additional class significantly improves the model fit, thus providing another possible layer of validation to AIC and BIC. LCA was implemented in R using the poLCA package. I ran models with a varying number of possible classes ranging between 1 and 7, with 5000 maximum iterations and 20 random starts to ensure reproducibility [329]. The LR test had 50 bootstrap replications. poLCA uses the expectation-maximisation algorithm, which can handle missing data under the assumption that the data is missing at random and estimates the model parameters based on the available data.

Regression

A linear regression line was fit and the R^2 coefficient calculated to quantify the relationship between individual metabolic syndrome components and parameters of left ventricular function.

4.2.6 Subgroup and sensitivity analyses

To mitigate possible bias introduced by the consideration of patients with prevalent and incident HFpEF relative to the time of imaging, I carried out a subgroup analysis comparing the cardiometabolic profile of those two subgroups to identify possible differences between them, which may affect the way results are interpreted. To achieve this, I repeated the statistical tests carried out to compare cardiometabolic comorbidities, abnormal metabolic syndrome components and clinically-defined metabolic syndrome in patients with prevalent vs incident HFpEF at the time of assessment. As described above, I also compared the results of outcome analyses with and without discarding cases with missing imaging biomarker data.

4.3 Results

4.3.1 Cohort description

The study cohort consisted of 445 participants with HFpEF. The cohort was mostly male (n = 229, 67%), elderly (median age of 70 years), of white ethnicity (n = 432, 97.1%), overweight with a BMI ≥25 kg/m^2 (median BMI 27.9 kg/m^2), regular or occasional alcohol drinkers (n = 326, 73%), current or previous smokers (n = 226, 51%), on cholesterol lowering medication (n = 246, 55.3%) and on blood pressure medication (n = 245, 55.1%). Blood pressure IQRs were close to or the same as elevated blood pressure ranges, i.e. SBP of 120-139 mmHg or DBP of 70-89 mmHg, as defined by the most recent European Society of Cardiology guidelines [58]. These trends were conserved in both sexes, though a higher proportion of males compared to females consumed alcohol more frequently, were current or previous smokers, and took either insulin, blood pressure medication or cholesterol-lowering medication. However, there was a higher proportion

of women with a history of cancer compared to men (25.3% vs. 16.7%, p<0.05). Full cohort characteristics are reported in Table 4.1.

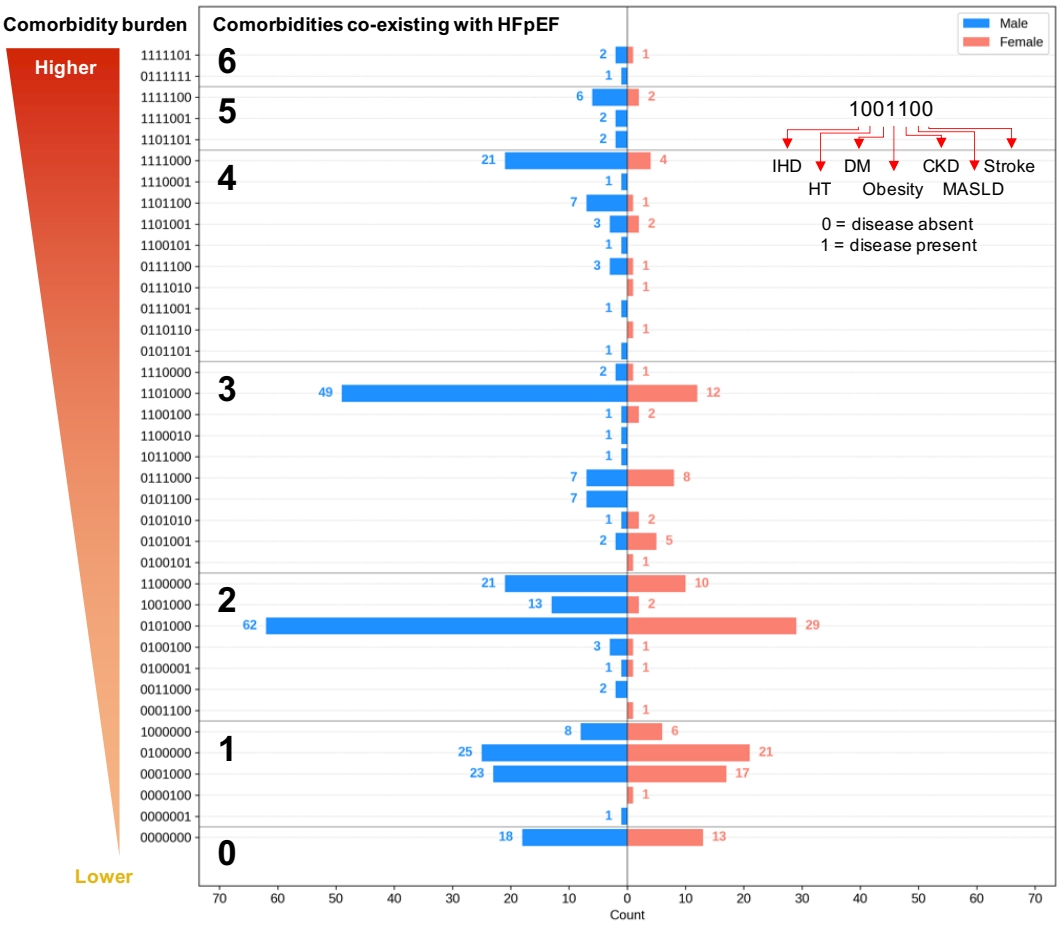

Figure 4.2: Cardiometabolic comorbidity burden in male (n=299) and female (n=146) HFpEF patients. Labels on the y-axis are coded with binary digits indicating the presence (1) or absence (0) of cardiometabolic comorbidities in the following order: ischemic heart disease, hypertension, diabetes, obesity, chronic kidney disease, MASLD, stroke. HFpEF: heart failure with preserved ejection fraction, IHD: ischemic heart disease, HT: hypertension, DM: diabetes mellitus, CKD: chronic kidney disease, MASLD: metabolic dysfunction-associated fatty liver disease.

Table 4.1: Demographic, socio-economic, lifestyle, and clinical characteristics of participants in the study sample, stratified by sex. Statistical tests were used to compare characteristics between male and female subgroups. All continuous variables were compared using the Mann-Whitney U-test aside from waist-to-hip ratio and SBP which were assessed using the independent t-test. Categorical variables were compared using the Chi-square test. The p-values reported refer to the individual results of those respective tests. For sex-specific columns, percentage for each count is reported as the percent of that sex subgroup. HFpEF: heart failure with preserved ejection fraction; IQR: inter-quartile range; BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; HDL-C: high-density lipoprotein cholesterol.

	All		**Males**		**Females**		**p-value**
	Count (%) or Median (Q1-Q3)	**Missing (%)**	**Count (%) or Median (Q1-Q3)**	**Missing (%)**	**Count (%) or Median (Q1-Q3)**	**Missing (%)**	
Demographics							
Age, years	70.0 (66.0-74.0)	0 (0.0%)	71.0 (66.0-74.0)	0 (0.0%)	70.0 (65.2-73.0)	0 (0.0%)	0.09
Ethnicity	-	0 (0.0%)	-	0 (0.0%)	-	0 (0.0%)	0.16
White	432 (97.1%)	-	289 (96.7%)	-	143 (97.9%)	-	-
Black	1 (0.2%)	-	0 (0.0%)	-	1 (0.7%)	-	-
Asian	6 (1.3%)	-	6 (2.0%)	-	0 (0.0%)	-	-
Mixed	1 (0.7%)	-	0 (0.0%)	-	1 (0.7%)	-	-
Other	5 (1.1%)	-	4 (1.3%)	-	1 (0.7%)	-	-
Socio-economic background							
Have college or university degree	66 (14.8%)	0 (0.0%)	45 (15.1%)	0 (0.0%)	21 (14.4%)	0 (0.0%)	0.97
Employment status	-	0 (0.0%)	-	0 (0.0%)	-	0 (0.0%)	0.64
Employed	73 (16.4%)	-	52 (17.4%)	-	21 (14.4%)	-	-
Retired	329 (73.9%)	-	217 (72.6%)	-	112 (76.7%)	-	-
Use private healthcare	72 (16.2%)	0 (0.0%)	48 (16.1%)	0 (0.0%)	24 (16.4%)	0 (0.0%)	1.00
Lifestyle characteristics							
BMI, kg/m^2	27.9 (24.9-32.0)	15 (3.4%)	28.4 (25.1-31.6)	10 (3.3%)	27.0 (24.4-32.1)	5 (3.4%)	0.09
Waist circumference, cm	95.0 (85.4-105.0)	14 (3.1%)	98.0 (90.0-107.0)	10 (3.3%)	87.0 (78.0-98.0)	4 (2.7%)	<0.001
Waist-to-hip ratio	0.9 (0.9-1.0)	14 (3.1%)	1.0 (0.9-1.0)	10 (3.3%)	0.8 (0.8-0.9)	4 (2.7%)	<0.001
Alcohol	-	0 (0.0%)	-	0 (0.0%)	-	0 (0.0%)	<0.001
Never	30 (6.7%)	-	16 (5.4%)	-	14 (9.6%)	-	-
$\leq 3 times a month$	89 (20.0%)	-	46 (15.4%)	-	43 (29.5%)	-	-
1-2 times weekly	99 (22.2%)	-	59 (19.7%)	-	40 (27.4%)	-	-
3-4 times weekly	123 (27.6%)	-	96 (32.1%)	-	27 (18.5%)	-	-
Daily	104 (23.4%)	-	82 (27.4%)	-	22 (15.1%)	-	-
Smoking	-	9 (2.0%)	-	5 (1.7%)	-	4 (2.7%)	0.001
Never	210 (47.2%)	-	124 (41.5%)	-	86 (58.9%)	-	-
Previous	208 (46.7%)	-	158 (52.8%)	-	50 (34.2%)	-	-
Current	18 (4.0%)	-	12 (4.0%)	-	6 (4.1%)	-	-
Unknown	9 (2.0%)	-	5 (1.7%)	-	4 (2.7%)	-	-
Clinical measurements							
SBP, mm/Hg	142.0 (130.0-155.1)	69 (15.5%)	142.5 (130.0-157.0)	48 (16.1%)	140.0 (130.5-151.5)	21 (14.4%)	0.18
DBP, mm/Hg	76.0 (70.0-84.0)	69 (15.5%)	76.5 (71.0-84.0)	48 (16.1%)	75.0 (68.0-83.5)	21 (14.4%)	0.17
HDL-C, mmol/L	1.3 (1.1-1.5)	54 (12.1%)	1.2 (1.0-1.4)	34 (11.4%)	1.5 (1.2-1.8)	20 (13.7%)	<0.001
Blood glucose, mmol/L	5.0 (4.7-5.5)	54 (12.1%)	5.0 (4.7-5.5)	34 (11.4%)	5.0 (4.7-5.4)	20 (13.7%)	0.35
Triglycerides, mmol/L	1.6 (1.1-2.3)	23 (5.2%)	1.7 (1.2-2.4)	11 (3.7%)	1.4 (1.0-1.9)	12 (8.2%)	<0.001
Medication							
Cholesterol lowering medication	246 (55.3%)	0 (0.0%)	188 (62.9%)	0 (0.0%)	58 (39.7%)	0 (0.0%)	<0.001
Blood pressure medication	245 (55.1%)	0 (0.0%)	179 (59.9%)	0 (0.0%)	66 (45.2%)	0 (0.0%)	0.005
Insulin	13 (2.9%)	0 (0.0%)	9 (3.0%)	0 (0.0%)	4 (2.7%)	0 (0.0%)	1.00
Medical history							
Cardiomyopathy	14 (3.1%)	0 (0.0%)	9 (3.0%)	0 (0.0%)	5 (3.4%)	0 (0.0%)	1.00
Cancer (all types)	87 (19.6%)	0 (0.0%)	50 (16.7%)	0 (0.0%)	37 (25.3%)	0 (0.0%)	0.04
Lung cancer	2 (0.4%)	0 (0.0%)	1 (0.3%)	0 (0.0%)	1 (0.7%)	0 (0.0%)	1.00
Breast cancer	16 (3.6%)	0 (0.0%)	-	-	16 (11.0%)	0 (0.0%)	-

Table 4.2: Cardiometabolic comorbidity burden and metabolic syndrome in the study sample, stratified by sex. The Chi-square test was used to compare categorical variables between male and female subgroups. The p-values reported refer to the individual results of those respective tests. For sex-specific columns, percentage for each count is reported as the percent of that sex subgroup. The asterisk refers to the number of additional comorbidities refers to the number of cases with *n* cardiometabolic diseases as defined in main text, additionally to HFpEF, and excluding MetS which I consider separately. HFpEF: heart failure with preserved ejection fraction, F: female, M: male, MetS: metabolic syndrome.

	All (n = 445)	**Males (n = 299)**	**Females (n = 146)**	**p-value**
MetS components and clinically-defined syndrome				
Waist circumference >102 cm (M) or >89 cm (F)	180 (40.4%)	113 (37.8%)	67 (45.9%)	0.13
Blood pressure >130/85 mmHg	335 (75.3%)	226 (75.6%)	109 (74.7%)	0.92
Triglycerides >150 mg/dL	197 (44.3%)	148 (49.5%)	49 (33.6%)	0.002
HDL-C <40 mg/dL (M) or <50 mg/dL (F)	120 (27.0%)	81 (27.1%)	39 (26.7%)	1.00
Glucose >100 mg/dL	86 (19.3%)	63 (21.1%)	23 (15.8%)	0.23
Clinically-defined MetS (≥ 3 components)	146 (32.8%)	103 (34.4%)	43 (29.5%)	0.34
Cardiometabolic comorbidities				
Ischemic heart disease	184 (41.3%)	141 (47.2%)	43 (29.5%)	0.001
Hypertension	339 (76.2%)	233 (77.9%)	106 (72.6%)	0.26
Diabetes	68 (15.3%)	49 (16.4%)	19 (13.0%)	0.43
Obesity	304 (68.3%)	216 (72.2%)	88 (60.3%)	0.02
CKD	46 (10.3%)	34 (11.4%)	12 (8.2%)	0.39
MASLD	7 (1.6%)	3 (1.0%)	4 (2.7%)	0.33
Stroke	28 (6.3%)	18 (6.0%)	10 (6.8%)	0.90
Number of additional comorbidities*				0.05
0 (HFpEF only)	31 (7.0%)	18 (6.0%)	13 (8.9%)	-
1	102 (22.9%)	57 (19.1%)	45 (30.8%)	-
2	146 (32.8%)	102 (34.1%)	44 (30.1%)	-
3	102 (22.9%)	71 (23.7%)	31 (21.2%)	-
4	48 (10.8%)	38 (12.7%)	10 (6.8%)	-
5	12 (2.7%)	10 (3.3%)	2 (1.4%)	-
6	4 (0.9%)	3 (1.0%)	1 (0.7%)	-

4.3.2 Multimorbid cardiometabolic disease patterns differ in men and women with HFpEF

The sex-specific distribution of cardiometabolic diseases and metabolic syndrome co-existing with HFpEF in the study cohort is reported in Table 4.2, with the exact comorbidity burden in men and women presented in Figure 4.2. In this cohort, 281 of 299 (94%) men and 133 of 146 (91%) women were found to have at least one other cardiometabolic disease co-existing with HFpEF, and 224 of 299 (75%) men and 88 of 146 (60%) women had two or more cardiometabolic comorbidities. I also found that the most likely number of comorbidities co-existing with HFpEF was differently distributed between men and women (p=0.05), with the majority of men having 2 (n=102, 34.1%) or 3 (n=71, 23.7%) co-morbidities while most women had 1 (n=45, 30.8%) or 2 (n=44, 30.1%).

The most common specific comorbidity pattern was hypertension and obesity combined, both in men (n=62, 21%) and women (n=29, 20%). Other comorbidity patterns differed in men vs. women (Figure 4.2). I found a significant difference in the distribution of obesity and ischemic heart disease cases in men vs. women (n=141 (47.2%) vs n=43 (29.5%), p=0.001, and n=216 (72.2%) vs 88 (60.3%), p=0.02, respectively).

4.3.3 Latent class analysis reveals three distinct clusters based on cardiometabolic diseases and metabolic syndrome components in HFpEF

No pair-wise associations with a Phi coefficient >0.4 were observed so I retained all candidate indicator variables for model optimisation. The optimal number of latent classes based on the given indicator variables was found to be 3, based on agreement between AIC and BIC. The results of the LR test were not significant but I have included test results for completeness and transparency. See Appendix A, Figures A.1 and A.2 for more detail.

Of the 445 patients in the cohort, 117 (26%) were assigned to Group 1, 116 (26%) to group 2 and 212 (48%) to group 3. The breakdown of class-conditional probability estimates of each variable used to perform LCA is given in Figure 4.3. Group 1 was characterised by the highest probability of a patient having a record of singular or co-morbid ischemic heart disease, hypertension, diabetes, chronic kidney disease (CKD) and stroke,

as well as the highest likelihood of any abnormal metabolic syndrome components, aside from waist circumference. Group 2 was characterised by the highest probability of a patient having obesity and an abnormally high waist circumference. Group 3 was characterised by absolute or relatively low probability of a patient having any cardiometabolic disease or abnormal metabolic syndrome components, aside from hypertension and high BP which were comparable to Group 2. Group 1 had the highest likelihood of a patient being male (84%). In proportion, there were fewer men in Groups 2 (55%) and 3 (65%) compared to the number of men in the entire cohort (299 of 445 patients, 67%).

I found several significant differences in socio-demographic, lifestyle, and clinical characteristics across the three subgroups (Table 4.3). Group 1 was found to be the most elderly subgroup, also with the highest median SBP, blood glucose and triglycerides, lowest HDL-C levels, and highest proportion of medicated patients including cholesterol lowering medication, blood pressure medication and insulin. Patients in Group 2 had the highest BMI and waist circumference, which is expected considering the obese profile of the group. Finally, patients in Group 3 had the highest proportion of non-smokers, both previous and current, highest HDL-C and lowest triglyceride levels.

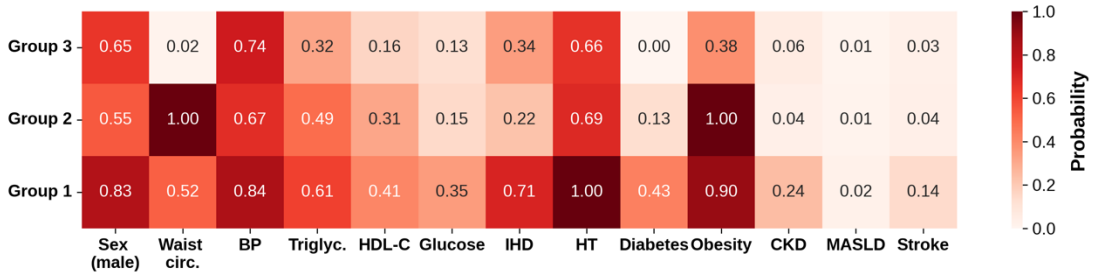

Figure 4.3: Latent class analysis results: estimated class-conditional probability of each variable being positive in respective classes. By a positive variable response, I refer here to sex being male, individual MetS components being beyond their respective normal clinical thresholds, and individual cardiometabolic diseases being present. Group 1 n = 117, group 2 n = 116, group 3 n = 212. MetS: metabolic syndrome, circ.: circumference, BP: blood pressure, triglyc.: triglycerides, HDL-C: high-density lipoprotein cholesterol, IHD: ischemic heart disease, HT: hypertension, DM: diabetes mellitus, CKD: chronic kidney disease, MASLD: metabolic dysfunction-associated fatty liver disease.

Table 4.3: Latent class analysis results: socio-demographic, lifestyle, and clinical characteristics compared across the different clusters identified. Statistical tests were used to compare characteristics across the three groups. Adjusted p-values after Bonferroni correction are reported for comparisons between groups 1 vs. 2, 2 vs. 3, and 1 vs. 3, respectively. Percentage for each count is calculated relative to each respective subgroup. HFpEF: heart failure with preserved ejection fraction; IQR: inter-quartile range; BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; HDL-C: high-density lipoprotein cholesterol.

	Group 1 n = 117 Median (Q1–Q3) or Count (%)	**Group 2** n = 116 Median (Q1–Q3) or Count (%)	**Group 3** n = 212 Median (Q1–Q3) or Count (%)	**p-value** (raw)	**p-value** (adjusted)
Socio-demographics					
Sex, male	102 (87.2%)	60 (51.7%)	137 (64.6%)	<0.001	<0.001
Age, years	71.0 (68.0-75.0)	69.0 (64.0-72.0)	71.0 (66.0-74.0)	0.002	0.183, 0.001, 0.095
Ethnicity	-	-	-	0.58	1.0
White	112 (95.7%)	115 (99.1%)	205 (96.7%)	-	-
Black	1 (0.9%)	0 (0.0%)	0 (0.0%)	-	-
Asian	2 (1.7%)	0 (0.0%)	4 (1.9%)	-	-
Mixed	0 (0.0%)	0 (0.0%)	1 (0.5%)	-	-
Other	2 (1.7%)	1 (0.9%)	2 (0.9%)	-	-
Lifestyle characteristics					
BMI, kg/m^2	30.4 (27.5-33.6)	32.7 (29.6-35.8)	24.9 (23.2-26.9)	<0.001	<0.001, 0.001, <0.001
Waist circumference, cm	101.0 (95.0-110.0)	105.0 (98.8-114.0)	85.0 (80.0-92.0)	<0.001	<0.001, 0.028, <0.001
Waist to hip ratio	1.0 (0.9-1.0)	0.9 (0.9-1.0)	0.9 (0.8-0.9)	<0.001	<0.001, 0.113, <0.001
Alcohol	-	-	-	0.21	0.63
Never	10 (8.5%)	6 (5.2%)	14 (6.6%)	-	-
≤3 times a month	19 (16.2%)	29 (25.0%)	41 (19.3%)	-	-
1-2 times weekly	28 (23.9%)	33 (28.4%)	38 (17.9%)	-	-
3-4 times weekly	33 (28.2%)	24 (20.7%)	66 (31.1%)	-	-
Daily	27 (23.1%)	24 (20.7%)	53 (25.0%)	-	-
Smoking	-	-	-	0.02	0.05
Never	45 (39.1%)	55 (47.8%)	110 (53.4%)	-	-
Previous	67 (58.3%)	51 (44.3%)	90 (43.7%)	-	-
Current	3 (2.6%)	9 (7.8%)	6 (2.9%)	-	-
Clinical measurements					
SBP, mmHg	150.2 (138.6-160.0)	138.8 (125.8-153.6)	139.0 (129.0-152.5)	<0.001	<0.001, 0.001, 1.0
DBP, mmHg	76.5 (70.5-84.0)	79.5 (70.0-86.6)	75.0 (69.6-82.0)	0.02	1.0, 0.68, 0.05
HDL-C, mmol/L	1.1 (1.0-1.3)	1.3 (1.0-1.5)	1.4 (1.2-1.7)	<0.001	<0.001, 0.001, 0.039
Blood glucose, mmol/L	5.4 (4.8-5.9)	4.9 (4.7-5.3)	4.9 (4.6-5.3)	<0.001	<0.001, <0.001, 1.0
Triglycerides, mmol/L	2.0 (1.4-2.5)	1.8 (1.4-2.3)	1.4 (1.0-1.9)	<0.001	<0.001, 0.69, <0.001
Medication					
Cholesterol lowering medication	96 (82.1%)	52 (44.8%)	98 (46.2%)	<0.001	<0.001
Blood pressure medication	94 (80.3%)	60 (51.7%)	91 (42.9%)	<0.001	<0.001
Insulin	12 (10.3%)	1 (0.9%)	0 (0.0%)	<0.001	<0.001

4.3.4 Left ventricular function across identified HFpEF cardiometabolic phenogroups

I compared LVEF, global longitudinal strain, LVEDV, LV stroke volume, and cardiac output across the three groups identified by LCA in the HFpEF cohort in the previous stage. Three outlier cases (LVEDV ≥500 ml) were removed and 30 cases had one or more missing parameter values, which I imputed. Results of this analysis are visualised in Figure 4.4. Results of the same analysis, repeated after excluding cases without a full set of imaging parameters instead of imputation, are displayed in Appendix A Figure A.3. Both analyses yielded very similar results.

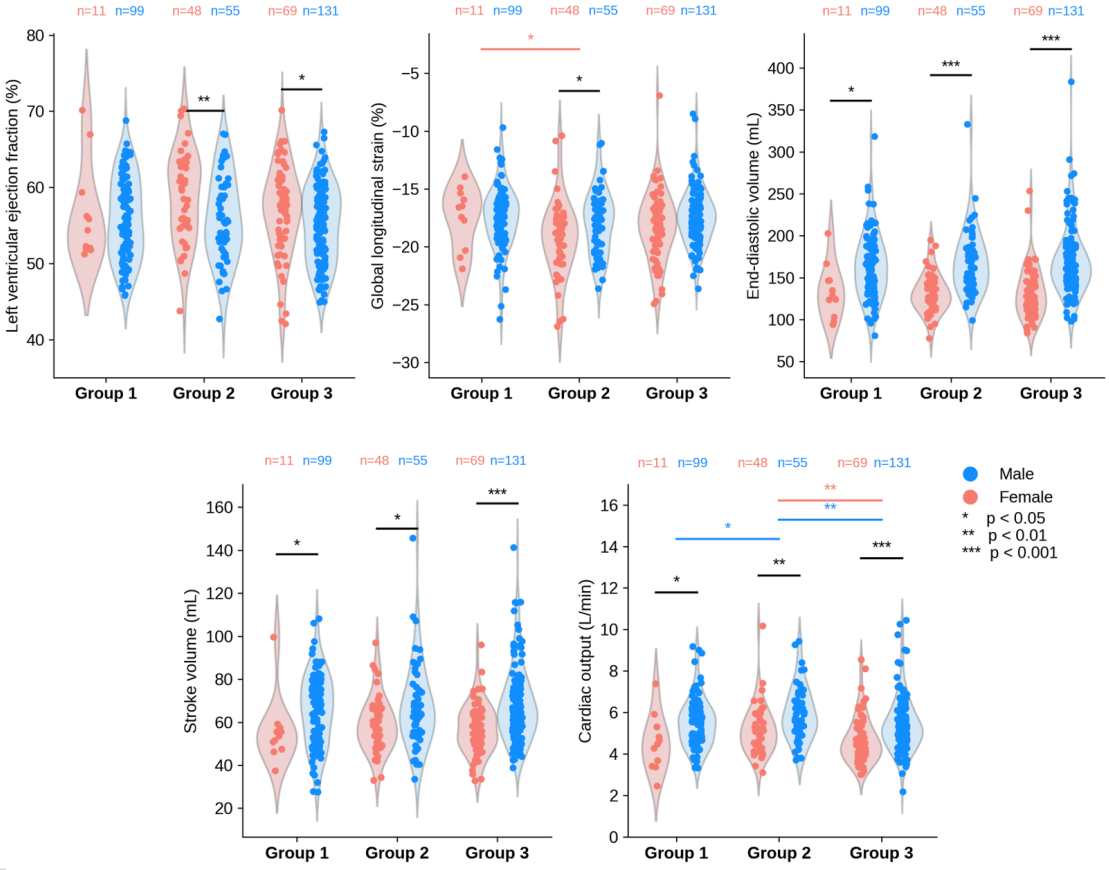

Figure 4.4: Distribution of CMR image-derived left ventricular function parameters in the HFpEF cardiometabolic phenogroups identified. CMR: cardiac magnetic resonance; HFpEF: heart failure with preserved ejection fraction.

Cardiac output was significantly higher in men and women in Group 2 vs. Group 3 (males: median 5.6 L/min vs. 5.2 L/min, p<0.05; females: 5.1 L/min vs. 4.4 L/min, p<0.01). Absolute global longitudinal strain was significantly lower, i.e. contractility

was worse, in women in Group 1 vs. Group 2 (-17.6% vs. -18.5%, p<0.05) (Figure 4.4). There were no statistically significant differences in LVEF, LVEDV or stroke volume across the three groups.

Significant sex-specific differences within each group were observed for LVEDV, stroke volume and cardiac output. These parameters in men were consistently higher than in women, consistent with known underlying sex-dependent anatomical differences in cardiac chamber volumes. Men and women had significantly different distributions of LVEF in Group 2, with higher LVEF in women (59% vs. 56%, p<0.01). Global longitudinal strain differed significantly within Group 2, with lower absolute values observed in men (-17.7% vs. -18.5%, p<0.05), suggesting that contractility in males with HFpEF is similar to females with HFpEF, except from Group 2 where it is worse in males (Figure 4.4).

Finally, I did not observe any statistically strong relationships between individual metabolic syndrome components i.e. waist circumference, blood pressure, triglycerides, HDL cholesterol and blood glucose, with LVEF, global longitudinal strain, LVEDV, stroke volume, and cardiac output in the study cohort (Figure 4.5).

4.3.5 Subgroup analysis: prevalent vs incident HFpEF

I compared the cardiometabolic profile of patients with prevalent and incident HFpEF at the time of imaging (Appendix A, Table A.2). There were very little differences between the two, with the only significant differences found between those two subgroups being a higher proportion of patients in the incident HFpEF subgroup, compared to prevalent HFpEF, with BP levels considered abnormally high in metabolic syndrome terms (n = 227 (79.1%) vs n = 108 (68.4%)), and more CKD diagnoses (n = 37 (12.9%) vs n = 9 (5.7%)). These findings suggest that patients with prevalent HFpEF at the time of imaging and incident HFpEF (mean follow-up of 3.3 years), have a similar cardiometabolic burden at the time of imaging. This strongly supports the rationale of considering these jointly for the purposes of this study and its design. However, these results must still be interpreted within the context and limitations of this study and the wider UK Biobank, which I discuss in the following section.

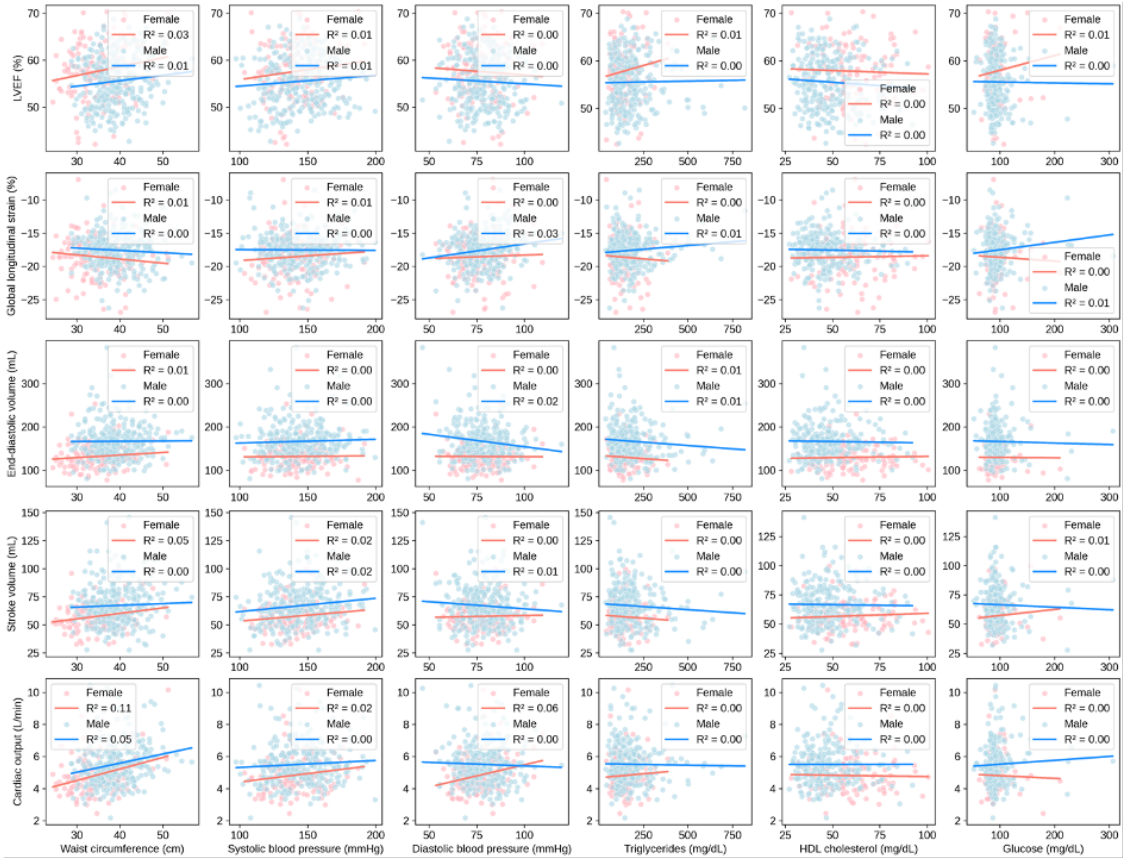

Figure 4.5: HFpEF: heart failure with preserved ejection fraction; LVEF: left ventricular ejection fraction; HDL: high-density lipoprotein.

4.4 Discussion

This Chapter provides an in-depth analysis of cardiometabolic HFpEF by examining the cardiometabolic comorbidity burden and metabolic syndrome in a retrospective cohort of elderly men and women with HFpEF, and assessing the relation between different cardiometabolic profiles and cardiac function and hemodynamics. I found that the combination of hypertension and obesity was the most common comorbidity pattern both in men and women with HFpEF, but that the other most common comorbidity patterns in men and women were different. Most men had 2 to 3 clinical cardiometabolic comorbidities while most women had 1 to 2, despite a similar profile in terms of abnormal metabolic syndrome components.

Given patients' cardiometabolic diseases and metabolic syndrome profile, I identified three distinct clusters of HFpEF patients, namely one with the highest likelihood of being

male, elderly, and being multimorbid with an abnormal metabolic profile (n= 117); a group with a high prevalence of severe obesity and abnormal waist circumference with the lowest relative proportion of females (n=116); and a final group with an apparently lower comorbidity burden aside from hypertension, generally "healthier" clinical measurements and the highest proportion of non-smokers (n=212). Certain characteristics observed in these three clusters overlap with some of the key clinical HFpEF phenotypes that have been proposed previously, namely complex multimorbid HFpEF [323, 324], obese HFpEF [330], and hypertensive HFpEF [5]. Importantly, however, these phenotypical cluster characteristics are not always mutually exclusive, hence the importance of probabilistic grouping presented in this study, and in more practical terms, a holistic consideration of different clinical characteristics of HFpEF when assessing individual patients.

I found that women had higher LVEF compared to men in the subgroup with high prevalence of severe obesity. Global longitudinal strain in women was similar to men except in the highly obese subgroup, where men had worse contractility. Cardiac output was significantly higher in the more obese group compared to the least comorbid group, and global longitudinal strain was lower, i.e. contractility was worse, in women in the highly comorbid subgroup compared to the more obese subgroup. I discuss these findings and their clinical implications in further detail in this section.

4.4.1 Challenges in management and treatment of multimorbid HFpEF

Cardiometabolic HFpEF is considered the most prevalent and highest-risk form of HFpEF [88, 304]. The results from this chapter highlight the existence of a broad variety of highly multimorbid, abnormal metabolic profiles within the HFpEF population studied. These systemic disturbances are likely to drive the poorer quality of life, symptoms, and higher non-cardiovascular death rates associated with HFpEF [6, 331].

SGLT2 inhibitors are one of the only existing therapy with proven benefit in HFpEF patients; a benefit that is due to lower heart failure hospitalisation rates, but with no significant reduction of cardiovascular mortality nor clinically meaningful improvements in quality-of-life scores [5, 10, 11]. Findings from my study strengthen the motivation for

a personalised, phenotype-specific approach to optimally manage HFpEF. Furthermore, patients with obesity and HFpEF have been shown to have an exacerbated symptom burden, be more prone to exercise intolerance, and have higher rates of hospitalisation for heart failure [40–42]. Semaglutide, a glucagon-like peptide-1 (GLP1) receptor agonist, was recently shown to improve symptoms and physical limitations of patients with obesity and HFpEF in the STEP-HFpEF trial, with a further link between body weight reduction and magnitude of treatment effect [65, 67]. These findings strengthen the notion of a central obesity-inflammatory HFpEF phenotype driven by dysfunctional visceral adipose tissue [43] (see Chapter 2 for more details). This evidence further supports the importance of identifying and addressing modifiable risk factors such as obesity to control disease progression and improve adverse outcomes in HFpEF.

Overall, the deep clinical characterisation of the cardiometabolic HFpEF phenotype presented in this Chapter supports our understanding of the variety of patient profiles that exist within cardiometabolic HFpEF, including obesity, and thus contributing to a better identification of different patient subgroups eligible for optimal therapeutic strategies.

4.4.2 Sex-specific differences in UK Biobank HFpEF rates and cardiometabolic profiles

A study by Fry et al. presented evidence of a "healthy volunteer" bias among the UK Biobank study population, with the prevalence of self-reported conditions being lower than that of the general UK population as reported in national health surveys [332]. One particular finding suggests that this bias is even more pronounced for cases of cardiovascular disease reported in women compared to men: only 7.4% of women aged 45 to 64 in the UK Biobank were reported as having cardiovascular disease, as opposed to 25.5% in the general population, while rates for men in the same age range were 16.1% in the UK Biobank vs 29.4% in the general population. The "healthy volunteer" bias explained by Fry et al., exacerbated in sex-specific reports of cardiovascular disease, could possibly explain the unexpected sex-specific distribution of HFpEF cases identified in the UK Biobank cohort I defined for my study, despite the subset of participants with imaging data, regardless of heart failure status, consisting of 52% women. However, I still found

that women with HFpEF were more likely to have a lower diagnosed comorbidity burden compared to men, which aligns with previously published studies [333, 334].

4.4.3 HFpEF, diabetes and chronic kidney disease

The pathophysiological pathways involved in the cardio-renal-metabolic axis all contribute to HFpEF and the adverse progression of connected systemic diseases, including diabetes and CKD [5, 323]. A recent meta-analysis of large-scale HFmrEF (heart failure with mildly reduced EF) and HFpEF clinical trials highlighted the high prevalence of comorbid diabetes and CKD in those cohorts, with rates as high as 35-40% in four major trials. I found comparatively low rates of diabetes (n = 68, 15.3%) and CKD (n = 46, 10.3%). One plausible explanation for this discrepancy is that HFpEF is commonly underdiagnosed in patients with type 2 diabetes and CKD in the general population [20, 335, 336]. Indeed, challenges in HFpEF diagnosis may be responsible for unrecognised HFpEF in UK Biobank participants with diabetes or CKD, and indeed vice versa. Also, the lack of complete linkage of the UK Biobank to primary care records (as of early 2025) could be responsible for missing cases of diabetes and CKD, as both diseases are more likely to be identified in primary care settings [25]. Thus, although this specific finding is controversial relative to existing literature, it may be simply a reflection of the study sample and dataset used, rather than reflecting de facto epidemiological differences in the underlying disease population.

4.4.4 Clinical relevance of global longitudinal strain and cardiac output in HFpEF

Impaired global longitudinal strain (GLS), i.e. a lower absolute value, is common in HFpEF. It is associated with reduced contractility and diastolic dysfunction, is a strong predictor of adverse events and disease progression, and is the most predictive of sudden cardiac death in patients with HFpEF in the TOPCAT study [319, 320, 337]. Studies have shown that obese patients and those with higher BMI have impaired GLS [338], and that metabolic syndrome is also associated with impaired GLS, with central obesity being the most strongly associated component [339]. I found that women in the most highly comorbid

subgroup had a significantly lower GLS absolute value compared to those in the subgroup with highest rates of obesity and abnormal waist circumference. Women in the obese subgroup had a higher absolute GLS value compared to men in the same subgroup, which aligns with previous studies showing that GLS tends to be higher in women than men [340, 341] and that male sex is an independent predictor of GLS decline [342]. This expected sex-specific difference, however, was not observed in subgroup 1 with high comorbidity burden nor in subgroup 3 with a high likelihood of hypertension. These heterogenous results highlight the importance of considering sex and existing comorbidities beyond obesity as important factors in the interpretation of GLS in the context of HFpEF.

A reduced cardiac output, notably during exercise, is one of the hallmarks of HFpEF [343, 344]. At rest, cardiac output tends to be normal in HFpEF, but variations may occur due to underlying or comorbid conditions [345]. I found that cardiac output measured non-invasively at rest was higher Group 2 (obese and abnormal waist circumference) compared to Group 3 (apparently "healthier" but hypertensive). This finding may be reflective of high-output heart failure in Group 2, a form of heart failure that typically results from the body's increased demand for blood resulting from increased metabolic activity and/or excessive vasodilatation, which are both common in obesity [346]. A previous study also showed that increasing severity of obesity in HFpEF patients was associated with a higher cardiac output at rest, which further supports my findings [347].

4.4.5 Novelty compared to existing HFpEF phenotyping studies

Unlike previous HFpEF studies, the clustering approach I used is based exclusively on sex and cardiometabolic disturbances, the main pathophysiological drivers of HFpEF. Previous studies approach clustering from a much wider angle, typically based a broader set of variables, aside from a few exceptions. One study clustered HFpEF patients based on >300 blood-based biomarkers [78], while another clustered based on non-cardiac conditions, frailty and nutrition indicators in the elderly [312]. I considered CMR imaging-derived parameters beyond the LVEF as study outcomes, to compare cardiac health in the different phenogroups identified. A couple of other studies also considered imaging parameters as outcomes, however neither included global longitudinal strain

despite its clinical relevance in HFpEF [78, 88]. In contrast, other studies considered these parameters as clustering variables rather than outcomes [77, 311, 313–316, 348], or did not include them in their analyses [86, 312, 317, 318]. Finally, by reflects the inherently continuous nature of physiology, my choice of LCA and probabilistic grouping in this study provides a major advantage over k-means or hierarchical clustering approaches previously used [78, 313, 314, 318].

4.4.6 Strengths, limitations, and future directions

A key strength of this study comes from the richness and quality of UK Biobank data, which was recorded following a robustly validated and systematic data recording protocol, helping to minimise data collection bias. The multi-centre nature of the UK Biobank allows for better geographical coverage and diversity of the study population at a national level. The depth of data available enabled an extensive characterisation of each case in my study sample.

The constraints imposed by the research question here limited my cohort to include only participants with CMR data, a record of heart failure, and an EF ≥50% to identify individuals who are most likely to have HFpEF i.e. those with a diagnosis of non-specific heart failure at or before imaging in combination with LVEF ≥50% at the imaging visit. This definition may have resulted in misclassification of certain cases, as some individuals may have had prior HFrEF with subsequent recovery of LVEF (HFrecEF) following treatment and were therefore classified as HFpEF at the time of imaging; treatment effects were not accounted for here and may have influenced phenotype classification. The relatively high prevalence of ischemic heart disease within the HFpEF cohort I identified may also point toward HFrecEF rather than "true" HFpEF. Alternative validation approaches, such as the examination of treatment and/or relevant procedure codes such as percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG), assessment of regional wall motion abnormalities on CMR, or stratified analyses according to the presence or absence of coronary artery disease. These analyses were beyond the scope of this thesis but could strengthen confidence in the case definition in future related work.

Due to the timing of this study in the lifespan of the UK Biobank, the strict ascertainment criteria led to a relatively small HFpEF cohort (n=445). This also meant that it was not possible to carry out a robust survival analysis considering incident adverse events such as mortality or hospitalisation outcomes. Nevertheless, I hope that this study may still pave the way for larger and possibly longitudinal, comparative studies of HFpEF in the UK Biobank in the future, as more outcome cases arise. Incomplete linkage to primary care data in the UK Biobank is another limitation which I discussed earlier in the context of ascertaining chronic diseases such as CKD and diabetes. The small cohort size may weaken the reliability and generalisability of statistical analyses, especially for subgroup analyses. In addition, HFpEF is a syndrome that develops very gradually and is challenging to diagnose, meaning that the date at which it was first recorded clinically may not always capture the condition at consistent stages of disease progression in different patients. I acknowledge the newly emerging definition of obesity as a clinical illness linked to organ dysfunction due to excess adiposity [349]. However, this definition has not been widely adopted in medical practice yet, and ascertaining "clinical obesity" is complex and beyond the scope of this thesis, therefore I decided to define obesity using traditional BMI-based criteria in my study. While BMI is simple and widely used, it does not distinguish between fat mass and fat-free mass nor does it capture fat distribution in different regions of the body. Alternative metrics including waist circumference, waist-to-hip ratio, waist-to-height ratio, or imaging-based quantification of visceral adipose tissue using MRI or DXA may better reflect adiposity types and patterns, and could improve phenotypic precision in future studies. Cardiac imaging data under stress conditions could provide more information on exercise intolerance in the different subgroups identified, as HFpEF symptoms do not always manifest at rest [9]. The UK Biobank did not collect cardiac imaging data under exercise conditions but this could be an interesting consideration for the design of future cohort studies of HFpEF. While exercise ECG data is available, it was collected at a much earlier timepoint, and ECG analysis is beyond the scope of this study. Finally, causal inferences cannot be made given the design of the study, and the study did not include an external validation cohort. Due to these limitations, further phenotyping studies are warranted to validate these results, especially when translating findings to

HFpEF populations of different sex, socio-economic background and ethnicity, to address any biases that may exist due to the demographics of the UK Biobank.

The challenges of obtaining both large cohorts and "deep", i.e. multi-modal, data for highly robust HFpEF phenotyping studies are widely acknowledged in the field. To tackle this, the UK HFpEF Collaborative Group and the United States National Heart, Lung and Blood Institute, for example, have set up specific consortia to aggregate multiple existing HFpEF data sources and collect new data across a range of clinical modalities in order to enable larger-scale, more effective HFpEF phenotyping research and trials to take place [350–352]. These represent very exciting future avenues for HFpEF research in parallel to future UK Biobank studies once more cases of the condition accrue over time.

4.5 Conclusion

In this study, I provided focused insights into the distribution of cardiometabolic multimorbidity burden and metabolic syndrome profiles of men and women with HFpEF in an elderly cohort from the UK Biobank. Latent class analysis based on patients' cardiometabolic profile yielded three distinct, clinically relevant phenogroups, namely: an older, more male and multimorbid subgroup; a subgroup with a high prevalence of severe obesity with the highest relative proportion of females; and a generally healthier yet hypertensive subgroup with a less severe cardiometabolic profile. I found significant differences in body shape and mass measurements, lipid biomarkers and medication across the different groups identified. Differential analysis of imaging-derived parameters of left ventricular function suggested a higher cardiac output in the most obese subgroup, as well as higher global longitudinal strain in women compared to men in this subgroup. Coupled with current challenges and limitations of HFpEF treatment, the work presented in this Chapter aligns with the vision of personalised medicine by providing further evidence that supports deeper context-aware characterisation of the condition and subsequent disease management routes, especially regarding the control of modifiable yet highly prevalent risk factors such as hypertension and obesity. Further studies in larger cohorts are warranted to validate these findings, especially in women and populations of different ethnicities, while application of the LCA model in an independent validation cohort would strengthen

the findings (with caution regarding the translation of the model to populations with a different underlying distribution of model input variables).

Although lower rates of type 2 diabetes were observed in our study, likely due to dataset and sample population limitations discussed earlier, type 2 diabetes remains a common comorbidity in HFpEF and one of its most important clinical risk factors. Despite this well-established association, the links between the two conditions remain underexplored and incompletely understood, especially in earlier stages of disease. In particular, early markers of cardiac dysfunction and underlying mechanisms of adverse cardiac remodelling in type 2 diabetes are poorly characterised. These critical gaps form the primary focus of the next two chapters.

Absence of evidence is not the evidence of absence.

— Carl Sagan, American astronomer and science communicator (1934-1996)

5

Multi-Modal Characterisation of Early-Stage Cardiac Deterioration in Patients with Type 2 Diabetes

Building on the phenotypic characterisation of cardiometabolic heart failure with preserved ejection fraction (HFpEF) presented in Chapter 4, this chapter focuses on type 2 diabetes mellitus (T2DM), a key HFpEF risk factor. It examines subclinical cardiac changes preceding overt cardiovascular disease, integrating electrocardiographic and cardiac magnetic resonance biomarkers in a large UK Biobank cohort. Through this multi-modal analysis, the study aims to identify signatures of early diabetic myocardial remodelling, characteristics that may precede the development of HFpEF and cardiac arrhythmias in T2DM.

This chapter is based on the following publication:

- **Bertrand A.**, Lewis, A., Camps, J. et al. Multi-modal characterisation of early-stage, subclinical cardiac deterioration in patients with type 2 diabetes. Cardiovasc Diabetol 23, 371 (2024). DOI:10.1186/s12933-024-02465-y

5.1 Introduction

Type 2 diabetes mellitus (T2DM) affects over half a billion adults worldwide and is a major known risk factor for heart failure with preserved ejection fraction (HFpEF), the focus of the previous Chapter. As discussed in Chapter 2 (Section 2.5), T2DM promotes cardiac remodelling at multiple biophysical scales including ionic remodelling, diffuse myocardial fibrosis, and impaired contractility. If unmanaged, these changes worsen over time, leading to diastolic dysfunction and eventually HFpEF. The combination of structural and electrophysiological alterations in T2DM also creates a pro-arrhythmic substrate. In individuals without diagnosed cardiovascular disease (CVD), these abnormalities are often termed diabetic myocardial disorder, though this concept remains to be fully validated [19, 106]. The growing recognition that subclinical cardiac dysfunction in T2DM often precedes HFpEF motivates the need to better characterise markers of cardiac deterioration, in the early stages of disease before overt CVD, in order to mitigate disease progression and ultimately reduce deaths.

Cardiac magnetic resonance (CMR) imaging studies have identified structural and functional changes in small cohorts of individuals with diabetes (n=46, n=143) [111, 353], while electrophysiological abnormalities in diabetes may be reflected in the electrocardiogram (ECG) [160, 354–356] (see Chapter 2, Section 2.5). However, large-scale evidence of these early abnormalities prior to overt HFpEF and indeed other forms of cardiac disease remains limited, reinforcing the need for further investigation into the subclinical cardiac effects of T2DM and their role in HFpEF and arrhythmia development.

To the best of my knowledge, no previous work has quantified ECG and CMR-derived biomarkers concurrently in a large, matched cohort of individuals with diabetes of type 2 specifically (as opposed to type 1 or a combination of both), in the stages before clinically diagnosed CVD. Furthermore, the availability of coupled ECG and imaging data collected at same timepoint for all the UK Biobank imaging participants is a rare resource that represents a unique, unprecedented opportunity to study subclinical cardiac changes in individuals with T2DM at such scale. Thus, the goal of this study is to identify markers of subclinical T2DM-driven cardiac remodelling and to assess the

predictive value of multi-modal biomarkers in predicting CVD; the overarching aim being to support CVD risk stratification in patients with T2DM. I hypothesise that, compared to healthy matched controls, biomarker differences in individuals with T2DM may provide further evidence pointing towards the existence of diabetic myocardial disorder and accompanying electrophysiological abnormalities. Using the UK Biobank, I investigated the effect of T2DM on ECG and CMR-derived biomarkers reflecting cardiac morphology, haemodynamic function and electrophysiology, and quantified these biomarkers' predictive value in relation to CVD development. Sex, age, body mass index (BMI), and ethnicity-related differences were investigated in subgroup analyses, as well as differences across the glycaemic spectrum.

5.2 Methods

5.2.1 Study design and population

The study sample was selected as a subset of the UK Biobank, distinct from the one considered in Chapter 4. The primary analysis carried out in this Chapter is a matched cross-sectional study examining the association between T2DM status and multi-modal cardiac biomarkers. As a secondary analysis, I carried out a case-control study quantifying the relationship between selected biomarkers in participants who did and did not develop CVD during follow-up, in two matched cohorts with and without T2DM. As in Chapter 4, this study is reported in line with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement. Ethical approval for the study was obtained as outlined in Chapter 3, Section 3.1.4.

5.2.2 Cohort definition

Disease ascertainment

The selection process and cohort sizes are summarised in Figure 5.1. I defined the study cohorts based on the clinical ascertainment of two disease phenotypes, namely T2DM and CVD. I chose to consider CVD more broadly as opposed to HFpEF only, due to limitations that would have arisen due to a smaller sample size. Relevant ICD-9 and

ICD-10 codes and their corresponding UK Biobank fields are listed in Appendix B, Table B.1. T2DM was determined by one or more of the following criteria: date of a T2DM ICD code first reported, HES record corresponding to a T2DM ICD code, HbA1c ≥48 mmol/mol, response to a patient touchscreen questionnaire [357]. CVD was determined by one or more of the following criteria: date of a CVD ICD code first reported, HES record corresponding to a CVD ICD code, date of myocardial infarction (MI), date of ST-elevated MI, and date of non-ST-elevated MI. The last three data fields are UK Biobank-specific algorithmically defined diagnoses of MI summarising information contained in other fields; cases identified via those fields may overlap with other routes of ascertainment but were included for completeness.

Exclusion criteria, exposure and outcomes

All participants with pre-existing CVD were excluded, as were censored all participants with CVD diagnosed after 21/09/2021, the date of the latest CVD diagnosis identified at the time of analysis. The median follow-up period between date of assessment and date of first CVD diagnosis is 1.5 years (min 8 days, max 5.9 years).

The exposure cohort consists of eligible participants with known T2DM at time of assessment. Pair-matching was carried out to select controls with no known T2DM based on age, sex, BMI, as well as incident CVD during follow-up (i.e. with vs without a CVD diagnosis). Age, sex and BMI as a proxy for obesity status were selected due to their clinically important determining effect of both diabetes status and cardiac structure and function, and because matching on these variables improves baseline comparability between groups. I selected BMI as a suitable matching variable as it is a commonly used marker of adiposity that modulates the ECG [358]. A parsimonious matching strategy was intentionally used to preserve an adequate pool of eligible controls and avoid overmatching. More extensive matching on additional comorbidities e.g. hypertension was not performed as this would have substantially reduced the number of eligible controls and may have introduced overmatching, thereby reducing sample size and statistical efficiency. Residual confounding by variables not used in matching was addressed in part in the stratified regression analyses (see Chapter 3 and following sections). I also chose

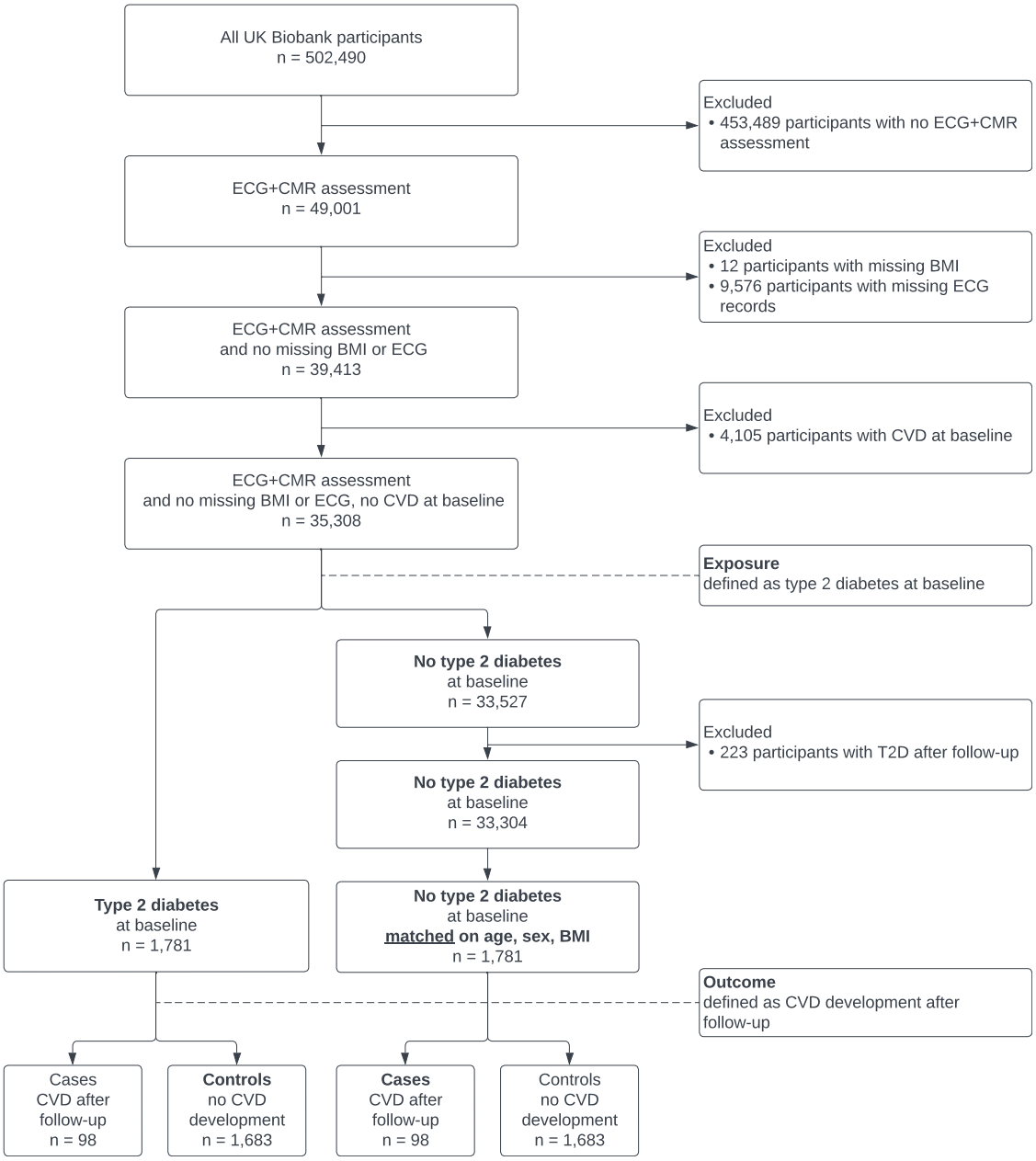

Figure 5.1: Cohort selection flowchart. CMR: cardiac magnetic resonance imaging, BMI: body mass index, CVD: cardiovascular disease, T2D: type 2 diabetes. Baseline refers to the time of recording of the ECG and CMR images. Follow-up refers to the period between baseline and censoring date (21/09/2021).

to match participants on incident CVD status to account for the potential confounding effects of undiagnosed or developing cardiovascular disease, which may independently influence cardiac structure and function. Thus, this approach aimed to better isolate associations attributable to T2DM, although some bias may remain due to undiagnosed or unrecorded cases. Euclidean distances were computed between the [age, BMI] vectors of each exposure participant and possible control participants. Each exposure participant was matched to a control participant of the same sex and follow-up CVD status with the lowest [age, BMI] Euclidian distance [359].

Outcomes were defined as ECG and CMR biomarker values for all analyses, apart from a secondary analysis considering CVD development, where an outcome was defined as a new case of incident CVD during follow-up.

5.2.3 Baseline covariates

Demographic, lifestyle, and clinical characteristics

I considered the following demographic, lifestyle, and clinical characteristics: age at assessment, sex, ethnicity, weight, height, smoking status, systolic blood pressure (SBP), diastolic blood pressure (DBP), medication. BMI was calculated using the ratio of weight in kilograms over squared height in meters. Information on medication was self-reported and divided into broad categories and/or using UK Biobank-specific individual medication codes recorded by nurses. I included cholesterol-lowering medication, blood pressure medication, insulin, metformin and rosiglitazone 1mg/metformin 500mg tablet. The following blood biochemistry measurements were included: high-density lipoprotein (HDL) cholesterol, total cholesterol, HbA1c, triglycerides, creatinine, C-reactive protein. These measurements were taken prior to the imaging assessment so I used the most recent one available. Estimated glomerular filtration rate (eGFR) was calculated using serum creatinine according to the CKD-EPI creatinine equation adjusted for age, sex and ethnicity [360]. Creatinine was converted to mg/dL by multiplying the given value in μmol/L by a factor of 0.0113 [361].

ECG and CMR-derived biomarkers

Resting ECG recordings were available in .xml file format. Each file included the raw ECG signal as well as automatically extracted key ECG markers. Wave amplitudes were available for all 12 leads, while other markers were available for lead I only. I used the following markers: ventricular rate, QRS duration, QTc interval, T wave offset, R-wave amplitude, S-wave amplitude, J point amplitude (equivalent to ST segment elevation or depression), and T wave amplitude. Pairwise Pearson correlation coefficients were calculated to assess inter-covariate correlations among T wave and J point amplitudes in different ECG leads (Appendix B, Figure B.3). Two representative leads were retained, one limb lead (aVL) and one precordial lead (V3). I computed the Sokolow-Lyon index for left ventricular hypertrophy as the sum of the S amplitude in V1 and the maximum value of R amplitude in V5 or V6 [362].

Some key CMR-derived biomarkers were directly available in the UK Biobank database. Manual analyses, algorithms and validation of these biomarkers are described in detail in [232, 233]. I considered the following markers: left ventricular ejection fraction (LVEF), LV end-diastolic volume, LV end-systolic volume, LV stroke volume, cardiac output, LV myocardial mass, LV global average wall thickness. I calculated body surface area (BSA) using Du Bois' formula, BSA = $0.007184 \times w^{0.425} \times h^{0.725}$ where w is weight in kilograms and h is height in cm, to additionally obtain LV mass index (LVMI = LV mass / BSA).

5.2.4 Statistical analysis

All statistical analyses were performed using RStudio with R version 4.1.1.

Comparison of ECG and CMR biomarker distributions between cohorts

Baseline covariates were compared between the exposure and control cohorts following the steps described in Chapter 3 (section 3.3.1). I performed comparisons of ECG and CMR biomarkers between groups to characterise overall differences between participants with and without T2DM. Statistical significance was defined as $p < 0.05$, without multiple-comparison correction, as each ECG and CMR biomarker represents a distinct physiological parameter and the comparisons are primarily descriptive and exploratory

in nature. Missing data were not imputed. All statistical tests were two-sided. Median, interquartile range (IQR) and missing measurements in the form of count and percentage of the full sample, are reported for each variable.

Linear regression modelling to assess effect of T2DM on ECG and CMR biomarkers

Formal assessment of the independent association between T2DM and individual ECG and CMR biomarkers was subsequently performed using multiple linear regression models, used to estimate adjusted mean differences in continuous biomarkers by T2DM status. With a binary exposure, linear regression estimates the adjusted difference in mean biomarker levels between individuals with and without T2DM, conditional on the included covariates. Models were sequentially adjusted for relevant covariates. A crude, unadjusted model (Model 0) was used to quantify the association of T2DM only with each biomarker. Three additional models were constructed incrementally to account for selected socio-demographic (Model 1), lifestyle (Model 2), and clinical covariates detailed in the previous section (Model 3; fully adjusted). Inter-correlation between predictor variables can sometimes dilute true associations and lead to larger, imprecise confidence intervals. To mitigate this effect, I calculated pairwise Pearson correlation coefficients between potential model covariates and, for each covariate pair with a coefficient magnitude greater than 0.3, one covariate was excluded and the other retained in the adjusted models. This 0.3 threshold has been used to assess covariate correlation in a previous study [363] and is also widely used in the field of medicine as an indicator of "fair" correlation [364]. This determined the exclusion of the following covariates from our models: eGFR, HbA1c, systolic blood pressure, HDL cholesterol, cholesterol-lowering medication, and metformin. Although HbA1c was excluded (Pearson correlation coefficient with T2DM = 0.57, see Appendix B, Figure B.4), a subgroup analysis was carried out to assess the association of HbA1c, as a proxy for hyperglycaemia, with biomarker outcomes of interest. Models were fitted using the Ordinary Least Squares method, therefore scaling of variables was not required. Regression coefficients, their 95% confidence intervals, and p-values are reported for each outcome variable and model in the tables.

Model diagnostics

To assess linear regression assumptions, diagnostic plots were examined for each model. For every ECG and CMR outcome, I examined standard diagnostic plots (residuals vs. fitted values to assess linearity and homoscedasticity; normal Q–Q plots of residuals to assess normality; scale–location plots to assess homoscedasticity; and residuals vs. leverage plots to identify influential observations and outliers). Variance inflation factors (VIFs) were calculated to assess multicollinearity of predictors, with values below 5 considered acceptable. Representative diagnostic plots for primary analyses are provided in Appendix B, Figure B.1 and Figure B.2.

Outcome analyses

I compared biomarkers between cases who did and did not develop CVD, in both cohorts with and without T2DM. I assessed the association between biomarkers with significant differences, and CVD development. I built two sets of logistic regression models, one for each cohort, and examined how these associations differ in each cohort. I followed the same covariate processing and sequential adjustment approach as described previously.

5.2.5 Subgroup analyses

I examined the role of sex, age, BMI and ethnicity by comparing biomarkers in stratified subgroups. I investigated the effect of severity of disease on selected biomarkers by considering HbA1c as a proxy for hyperglycaemia. I built regression models as described previously, this time using HbA1c as a continuous exposure variable, and biomarkers that were associated with T2DM as outcomes. This analysis was performed within the T2DM cohort only, to avoid a bimodal HbA1c distribution due to lower HbA1c levels in the control cohort. Diagnostic plots and VIF tests were also run as detailed above; representative plots are provided in Appendix B, Figure B.1 and Figure B.2).

5.3 Results

5.3.1 Cohort description

Both the exposure and the control cohorts (T2DM vs no T2DM) consisted of 1,781 subjects matched by age, sex, BMI and CVD status. Cohorts were mostly male (n=1133, 63.6%), had a median age of 67 years and a median BMI of 27.8 kg/m^2, which is considered overweight by the UK National Health Service (geq25 kg/m^2) (Table 5.1).

Table 5.1: Socio-demographic, lifestyle, and clinical characteristics recorded for the cohorts with and without type 2 diabetes. IQR: inter-quartile range, BMI: body mass index, SBP: systolic blood pressure, DBP: diastolic blood pressure, HDL: high-density lipoprotein, HbA1c: glycated haemoglobin A1c, eGFR: estimated glomerular filtration rate.

	No type 2 diabetes		**Type 2 diabetes**			
	Median (IQR) or Count (%)	**Total count (%)**	**Median (IQR) or Count (%)**	**Total count (%)**	**t-statistic**	**p-value**
Full cohort	1781 (100)	1781 (100)	1781 (100)	1781 (100)	-	-
Socio-demographics						
Age, years	67 (60-71)	1781 (100)	67 (60-71)	1781 (100)	-	0.63
Sex: female	648 (36.4)	1781 (100)	648 (36.4)	1781 (100)	0	1
Ethnicity	-	-	-	-	66.8	< 0.001
White	1733 (97.3)	1781 (100)	1640 (92.1)	1781 (100)	-	-
Black	15 (0.8)	1781 (100)	37 (2.1)	1781 (100)	-	-
Other	33 (1.9)	1781 (100)	104 (5.8)	1781 (100)	-	-
Lifestyle characteristics						
BMI, kg/m^2	27.7 (24.8-31.2)	1781 (100)	28.0 (25.1-31.6)	1781 (100)	-	0.05
Smoking status	-	-	-	-	6.4	0.04
Never	1047 (59.4)	1762 (99.0)	994 (56.4)	1761 (98.9)	-	-
Previous	648 (36.8)	1762 (99.0)	694 (39.4)	1761 (98.9)	-	-
Current	67 (3.8)	1762 (99.0)	73 (4.1)	1761 (98.9)	-	-
Clinical characteristics						
Average SBP, mmHg	142 (131-155)	1577 (88.5)	142 (132-155)	1560 (87.6)	-	0.68
Average DBP, mmHg	81 (74-88)	1577 (88.5)	79 (72-85)	1560 (87.6)	-	< 0.001
Total cholesterol, mmol/L	5.7 (5.0-6.5)	1673 (93.9)	5.1 (4.2-6.0)	1686 (94.7)	-	< 0.001
HDL cholesterol, mmol/L	1.4 (1.1-1.6)	1557 (87.4)	1.2 (1.0-1.4)	1580 (88.7)	-	< 0.001
HbA1c, mmol/mol	34.7 (32.5-37.1)	1669 (93.7)	43.7 (38.9-51.0)	1672 (93.9)	-	0
Creatinine, µmol/L	75.6 (65.5-84.4)	1672 (93.9)	73.5 (63.5-83.4)	1684 (94.6)	-	< 0.001
Triglycerides, mmol/L	1.6 (1.1-2.2)	1672 (93.9)	1.9 (1.3-2.7)	1681 (94.4)	-	< 0.001
C-reactive protein, mg/L	1.4 (0.7-2.8)	1673 (93.9)	1.8 (0.9-3.5)	1681 (94.4)	-	< 0.001
eGFR, ml/min/1.73m^2	86.1 (77.5-92.5)	1672 (93.9)	87.8 (78.2-94.3)	1684 (94.6)	-	< 0.001
Medication						
Cholesterol-lowering	485 (27.2)	1781 (100)	1287 (72.3)	1781 (100)	1801.8	0
Antihypertensive	487 (27.3)	1781 (100)	1003 (56.3)	1781 (100)	607.7	< 0.001
Insulin	1 (0.1)	1781 (100)	244 (13.7)	1781 (100)	280.4	< 0.001
Metformin	1 (0.1)	1781 (100)	945 (53.1)	1781 (100)	2009	0

Individuals with T2DM were more likely to be non-white (7.9% vs 2.7%) and to be current or previous smokers (43.5% vs 40.6%). They tended to have lower diastolic blood pressure, total cholesterol, and HDL cholesterol, while HbA1c, triglycerides, C-reactive protein, and eGFR were generally more elevated. The latter may indicate hyperfiltration, defined by an elevated estimated eGFR >120-140 mL/min/1.73 m^2. It is common in early stages of diabetes including type 2, caused by increased pressure in the kidney's glomerulus due to high blood glucose, sodium-glucose reabsorption via the SGLT2 protein and renal growth. Future work could consider the use of other markers of kidney function such as proteinuria (presence of excess protein in the urine) to further characterise renal function in high-risk groups such as T2DM. Subjects in the T2DM cohort were also more likely to be on cholesterol-lowering drugs, anti-hypertensives, insulin and metformin. One patient in the control cohort reported taking insulin and metformin. This may be due to a lack of diabetes diagnosis at the time of visit, or an increase in HbA1c levels after the last measurement, leading to potential misclassification. Although rare, this case could also be due to the medication being used to treat another condition unrelated to diabetes, such as polycystic ovary syndrome [365] or advanced pancreatic failure.

5.3.2 Model diagnostics

Visual inspection of diagnostic plots demonstrated no major violations of linear regression assumptions (see representative plots for fully adjusted models in Appendix B, Figures B.1 and B.2). Although statistical tests suggested some heteroscedasticity, this was expected given the large sample size and was not considered to affect model estimates. Variance inflation factors were below the threshold of 5 across all models (maximum VIF was 1.31 for main analyses, 1.30 for subgroup HbA1c-stratified analyses), indicating no evidence of multicollinearity.

5.3.3 Type 2 diabetes is associated with a higher resting heart rate, longer QTc interval, reduced T wave amplitude and lower Sokolow-Lyon index

Compared to matched controls, ventricular rate in T2DM was higher (+5 bpm, median: 66 IQR: [59-74] vs 61 [55-68] bpm) and T wave offset was earlier (-12 ms, 842 [820-864]

vs 854 [834-874] ms) (Table 5.2). T2DM was strongly associated with both biomarkers in all models (Table 5.3). This consistency is expected, given the known inverse correlation between T wave offset and ventricular rate. QRS duration was shorter (-2 ms, 86 [79-93] vs 88 [82-96] ms) and QTc interval was longer (+4 ms, 424 [408-440] vs 420 [405-436] ms) in the T2DM cohort (Table 5.2). There was a statistically significant association of T2DM with QTc interval in all models but the fully adjusted one (beta=1.83, 95% CI=[0.13, 3.53], p=0.035 before adjustment vs beta=0.86, 95% CI=[-1.22, 2.95], p=0.416 after adjustment) (Table 5.3. However, the association of T2DM with QRS duration was significant only in the two last models adjusted for lifestyle and clinical factors (Model 2 and 3, Table 5.3). This suggests that the importance of this variable may increase relative to other covariates included in the later models. There were statistically significant differences in almost all wave amplitude biomarkers between the cohorts (Table 5.2). T2DM was associated with a lower T wave amplitude in V3 in all models, and a lower T wave amplitude and less elevated J-point in aVL in all models but the fully adjusted one (Table 5.3). T2DM was associated with a lower Sokolow-Lyon index in all models but the fully adjusted one (Tables 5.2 and 5.3).

5.3.4 Type 2 diabetes is independently associated with reduced stroke volume and higher left ventricular wall thickness

Left ventricular wall thickness was higher in the T2DM cohort (+2 mm, 6.1 [5.6-6.6] vs 5.9 [5.4-6.5] mm) (Table 5.2), and its association with T2DM was statistically significant in all models (Table 5.3). Left ventricular stroke volume was lower in the T2DM cohort (-6 ml, 72 [60-85] vs 78 [65-90] ml), as were end-diastolic and end-systolic volumes (Table 5.2). However, the association of T2DM was significant in all models only with stroke volume (Table 5.3). No statistically significant differences or associations were observed for LVEF nor cardiac output.

Table 5.2: ECG and CMR-derived biomarkers recorded for the cohorts with and without type 2 diabetes. IQR: inter-quartile range; ECG: electrocardiogram; CMR: cardiac magnetic resonance; bpm: beats per minute; LV: left ventricular; EF: ejection fraction; ED: end-diastolic; ES: end-systolic. All continuous variables are distributed non-normally and compared using the Mann–Whitney U-test.

	No type 2 diabetes		**Type 2 diabetes**		**p-value**
	Median (IQR)	**N (%)**	**Median (IQR)**	**N (%)**	
Full cohort	–	1781 (100)	–	1781 (100)	–
ECG biomarkers					
Ventricular rate, bpm	61 (55–68)	1781 (100)	66 (59–74)	1781 (100)	<0.001
QRS duration, ms	88 (82–96)	1781 (100)	86 (79–93)	1781 (100)	0.005
QTc interval, ms	420 (405–436)	1781 (100)	424 (408–440)	1781 (100)	<0.001
T-wave offset, ms	854 (834–874)	1781 (100)	842 (820–864)	1781 (100)	<0.001
T-wave amplitude (V3), mV	0.37 (0.23–0.54)	1766 (99.2)	0.33 (0.21–0.49)	1765 (99.1)	<0.001
T-wave amplitude (aVL), mV	0.11 (0.05–0.17)	1718 (96.5)	0.10 (0.04–0.15)	1669 (93.7)	<0.001
J-point amplitude (V3), mV	–0.02 ((–0.04)–0.02)	1781 (100)	–0.02 ((–0.04)–0.02)	1781 (100)	0.89
J-point amplitude (aVL), mV	0.01 ((–0.01)–0.03)	1781 (100)	0.00 ((–0.02)–0.03)	1781 (100)	0.016
Sokolow–Lyon index, mm	20.2 (16.2–24.8)	1650 (92.6)	19.1 (15.2–23.5)	1612 (90.5)	<0.001
CMR biomarkers					
LVEF, %	56 (52–59)	1533 (86.1)	55 (51–59)	1501 (84.3)	0.066
LV ED volume, mL	140 (119–163)	1533 (86.1)	130 (109–155)	1501 (84.3)	<0.001
LV ES volume, mL	61 (51–74)	1533 (86.1)	58 (47–71)	1501 (84.3)	<0.001
LV stroke volume, mL	78 (65–90)	1533 (86.1)	72 (60–85)	1501 (84.3)	<0.001
Cardiac output, L/min	4.7 (4.1–5.5)	1533 (86.1)	4.7 (4.0–5.5)	1501 (84.3)	0.53
LV mass, g	92 (76–109)	1518 (85.2)	91 (77–108)	1518 (85.2)	0.87
LV mass index, g/m^2	47 (41–53)	1518 (85.2)	47 (41–52)	1518 (85.2)	0.48
LV global average wall thickness, mm	5.9 (5.4–6.5)	1516 (85.1)	6.1 (5.6–6.6)	1516 (85.1)	<0.001

Table 5.3: Multiple linear regression models used to quantify the association of type 2 diabetes with selected ECG and imaging-derived biomarkers. Models are adjusted sequentially for different types of confounding factors. Socio-demographic factors include age, sex, ethnicity; lifestyle factors include body mass index (BMI), smoking; clinical factors include diastolic blood pressure, total cholesterol, triglycerides, C-reactive protein, anti-hypertensive medication and insulin. Bpm: for beats per minute, LV: left ventricular, EF: ejection fraction, ED: end-diastolic, ES: end-systolic.

	Model 0 **Unadjusted** **N = 2577**		**Model 1** **Adjusted for socio-demographic factors** **N = 2577**		**Model 2** **Additionally adjusted for lifestyle factors** **N = 2544**		**Model 3** **Additionally adjusted for clinical factors** **N = 2201**	
Outcome	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**
ECG								
Ventricular rate, bpm	4.14 (3.30–4.98)	<0.001	4.15 (3.31–4.98)	<0.001	3.84 (3.01–4.67)	<0.001	3.11 (2.11–4.10)	<0.001
QRS duration, ms	-0.76 (-1.81–0.28)	0.151	-0.89 (-1.89–0.10)	0.079	-1.14 (-2.14—0.14)	0.026	-1.81 (-3.01—0.61)	0.003
QTc interval, ms	2.57 (0.78–4.37)	0.005	2.48 (0.76–4.20)	0.005	1.83 (0.13–3.53)	0.035	0.86 (-1.22–2.95)	0.416
T-wave offset, ms	-10.5 (-12.9—8.0)	<0.001	-10.6 (-13.0—8.2)	<0.001	-10.1 (-12.5—7.6)	<0.001	-8.3 (-11.2—5.3)	<0.001
T-wave amplitude (V3), mV	-0.04 (-0.06—0.02)	<0.001	-0.038 (-0.06—0.02)	<0.001	-0.031 (-0.048—0.014)	<0.001	-0.025 (-0.047—0.004)	0.021
T-wave amplitude (aVL), mV	-0.01 (-0.02—0.01)	<0.001	-0.01 (-0.02—0.01)	<0.001	-0.01 (-0.02—0.01)	<0.001	-0.005 (-0.015–0.004)	0.285
J-point amplitude (V3), mV	0.002 (-0.003–0.006)	0.437	0.002 (-0.003–0.006)	0.470	0.003 (-0.001–0.007)	0.198	0.003 (-0.002–0.008)	0.218
J-point amplitude (aVL), mV	-0.004 (-0.007—0.001)	0.007	-0.004 (-0.006—0.001)	0.009	-0.003 (-0.006—0.001)	0.019	-0.001 (-0.004–0.002)	0.573
Sokolow-Lyon index, mm	-0.83 (-1.34—0.32)	0.001	-0.80 (-1.29—0.30)	0.002	-0.61 (-1.09—0.12)	0.015	-0.39 (-0.99–0.21)	0.199
CMR								
LVEF, %	-0.415 (-0.952–0.123)	0.131	-0.371 (-0.902–0.161)	0.171	-0.289 (-0.825–0.247)	0.290	-0.195 (-0.839–0.449)	0.553
LV ED volume, mL	-8.63 (-15.2—2.09)	0.010	-8.44 (-14.8—2.03)	0.010	-9.11 (-15.6—2.61)	0.006	-6.86 (-15.3–1.55)	0.110
LV ES volume, mL	-3.47 (-9.09–2.15)	0.226	-3.48 (-9.06–2.10)	0.221	-3.95 (-9.62–1.71)	0.171	-2.79 (-10.2–4.59)	0.459
LV stroke volume, mL	-5.18 (-6.87—3.50)	<0.001	-4.98 (-6.53—3.42)	<0.001	-5.17 (-6.74—3.60)	<0.001	-4.11 (-6.03—2.19)	<0.001
Cardiac output, L/min	-0.0517 (-0.158–0.055)	0.341	-0.038 (-0.137–0.061)	0.453	-0.0622 (-0.161–0.037)	0.219	-0.0429 (-0.165–0.080)	0.493
LV mass, g	0.359 (-1.38–2.10)	0.687	0.466 (-0.846–1.78)	0.487	-0.388 (-1.59–0.811)	0.525	-0.165 (-1.63–1.30)	0.825
LV mass index, g/m²	-0.17 (-0.83–0.49)	0.612	-0.16 (-0.70–0.38)	0.562	-0.24 (-0.78–0.31)	0.393	-0.10 (-0.77–0.56)	0.763
LV global average wall thickness, mm	0.173 (0.113–0.232)	<0.001	0.174 (0.125–0.223)	<0.001	0.139 (0.094–0.183)	<0.001	0.133 (0.081–0.186)	<0.001

Table 5.4: Logistic regression models used to quantify the association of selected biomarker predictors with incident cardiovascular disease (single binary outcome). Models are adjusted sequentially for different types of confounding factors. Socio-demographic factors include age, sex, ethnicity; lifestyle factors include body mass index (BMI), smoking; clinical factors include diastolic blood pressure, total cholesterol, triglycerides, C-reactive protein, anti-hypertensive medication and insulin.

		Model 0 Unadjusted		**Model 1 Adjusted for socio -demographic factors**		**Model 2 Additionally adjusted for lifestyle factors**		**Model 3 Additionally adjusted for clinical factors**	
Predictor	**N**	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**
No type 2 diabetes									
QRS duration, ms	1781	0.016 (0.004–0.027)	0.007	0.015 (0.002–0.027)	0.019	0.014 (0.002–0.026)	0.022	0.0137 (0–0.027)	0.046
Sokolow-Lyon index, mm	1650	0.049 (0.019–0.079)	0.001	0.049 (0.017–0.081)	0.003	0.051 (0.017–0.084)	0.003	0.051 (0.015–0.086)	0.005
LV mass, g	1518	0.017 (0.008–0.025)	<0.001	0.025 (0.014–0.036)	<0.001	0.027 (0.015–0.039)	<0.001	0.029 (0.017–0.042)	<0.001
LV mass index, g/m^2	1518	0.047 (0.024–0.068)	<0.001	0.058 (0.032–0.083)	<0.001	0.054 (0.027–0.080)	<0.001	0.059 (0.031–0.087)	<0.001
LV global average wall thickness, mm	1516	0.508 (0.253–0.760)	<0.001	0.623 (0.316–0.926)	<0.001	0.707 (0.357–1.050)	<0.001	0.677 (0.293–1.060)	<0.001
Type 2 diabetes									
QRS duration, ms	1781	0.014 (0.002–0.026)	0.020	0.013 (0–0.025)	0.045	0.013 (0–0.025)	0.042	0.009 (-0.007–0.023)	0.263
Sokolow-Lyon index, mm	1612	0.031 (-0.001–0.062)	0.053	0.030 (-0.003–0.062)	0.071	0.036 (0.002–0.069)	0.032	0.021 (-0.018–0.058)	0.280
LV mass, g	1518	0.010 (0.001–0.019)	0.022	0.014 (0.003–0.024)	0.013	0.016 (0.003–0.027)	0.011	0.019 (0.006–0.032)	0.003
LV mass index, g/m^2	1518	0.034 (0.009–0.056)	0.006	0.038 (0.011–0.064)	0.005	0.039 (0.012–0.066)	0.004	0.047 (0.017–0.075)	0.001
LV global average wall thickness, mm	1516	0.354 (0.091–0.614)	0.008	0.413 (0.107–0.710)	0.007	0.457 (0.115–0.790)	0.008	0.561 (0.191–0.923)	0.003

5.3.5 QRS duration has a positive association with cardiovascular outcomes in type 2 diabetes

The breakdown of follow-up CVDs in each cohort is presented in Figure 5.2. T2DM individuals who went on to develop CVD (n=98, Figure 5.2) compared to those who did not (n=1683) were more likely to be male (70% vs 64%) and more likely to take metformin (57% vs 53%). In this cohort of subjects with diabetes, there were no changes observed in blood pressure, cholesterol, HbA1c, or eGFR between individuals who do and do not develop CVD. In the T2DM cohort, cases who developed CVD during follow-up (n=98) had a longer QRS duration (88 [82-98] vs 86 [80-94] ms, p=0.03) (Appendix B, Table B.3). There was a statistically significant association between QRS duration and CVD development in all models but the fully adjusted one (Table 5.4). In subjects without T2DM, cases of CVD (n=98) had a higher Sokolow-Lyon index (23.4 [17.7-27.1] vs 20.1 [16.2-24.7] mm, p=0.005) (Appendix B, Table B.3). I found a statistically significant association between the Sokolow-Lyon index and CVD development in all models (Table 5.4).

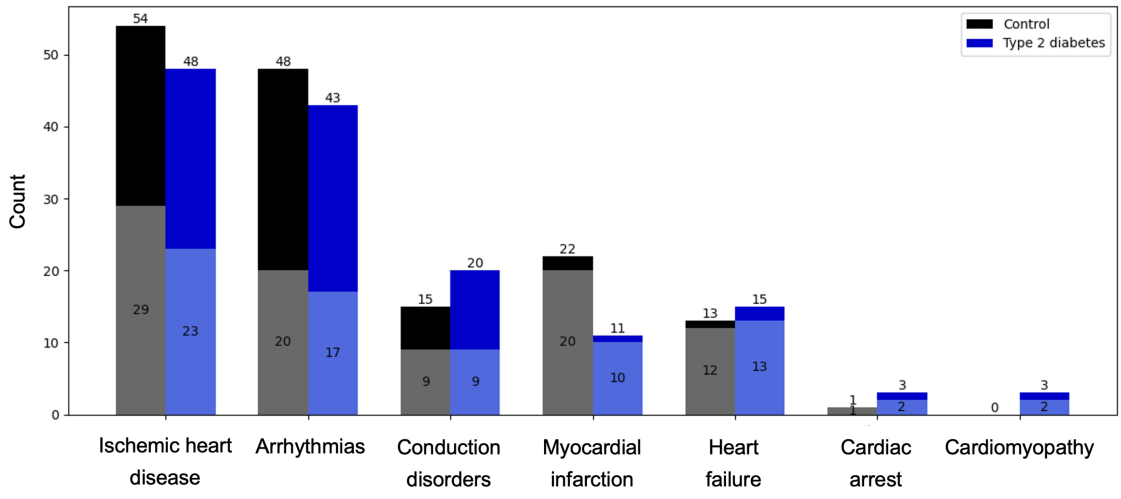

Figure 5.2: Incident cardiovascular disease diagnoses in the control and type 2 diabetes cohorts during the follow-up period after the imaging and ECG assessment. Total count for each CVD category is indicated at the top of each bar. The shaded areas represent cases where this CVD exists in conjunction with one or more other CVD diagnoses. In each cohort, n=98 individuals developed one or more CVD(s) during the follow-up period. CVD: cardiovascular disease.

5.3.6 Left ventricular mass, left ventricular mass index and wall thickness have a positive association with cardiovascular outcomes in cohorts with and without type 2 diabetes

In both cohorts, cases with CVD had a higher left ventricular mass (T2DM: 100 [86-107] vs 91 [76-108] g, p=0.006; no T2DM: 101 [83-118] vs 92 [76-108] g, p<0.001), a higher LVMI (T2DM: 48.7 [45.3-54.0] vs 46.4 [41.2-52.1] g/m2, p=0.001; no T2DM: 50.3 [43.0-57.5] vs 46.7 [41.1-52.8] g/m2, p=0.001), and a thicker left ventricular wall (T2DM: 6.3 [5.9-6.8] vs 6.1 [5.6-6.6] mm, p=0.006; no T2DM: 6.3 [5.6-6.9] vs 5.9 [5.4-6.4] mm, p<0.001) (Appendix B, Tables B.3 and B.2). All three variables showed a statistically significant association with CVD in all models (Table 5.4).

5.3.7 Subgroup analyses

Sex, age, and body mass index

Similar trends were observed between biomarkers of subjects with and without T2DM, in females (n=648) and males (n=1133). Compared to the control cohort, both females and males in the T2DM group exhibited higher ventricular rates, a prolonged QTc interval, an earlier T wave offset, lower T wave amplitudes in leads V3 and aVL, a lower Sokolow-Lyon index, lower end-diastolic, end-systolic, and stroke volumes, and a thicker left ventricular wall (Appendix B, Tables B.5 and B.4). As expected, baseline biomarker values were different in males and females. Sex-specific results stratified by age and BMI are included in Appendix B, Figures B.5 and B.6. No notable differences were observed across age groups. However, it is important to note that due to the design of the UK Biobank, the study only included middle-aged and elderly participants. Potential implications of this demographic trait are considered in the Discussion, notably regarding women and the effects of menopause on CVD risk. In different BMI groups, there was a gradual decrease in Sokolow-Lyon index with increasing BMI in both cohorts, likely explained by higher electrical impedance, and a steady increase in wall thickness with increasing BMI (Appendix B, Figure B.6).

Ethnicity

A higher proportion of participants from non-white ethnic backgrounds had T2DM (141 of 189, 74.6%), compared to participants of white ethnicity with T2DM (1640 of 3373, 48.6%). In both ethnicity subgroups, participants with T2DM showed a significantly higher ventricular rate and an earlier T wave offset (Appendix B, Tables B.6 and B.7). A significant QRS shortening, QTc prolongation, lower T wave amplitude, lower Sokolow-Lyon score, lower end-diastolic, end-systolic, and stroke volumes, and lower left ventricular wall thickness were only observed in white participants. In contrast, no significant changes were observed in the non-white subgroup for these biomarkers.

Association of HbA1c with ECG and CMR-derived biomarkers

In individuals with T2DM, I found a statistically significant association of higher levels of HbA1c with both a higher ventricular rate and an earlier T wave offset in all models (Appendix B, Table B.8). As noted earlier, heart rate and T wave offset are strongly correlated, so this consistent association is expected. Higher levels of HbA1c were also associated with a lower T wave amplitude in lead V3, and a lower stroke volume in all but the fully adjusted model.

5.4 Discussion

This study is the first to-date to provide a large-scale, concurrent analysis of ECG and CMR-derived biomarkers in a population with T2DM and no prior history of diagnosed CVD. Compared to matched controls, individuals in the T2DM cohort had a higher resting ventricular rate, prolonged QTc interval, reduced T wave amplitude, thicker left ventricular wall and lower end-diastolic, end-systolic and stroke volumes. T2DM was associated with all these changes, notably showing a strong independent association with heart rate, left ventricular wall thickness and stroke volume after full model adjustment. These findings may reflect subclinical diabetes-induced pathophysiological changes, notably the presence of diabetic myocardial disorder, which has been debated for many years [19, 106]. I discuss the clinical implications of these findings below.

5.4.1 Implications of a higher resting heart rate and QTc prolongation on cardiac autonomic neuropathy and arrhythmias

T2DM showed a strong, independent association with a higher resting heart rate, even after adjustment for potential confounders including medication. High heart rates have been associated to hypertension and metabolic conditions for many decades [366, 367]. In T2DM subjects specifically, this is typically a consequence of cardiac autonomic neuropathy [368] (see Chapter 2, Section 2.5.3). Cardiac autonomic neuropathy causes heart rate to increase in the early stages of disease and gradually decrease back to normal, but with reduced variability and an increased likelihood of arrhythmias [149]. This may explain why heart rates in our T2DM cohort lie mostly within the normal clinical range of 60-100 beats per minute. Additionally, I found a significant association between HbA1c and resting heart rate, supporting evidence linking dysregulations in glucose metabolism with increased sympathetic nervous system activity and severity of cardiac autonomic neuropathy [369–371].

I also found a significant positive association of T2DM with the QTc interval. The strength of this association decreased after adjusting for clinical covariates, suggesting that other factors may be partly responsible for the observed QTc interval prolongation. In the T2DM cohort, 113 women and 159 men (n=272 of 1781 [15%]) met sex- and age-specific criteria for clinical QTc prolongation [372], compared to n=211 (12%) in the control cohort. QTc prolongation is a well-established indicator of an increased risk of cardiac arrhythmias, and was also correlated with a higher cardiac autonomic neuropathy score in patients with T2DM [160].

Potential underlying ionic-level drivers of QT prolongation in T2DM were presented in detail in Chapter 2 (Section 2.5.4), and we will explore this further in the next chapter.

5.4.2 T wave amplitude reduction, asymptomatic ischemic heart disease and metabolic disturbances

T2DM subjects exhibited lower T wave amplitudes in leads V3 and aVL. BMI was included both as a matching variable and as a model covariate, to minimise the effect of

body composition on electrical impedance and subsequent reductions in ECG amplitude. Thus, our results suggest a plausible direct effect of T2DM on T wave amplitude. A recent ECG study also showed that T wave amplitudes were reduced in individuals with T2DM [373]. However, some subjects in this study had a history of myocardial infarction, which itself is likely to affect the T wave amplitude; our study overcomes this limitation as the study sample considered here is free of prior diagnosed cardiovascular conditions. That being said, asymptomatic ischemic heart disease may be partly responsible for changes observed. A reduced T wave amplitude has previously been associated with hypokalaemia, as well as hyperinsulinemic hypoglycaemia, which is common in T2DM [374, 375]. Thus, the lower T wave amplitudes observed may reflect underlying changes in potassium, insulin or glucose levels, or silent ischemia.

5.4.3 Lower stroke volume and thicker ventricular wall, subtle changes in the trajectory of heart failure with preserved ejection fraction

I observed a significantly reduced stroke volume in the T2DM cohort, accompanying a higher heart rate, while no change was observed in cardiac output. A previous study examined hemodynamic parameters, including stroke volume, in a smaller cohort of 143 subjects with type 1 and 2 diabetes [353]. Based on a much larger matched cohort with type 2-specific diabetes, the results presented in this chapter corroborate these findings. I also found that T2DM was associated with a thicker left ventricular wall, independently of blood pressure. This is consistent with previous studies that have demonstrated that patients with diabetes, hypertension, or both, tend to develop a thicker myocardial wall [107, 376, 377]. In later stages of disease, left ventricular hypertrophy is also common in patients with T2DM due to increased myocardial steatosis which triggers hypertrophic signalling and concentric remodelling [111] (see Chapter 2, Section 2.5.2). Patients with HFpEF also tend to have a lower stroke volume and higher left ventricular wall thickness [39, 378]. Thus, the changes observed in this study may represent early, subclinical stages of adverse remodelling in the trajectory toward HFpEF. However, measurements of left ventricular end-diastolic volume in our T2DM cohort remained within the normal sex-specific reference ranges that have been established for healthy

adults without diabetes in the UK Biobank [233]. In the T2DM cohort, the median wall thickness (6.1 mm) remained well below the clinical threshold for left ventricular hypertrophy (end-diastolic maximal wall thickness 15.0 mm) [379], and the absolute change compared to the control cohort was subtle (+0.2 mm).

Despite a thicker left ventricular wall, the Sokolow-Lyon index was lower in the T2DM cohort, suggesting lower R and S wave amplitudes. The discordance between this index and left ventricular hypertrophy as assessed by CMR is interesting but plausible. In obese patients, QRS voltages are artificially reduced due to higher body fat causing electrical impedance [358]. Despite individuals having similar BMI and left ventricular mass, this observed difference in Sokolow-Lyon index may be driven by a number of factors including varying chest wall shapes and subcutaneous fat, additional myocardial fat infiltration and deposition contributing to non-electrically active left ventricular mass, cardiac amyloidosis (causing a thicker heart with low voltages), myocardial diffuse fibrosis also replacing conductive tissue, and voltage criteria inaccuracy in metabolically abnormal populations.

5.4.4 Higher QRS duration, higher left ventricular mass and higher left ventricular wall thickness are associated with the development of cardiovascular disease in type 2 diabetes

Considering CVD outcomes in the T2DM cohort, no statistically significant differences were observed in blood pressure, cholesterol, HbA1c, or eGFR between subjects who did and did not develop CVD. These biomarkers are clinical risk factors used in the calculation of SCORE2-Diabetes, a 10-year T2DM-specific cardiovascular risk prediction score developed by a working group of the European Society of Cardiology [380]. Here, the lack of significant differences may be due to a shorter follow-up time or a small sample size (n=98 CVD cases). I found that changes in ECG biomarkers between CVD cases and controls differed in the cohorts with and without T2DM. In the T2DM cohort, 668 subjects (38%) had a heart rate above 70 bpm, a threshold associated with a higher risk of cardiovascular events specifically in T2DM [381], however I did not observe significant changes in heart rate between cases and controls in both cohorts. A higher QRS duration was the only ECG biomarker significantly associated with the development of CVD in the T2DM cohort. In the ACCORD trial, QRS duration was

higher in patients with diabetes and incident heart failure compared to those without heart failure [80]. Additionally, a longer QRS complex has been shown previously to be an independent predictor of cardiovascular events in middle- to older-aged men and is associated with all-cause mortality in T2DM [102, 382]. In contrast, in our control cohort, the Sokolow-Lyon index was the only ECG biomarker showing a significant association with the development of future CVD. This is in line with previous studies linking left ventricular hypertrophy with adverse cardiovascular events [383–387]. The hypothesis behind the hypertrophic CVD phenotype is strengthened by the strong independent association of higher left ventricular mass, LVMI and wall thickness with development of CVD, which was present in both cohorts.

5.4.5 Glycemic control and cardiac remodelling in type 2 diabetes

In T2DM management, HbA1c and blood glucose levels are key clinical targets to achieve effective control of the disease [19, 388]. Findings from the study presented in this Chapter suggest that, beyond binary T2DM status, higher HbA1c levels are associated with a higher ventricular rate, reduced T wave amplitude and lower stroke volume, all possible indicators of early cardiac dysfunction.

As described in Chapter 2 (Section 2.5.4), experimental studies have shown that acute hyperglycemia activates the hexosamine biosynthetic pathway, leading to increased O-GlcNAcylation and autonomous activation of CaMKII, a key mediator of ionic dynamics in T2DM including calcium handling abnormalities [171]. If glucose is well controlled, those acute effects may be avoided, preventing maladaptive CaMKII-mediated phosphorylation of targets such as L-type Ca^{2+} channels and ryanodine receptors. However, even if glucose levels are eventually stabilised in a patient with T2DM, chronic hyperglycemia in early disease may cause sustained or partially irreversible modifications, such as CaMKII oxidation at its Met281/282 site, which can persist even after glucose levels are stabilised [389–391].

Therefore, tight glycaemic control in patients with T2DM as early as possible in disease is not only beneficial for metabolic regulation of diabetes itself, but may also be critical for limiting long-term cardiac remodelling and dysfunction.

5.4.6 Sex, age, body mass index and ethnicity-specific differences

In male and female subgroups, baseline biomarker values differed, which is expected [392–394]. The sex-specific subgroup analysis suggests that trends according to T2DM status were consistent across both sexes. Little to no differences in biomarkers were observed among different age groups, however this result is unlikely to generalise to younger individuals outside the age range of this study (<40 years old), particularly women. Evidence suggests that post-menopausal women, and those in the menopausal transition phase, are at higher risk of CVD compared to those in their reproductive years [395]. This risk is further increased by an earlier age at menopause, Black ethnicity, and T2DM [396]. Fluctuations in hormones and lipid profile that occur during menopause may affect cardiac function and electrophysiology, however this was a challenge to assess in this study, given that women recruited in the UK Biobank were all above 40 years old, i.e. already in the peri- or post-menopause stages. Moreover, I found a notable imbalance of ethnicity groups, with 1733 (97.3%) and 1640 (92.1%) of all participants with and without T2DM, respectively, being from a white ethnic background (Table 5.1). Although this is reasonably representative of the national distribution for different ethnic groups in the UK population at the time of recruitment [229], the small sample size of individuals from a non-white ethnic background (n=189) may have driven the lack of significant differences found between those individuals with (n=141) and without T2DM (n=48) (Appendix B, Tables B.6 and B.7). A larger, more ethnically diverse cohort spanning a wider age range would be beneficial to investigate age- and ethnicity-related differences, especially in women.

5.4.7 Strengths and limitations

The size of the study sample used strengthens the credibility of the study's results. Thanks to the range of data available in the UK Biobank, I was able to characterise our cohorts in depth by capturing detailed demographical and clinical data of all participants studied. These data were used to match the exposure and control groups, and were included as covariates in the regression analyses, thus accounting for any confounding effects on

biomarker outcomes. The UK Biobank's robustly validated and systematic data recording protocol ensured that data collection bias was minimised [229].

This study also has several limitations. Firstly, statistical significance does not necessarily equate to clinical relevance. Small changes in certain biomarkers may fall within the range of natural population variation. Although a shift in the population mean does not imply consistent changes at the individual level, it does indicate that some individuals have experienced changes. This is particularly relevant to subgroup analyses involving smaller sample sizes, which may result in imprecise estimates due to random error. This applies specifically to groups with CVD outcomes (n=98 of which HF cases n=15 in T2DM cohort, n=13 in control), limiting my ability to conduct rigorous outcome analyses, and non-white ethnicity (n=289, of which n=48 without T2DM). Regardless of sample size, I strived to interpret all statistically significant results cautiously and within the context of existing knowledge established by previously published research. CMR-derived global longitudinal strain (GLS) was unfortunately not included in this Chapter as the UK Biobank project used to carry out these analyses was different from the first Chapter and did not contain this data field. The broad UK Biobank medication category used in this study does not contain information on specific compounds, dose, or treatment duration. This limits our ability to assess the impact of common drugs with a notable effect on the heart, such as beta-blockers or SGLT2 inhibitors, on the study's results, which may cause the effect of T2DM to be underestimated in ECG and possibly CMR-derived biomarkers. This could be explored in future work, considering more granular information on medication. Another limitation is the lack of widespread UK Biobank linkage to primary care records. As of 2024, only 45% of the UK Biobank cohort was linked to these records for general research purposes [25]. This directly impacts cohort size and composition, especially for chronic conditions like diabetes which tend to be diagnosed and recorded within a primary care setting as opposed to hospital admissions. Linkage for the entire UK Biobank cohort would increase statistical power and robustness of future population studies. I believe that matching participants by BMI was adequate for the reasons previously stated in the Methods section, however I acknowledge that waist-to-hip ratio is a better indicator of general health and mortality

[397]. Finally, I recognise that including multiple distinct conditions within the broad umbrella of CVD may dissolve opposing trends in certain markers. This could be tackled by focusing on a single condition, albeit with small sample sizes.

5.5 Conclusion

This study examined multi-modal cardiac biomarkers of a large cohort of individuals with T2DM, a well-established high-risk factor for the development of HFpEF and cardiac arrhythmias. Despite a lack of diagnosis of cardiovascular disease in all subjects at the time of assessment, T2DM was associated with a higher resting ventricular rate and an altered ECG including a prolonged QTc interval and lower T wave amplitude, compared to controls matched by sex, age and BMI. ECG changes were accompanied by an independent association of T2DM with a lower stroke volume and thicker ventricular wall. My results provide strong evidence supporting current hypotheses on diabetes-induced pathophysiological alterations in the heart, specifically diabetic myocardial disorder and the development of a pro-arrhythmic substrate. Subject to further validation, these findings support the importance of ECG and cardiac imaging to identify subclinical cardiac abnormalities in patients with T2DM, ultimately aiming to improve cardiovascular risk management in this high-risk population and preventing the development of more severe cardiac complications.

While the results from this Chapter suggest the presence of early cardiac involvement in T2DM even in the absence of overt CVD, the mechanisms driving these changes remain unexplained. Chapter 6 addresses this by using human-based cardiac modelling and simulation to quantify the contributions of multi-scale cardiac remodelling to these changes, offering complementary mechanistic insights beyond what population-level, data-driven methods alone can provide.

It does not appear that the physician studies (even) health in the abstract: he studies the health of the human being, or rather of some particular human being, for it is individuals he has to cure.

— Aristotle, Nicomachean Ethics (I. vi. 14-16)

6

Investigating the Effects of Adverse Cardiac Remodelling on Cardiac Electrophysiology in Type 2 Diabetes Using Human-Based Multi-Scale Computational Modelling and Simulation

Building on the cohort-level findings of early cardiac involvement in type 2 diabetes presented in the previous chapter, this chapter explores the mechanistic underpinnings of these changes, notably electrophysiological, through human-based cardiac modelling and simulation. Using a multi-scale in silico framework deployed on selected UK Biobank cases, I explore how ionic, structural, and electrophysiological remodelling in type 2 diabetes contribute to ECG abnormalities observed in patients. This mechanistic perspective complements the data-driven analyses of previous chapters and provides a foundation for future in silico investigations into arrhythmic risk in type 2 diabetes and heart failure with preserved ejection fraction.

6.1 Introduction

Sudden cardiac death is considered one of the most common modes of death in heart failure with preserved ejection fraction (HFpEF) patients [398, 399]. Sudden cardiac

death can be due to acute pump failure, or stem from a ventricular arrhythmia, which is common in HFpEF [142, 400]. As discussed in Chapter 2 (Sections 2.4 and 2.5), there are many parallels in adverse cardiac remodelling occurring at different scales in type 2 diabetes mellitus (T2DM) and HFpEF which create a pro-arrhythmic substrate through a number of mechanisms. Among others, ionic level disturbances causing delayed repolarisation increases the risk of early afterdepolarisation, a trigger of ventricular arrhythmias, while increased calcium leakage from the sarcoplasmic reticulum increases the risk of delayed afterdepolarisation, another key arrhythmic trigger.

Thus, gaining a better understanding of the electrophysiological substrate in the early stages of disease prior to overt HFpEF represents a strategic approach to eventually determine which settings are most pro-arrhythmic and therefore which patients are at highest risk of developing potentially lethal arrythmias, notably in patients with T2DM where adverse electrophysiological remodelling is of particular clinical relevance. As seen in Chapter 2 (Section 2.5.1), most recent European Society of Cardiology clinical guidelines do not contain diabetes-specific protocols for sudden cardiac death screening; rather, they rely heavily on underlying structural cardiac conditions. However, not all patients with diabetes have established underlying structural heart disease, and clinical characteristics such as sex and body mass index (BMI) are not considered [19, 104].

Building on the population-level investigations presented in previous chapters, this chapter focuses on T2DM through a mechanistic lens. In Chapters 4 and 5, I employed computational statistics and machine learning to quantify the relationships between multi-modal biomarkers, cardiometabolic phenotypes, and stages of disease from T2DM without overt cardiovascular disease to patients with more advanced HFpEF. While I observed relevant early-stage cardiac changes in T2DM, notably QTc prolongation, QRS duration reduction and T wave flattening, it is unclear whether these reflect intrinsic ionic remodelling, differences in myocardial tissue conduction velocity, anatomical variation, or a combination of these factors. Furthermore, due to the number and complexity of contributing factors at play, the relative contribution of these different aspects of cardiac remodelling leading to pro-arrhythmic risk in T2DM remains unclear and is a recognised need in the field [129]. While data-driven machine learning-enabled approaches are useful

to obtain insights at population level, the underlying pathophysiological processes driving these changes remain unexplained. Instead, mechanistic studies are required. In contrast to heart failure, for which computational models of cellular electrophysiology and mechanics already exist [262, 401, 402] and whole-organ modelling is better established [403–405], there is a comparative lack of well-developed modelling frameworks to simulate the specific effects of T2DM on cardiac electrophysiology across different biological scales.

This chapter aims to address these gaps through a detailed mechanistic investigation of the electrophysiological impact of T2DM-driven cardiac remodelling. Specifically, the objective of the study is to build a framework capable of examining the relative contributions of cardiac ionic remodelling, cardiac anatomy, and organ-level electrophysiological parameters on ECG abnormalities in T2DM. I hypothesise that the extent and pattern of ECG changes are modulated by clinical characteristics such as T2DM severity, body mass index, and sex. Specifically, T2DM severity influences ionic-level electrophysiological remodelling and calcium handling, while variations in body size and sex contribute to differences in cardiac and torso anatomy, tissue conduction properties, and electrical activation dynamics.

To address this hypothesis, I develop a highly controllable, biophysically-detailed in silico modelling and simulation framework that builds upon the multi-scale computational modelling approaches for human cardiac electrophysiology introduced in Chapter 3. The goal of this framework is not to construct personalised digital twins, but rather to systematically investigate how variations in anatomy and electrophysiology influence ECG manifestations of T2DM-related remodelling. Using representative cases selected from the UK Biobank cohort examined in Chapter 5, I integrate in the framework characteristics relevant to heterogeneities in sex, body morphology, and cardiac geometry. Based on experimental and clinical data reviewed in Chapter 2, I model the cardiac effects of T2DM across different scales, from single-cell simulations of ionic remodelling to 3D organ-level electrophysiology simulations and ECG reconstruction.

This multi-scale approach allows to dissect the contribution of each component of remodelling under controlled conditions and investigate their combined effects on ECG biomarkers. Ultimately, this work aims to pave the way toward a better understanding of

the arrhythmogenic substrate in T2DM and eventually HFpEF, supporting patient risk stratification and personalised therapeutic strategies in this high-risk clinical population.

6.2 Methods

In a similar way to previous computational studies presented in Chapter 3, I developed and evaluated a multi-scale framework to model and simulate the effects of T2DM on human cardiac electrophysiology. To summarise, I reconstructed patient-specific, virtual, physiologically accurate 3D representations of the heart, the torso, and ECG electrode positions by adapting and applying previously developed and validated pipelines to selected participant data from the UK Biobank [290, 406–408]. I then simulated electrophysiological behaviour at single-cell level using the human ventricular myocyte ToR-ORd model [26], which I adapted to T2DM. To simulate electrical propagation at tissue level, I used MonoAlg3D, an open-source graphics processing unit (GPU)-enabled software [270] that solves the monodomain model introduced in Chapter 3, Section 3.4.3. The baseline ToR-ORd cellular model was already implemented in MonoAlg3D by the authors of the original MonoAlg3D publication and their collaborators [270]. To reflect modifications in T2DM, I implemented my own adaptation of the cellular model with T2DM-specific adaptations in MonoAlg3D. I then used a novel "healthy twin" approach to calibrate certain model parameters affecting the ECG, namely position and local activation time of cardiac activation root nodes and tissue conduction speeds, by applying a Bayesian inference method previously developed by Camps et al. [280]. Simulated ECG signals were reconstructed based on the cardiac activation sequences simulated in 3D using the pseudo-bidomain approach [288]. Each specific step of this framework is detailed in this section.

6.2.1 Subject selection

I selected six representative cases from the UK Biobank cohort of participants with T2DM described previously in Chapter 4. I aimed to capture anatomical heterogeneity across a range of relevant factors: sex, stage of T2DM disease, and BMI as a measure

of body weight and composition. Individual stages of selection and corresponding sample sizes are summarised in Figure 6.1.

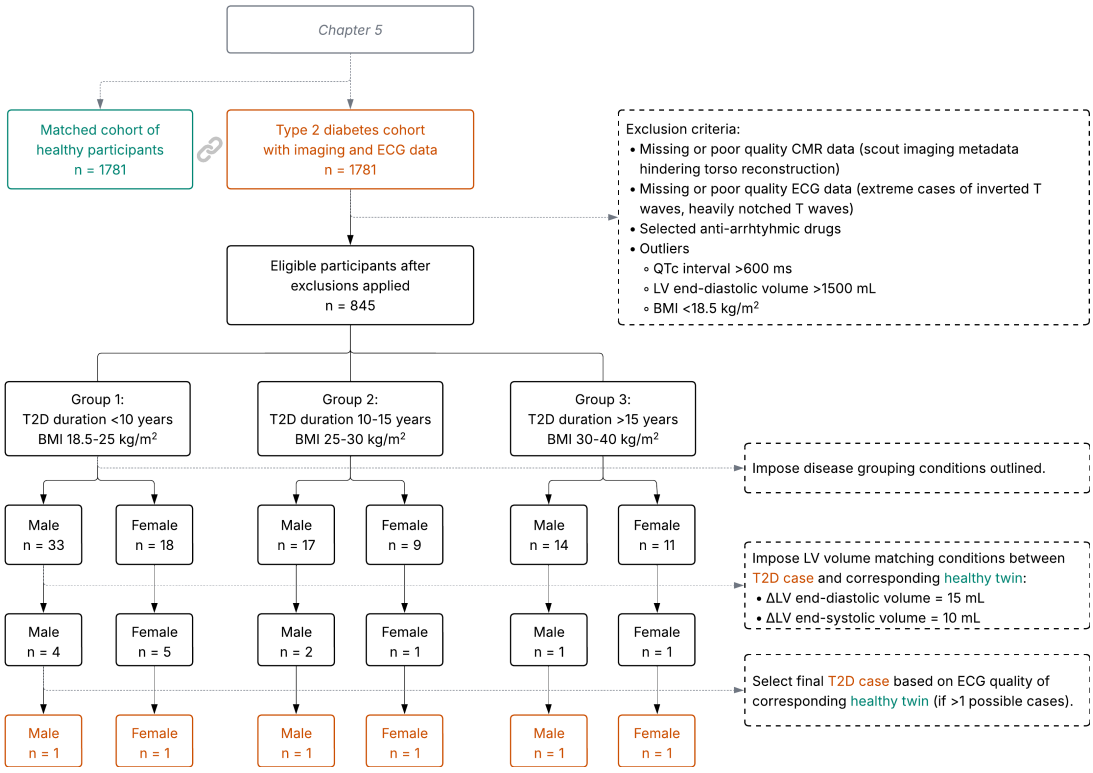

Figure 6.1: Cohort selection flowchart. See Chapter 5 for the selection of the original T2D cohort. Six representative cases were selected for modelling and simulation, following imposed criteria summarised here and in the main text of this Chapter. BMI: body mass index, CMR: cardiac magnetic resonance, LV: left ventricular, T2D: type 2 diabetes.

Exclusion criteria

Cases with missing relevant CMR information were excluded. Patients with anti-arrhythmic or common QTc prolonging medication were also excluded (Appendix C, Table C.1). Anatomical Therapeutic Chemical medication codes were converted from UK Biobank-specific medication codes in field 20003 using a conversion table from previously published work [409]. Patients with very poor ECG quality, assessed visually, were also excluded (too much heterogenous noise and/or baseline wander, unphysiological or fractionated QRS complex). Considering the size of the cohort, this specific criterion was verified in the last stages of the process, immediately before final subject selection.

Stratifying by disease stage

The remaining eligible subjects were first split into male and female group, then further stratified into the following three disease groups which I defined to represent concurrently progressive severity of T2DM and BMI:

- **Group 1:** early T2DM <10 years and normal weight (BMI 18.5–25 kg/m^2)
- **Group 2:** chronic T2DM 10–15 years and overweight (BMI 25–30 kg/m^2)
- **Group 3:** advanced T2DM >15 years and obese (BMI 30–40 kg/m^2)

Previously established anatomical differences across these different factors include larger cavity volume in males compared to females [394], increasing cavity volume with increasing BMI [410, 411], and increasing left ventricular wall thickness in later stages of T2DM [109, 110]. Although the latter two were not reflected in this particular cohort (Figures 6.2 and 6.3), selected cases still provide inherent anatomical variety in the cardiac geometries, to which ionic remodelling is later applied to.

Selecting "control twins"

In Chapter 5, I described the matching process developed to select a control cohort matched on sex, age, and BMI relative to the cohort of T2DM patients, with no history of cardiovascular disease. This effectively allocated a non-diabetic "control twin" for each T2DM case with similar clinical characteristics, including obesity status based on BMI. In the absence of T2DM and associated cardiac remodelling, these healthy twins are assumed to exhibit a normal ECG and can therefore serve as baseline references to calibrate the model before applying disease-related remodelling for each corresponding case. I will return to this assumption in later sections.

The selection of matching criteria (i.e. sex, age, BMI) was chosen primarily to represent cardiac and torso anatomical variability, as these have a direct influence on the ECG. Other clinical criteria such as blood pressure or lipid profile were not considered for the matching here as they do not fit within the scope of the modelling framework considered in this Chapter.

Final pair selection

For each of the three disease groups defined above, I narrowed down a pair of case-twin subjects to those with the best functional/anatomical similarity, i.e. retaining only the case-twin pairs with a left ventricular (LV) end-diastolic (ED) volume difference of ±15 mL and an LV end-systolic (ES) volume difference of ±10 mL. In subgroups with more than 1 selected case, I selected the pair whose twin had the ECG closest to normal (i.e. normal appearance of QRS complex, normal QT interval, least noisy). Final case-twin pairs are visualised in Figure 6.4. ECG chest leads of an illustrative pair are visualised in Appendix C, Figure C.1.

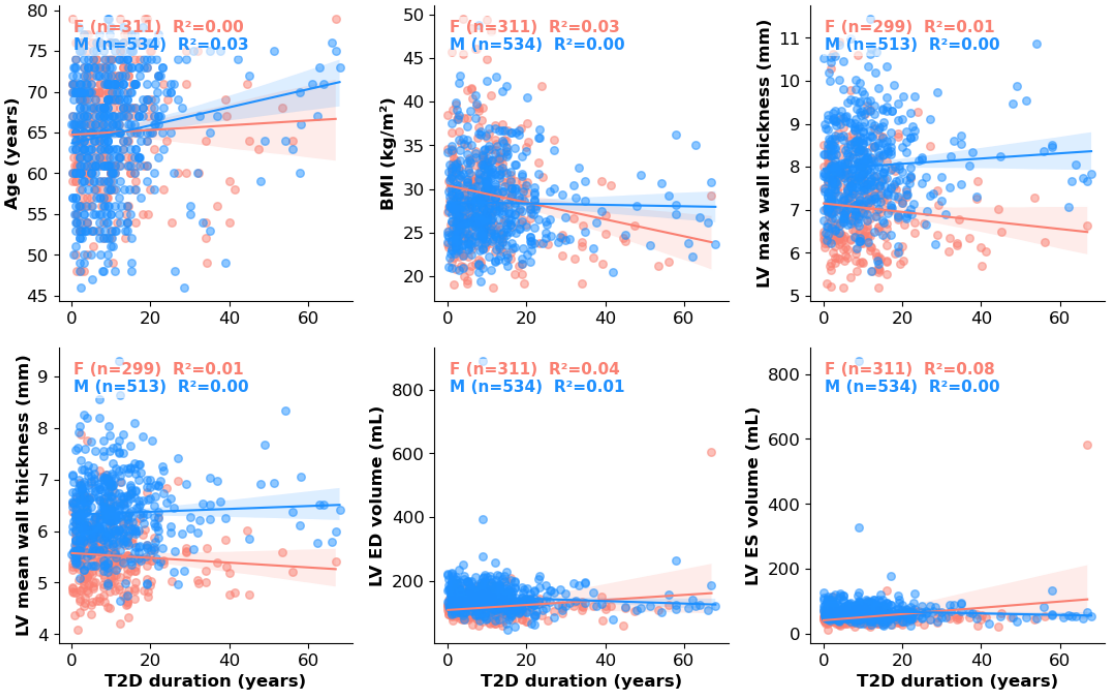

Figure 6.2: Relationship between type 2 diabetes duration and clinical characteristics of subjects in the UK Biobank T2DM cohort described in Chapter 4. BMI: body mass index, CMR: cardiac magnetic resonance, ED: end-diastolic, ES: end-systolic, LV: left ventricular, T2D: type 2 diabetes.

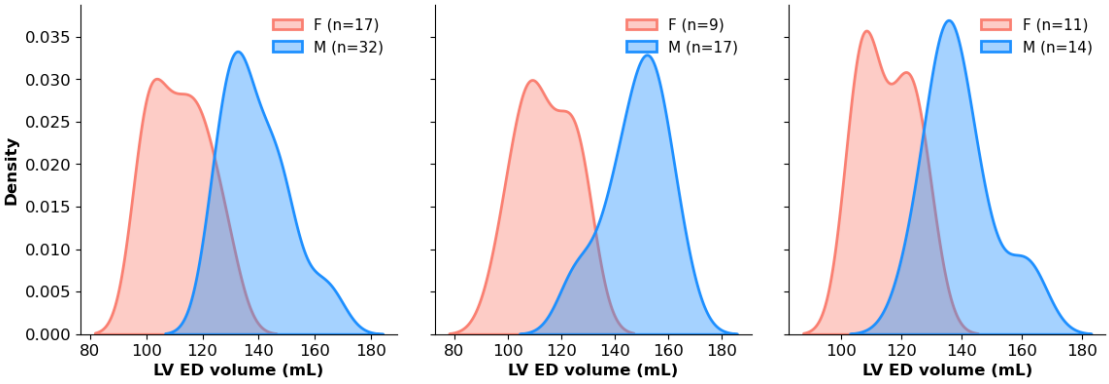

Figure 6.3: LVED volume visualised for men and women stratified by disease severity. BMI: body mass index, ED: end-diastolic, LV: left ventricular, T2D: type 2 diabetes.

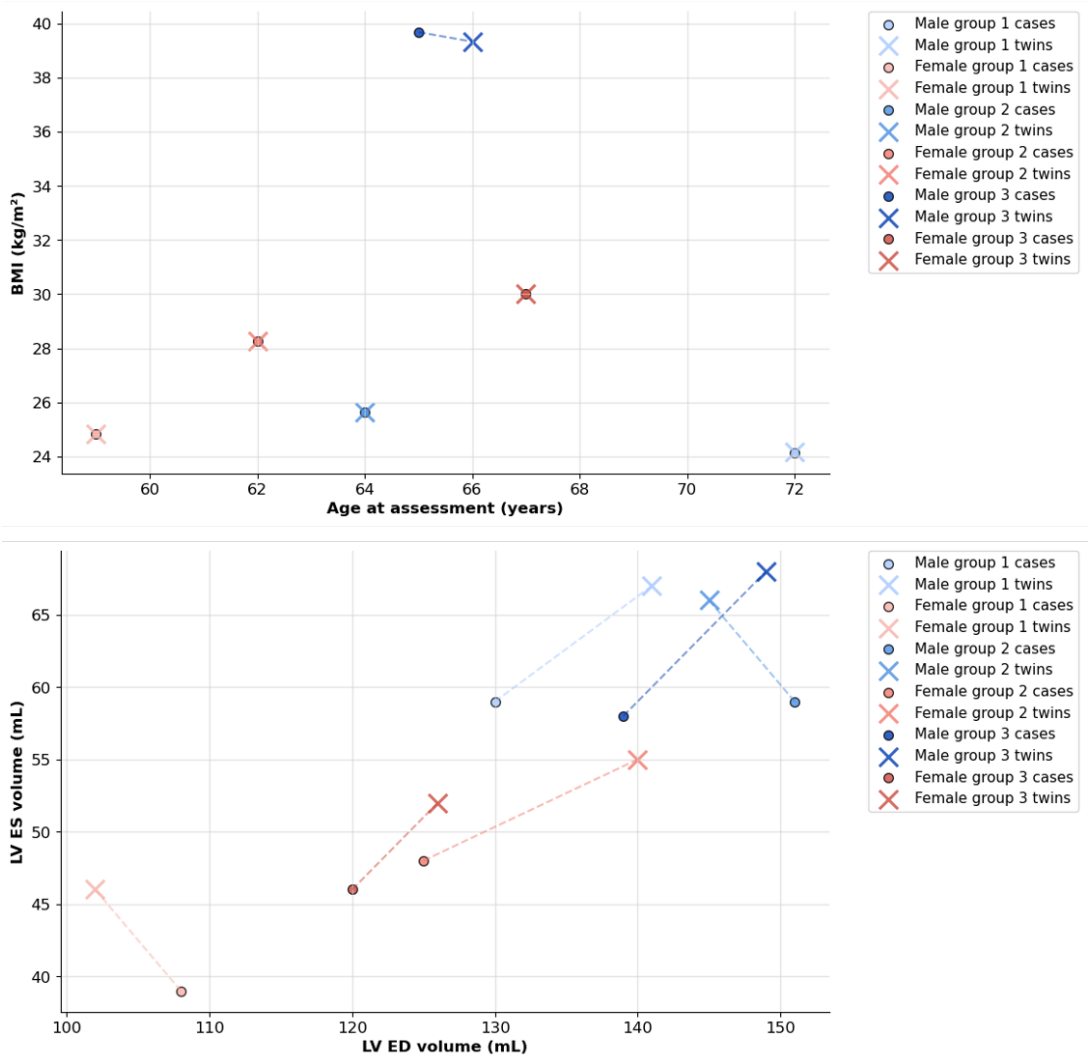

Figure 6.4: Clinical characteristics of pairs of selected type 2 diabetes cases with their corresponding "control twin". Top panel: Age and BMI similarity between eligible case-twin pairs, visualised for men and women in the three study groups of increasing T2D duration and BMI. Bottom panel: LVED and LVES volume as a proxy of cavity size and BMI similarity between eligible case-twin pairs, visualised for men and women in the three study groups of increasing T2D duration and BMI. BMI: body mass index, F: female, LVED: left ventricular end-diastolic, LVES: left ventricular end-systolic, M: male, T2D: type 2 diabetes.

6.2.2 Virtual torso and cardiac anatomy reconstruction

For each selected case with T2DM, a patient-specific 3D biventricular cardiac anatomy mesh was reconstructed from the patient's 2D cardiac magnetic resonance image slices, and embedded in a reconstructed torso with virtual ECG electrode positions based on the same images. To achieve this, I adapted, merged and applied previously validated pipelines [290, 406–408]. The key algorithmic components of this process, and relevant hardware and software requirements, are detailed in Appendix B (Section C.1.1).

6.2.3 Parameter calibration: tissue conduction speeds and early electrical activation sites

The ECG morphology depends not only on cardiac and torso anatomy, and underlying ionic-level electrophysiological behaviour, but also on the position and initial activation time of the sites of early electrical activation sites or "root nodes" in the heart, as well as tissue conduction speeds in the sheet (transmural), fibre (longitudinal) and normal directions. These parameters determine how the electrical signal propagates through the ventricles, thus influencing the final simulated activation time map and subsequently the simulated ECG.

To ensure that these parameters are well-calibrated to each specific case, I used a sequential Monte Carlo approximate Bayesian computation inference approach to estimate personalised root node positions, local activation times and tissue conduction speeds given a patient's cardiac anatomy and ECG QRS complex, following the method described in [280]. Briefly, the algorithm selects populations of candidate root node locations and conduction speeds using a Latin Hypercube Sampling approach and, for each parameter set in the generated population, simulates activation time maps and a pseudo-ECG using a fast Eikonal-based method (see Chapter 3, Section 3.4.3). The simulated ECGs are compared to the clinical ECG of the presented case, and this process is repeated iteratively until the parameter set with the best match is selected.

For each case-twin pair, I used the T2DM patient's anatomy in tetrahedral mesh form together with their control twin's ECG QRS complex. As previously described, the control twin serves as a baseline reference representing normal electrophysiology

before applying diabetes-related remodelling. After pre-processing the UK Biobank ECGs including up-sampling from 500Hz to 1000Hz, baseline wander removal and signal smoothing, I applied the algorithm described above to obtain root node positions, local activation times and anisotropic tissue conduction speeds in "healthy baseline" conditions. More specifically, conduction speeds CV_l in the fibre/longitudinal (l) and CV_n in the normal (n) directions were set according to Taggart et al. [412], while conduction speed CV_t in the sheet/transmural (t) direction was estimated using the Bayesian inference algorithm described above, following the original approach in [280]. For each case, I then computed corresponding anisotropic tissue conductivities σ(l,t,n) by "tuning" these to respective conduction speeds, by simulating electrical activation propagation in a short one-dimensional cable of equal resolution to the mesh, i.e. 400µm. Together, these estimated parameters are used to generate a baseline ECG for each case before applying ionic-level remodelling, acting as a control simulation which further simulations can be compared to.

6.2.4 Parameter calibration: fast-conducting endocardial layer

Alongside calibrating the positions of early activation sites and the conduction speeds across the bulk of the tissue, it is important to account for the presence of a fast-conducting endocardial layer. This thin layer plays a key role in helping the activation spread quickly across the ventricles, effectively modelling the final part of the Purkinje system's influence on electrical signal propagation.

In the 3D anatomical mesh, each element is tagged according to its location, with those lying within the endocardial layer tagged accordingly. These elements are then assigned higher conduction speeds by applying a scaling factor to the baseline tissue conductivity during the simulation. This allows the electrical signal to travel faster through this layer than in the surrounding myocardium. The value of this scaling factor needs to be tuned for each specific case as it has a significant impact on the activation pattern and the resulting ECG, notably the QRS complex. Thus, the QRS complex of the ECG simulated by MonoAlg3D needs to match the ECG simulated by the inference as closely as possible, i.e. the one used to estimate the tissue conduction speeds and root node positions and local activation times. To achieve this, I carried out the following steps, for each case:

1. **Simulate ECGs with distinct scaling factor values**. Consider candidate scaling factor values as integers between 1 and 6 inclusive (higher values would be unphysiological). For each candidate scaling factor, perform a MonoAlg3D biventricular simulation using the patient-specific mesh, ECG electrode positions, calibrated tissue conductivities, and calibrated root node positions and local activation times to obtain a simulated 12-lead simulated ECG.

2. **QRS complex detection** from the simulated ECG signal using an algorithm that I designed, tested and implemented for my cases (see Appendix C, Section C.1.2).

3. **Similarity index calculated relative to the inference-simulated ECG**. I compared all six sets of QRS complexes from the ECGs simulated by MonoAlg3D with the different scaling factors, to the QRS complexes from corresponding ECG leads in the inference-simulated ECG. For each lead pair, I computed Pearson's correlation coefficient as a measure of signal morphology similarity. The average coefficient was computed across all leads and the signal with the best average coefficient was selected, determining the value of the corresponding fast endocardial layer scaling factor to use.

6.2.5 Modelling type 2 diabetes ionic-level remodelling

Chapter 2 (Section 2.5.4) provided extensive background on available experimental data on cardiac ionic current and channel remodelling in T2DM. Based on this data, I propose three distinct models of T2DM ionic remodelling as identified in the literature, by adapting the ToR-ORd model of cardiac cellular electrophysiology. Relevant experimental findings in human data (and supporting animal data) considered in the ToR-ORd model adaptation are recapitulated and visualised in Figure 6.5.

Table 6.1 presents the scaling factors applied to the baseline model to create the three diabetic models. In the first model, numerical values for scaled current conductances (G_X) and RyR activity (J_{rel}) are mostly based on human mRNA data from Ashrafi et al. [161], supported by animal data, while SERCA pump activity (J_{up}) is based on data from [163, 168]. More detail on these data can be found in Chapter 2. I reduced the

extent of I_{Kr} decrease and I_{NaCa} increase compared to [161]; using the exact numbers from that study led to an exaggerated phenotype with an action potential too long to be physiological. The second model represents chronic CaMKII activation by increasing the rate of phosphorylation of CaMKII (α_{CaMK}). The third model also exhibits increased α_{CaMK} as well as additional changes in targets phosphorylated by CaMKII, namely a further decrease in J_{up} and a further increase in the late sodium current (G_{NaL}).

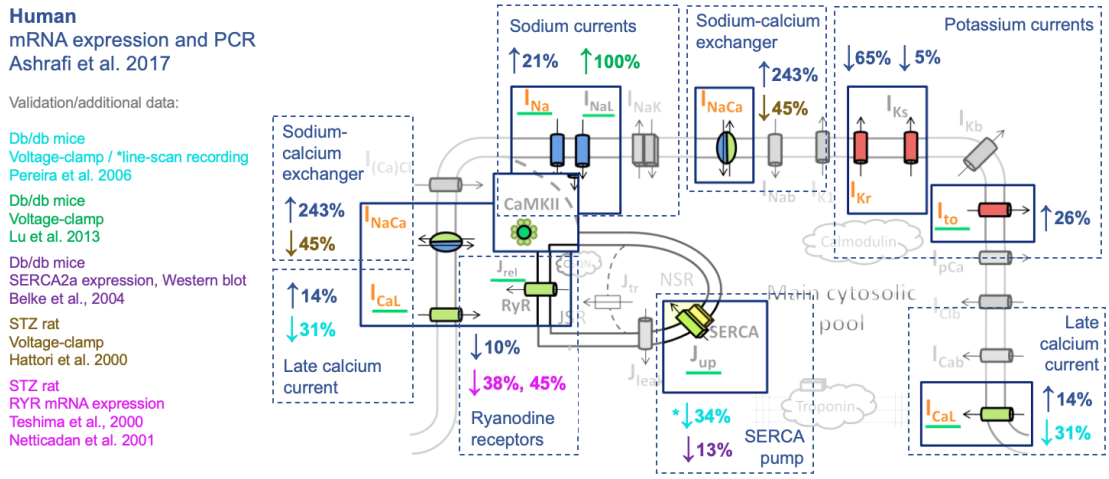

Figure 6.5: Diagram of the ToR-ORd model with relevant ionic channels and exchangers, and related currents and fluxes, adapted to reflect experimental findings of remodelling in T2DM.The colour coding relates ionic changes observed to respective studies, including experimental model and recording methods [161, 163–168]. Refer to Chapter 2 for more details. PCR: polymerase chain reaction, STZ: streptozotocin, RYR: ryanodine receptor, T2DM: type 2 diabetes mellitus.

6.2.6 ECG reconstruction and analysis

ECG signals were computed using the pseudo-bidomain ECG reconstruction approach described in Chapter 3. Personalised ECG lead positions were included for each case including four limb leads and six chest leads. MonoAlg3D generates 10 signals, one for each virtual lead location. These can then be used to reconstruct the signals into the standard 12-lead ECG format using the equations presented in Chapter 3 (Section 2.3.1). I processed and analysed the simulated ECGs using the automatic delineation tool that I designed for this purpose (see Appendix C, Section C.1.2).

Table 6.1: Proposed cardiac ionic remodelling scenarios in type 2 diabetes (T2D), reflecting disease progression, notably regarding CaMKII activation and its phosphorylating effects. Asterisk highlights scaling factor changes in the different models.

Scaling factor	Early stage	Mid stage	Late stage
GNaL	1.21	1.21	2.00*
GNa	1.21	1.21	1.21
GCaL	1.14	1.14	1.14
Gto,f	1.26	1.26	1.26
GKr	0.70	0.70	0.70
GKs	0.95	0.95	0.95
GNaCa	1.50	1.50	1.50
Jrel	0.90	0.90	0.90
Jup	0.80	0.80	0.65*
α_{CaMKII}	(1)	1.50*	1.50*

6.2.7 Model calibration and validation

I aimed to follow a similar approach previously adopted in [413], inspired by the American Society of Mechanical Engineers V&V 40 strategy, for model calibration, verification and validation. This framework provides a structured approach to assess the credibility of computational models through systematic verification and validation, ensuring their reliability and relevance for intended applications.

Question of interest

How can human-based multi-scale electrophysiological simulations reproduce and explain the cardiac electrophysiological phenotype of patients with T2DM?

Context of use

Human-based multi-scale electrophysiological biventricular simulations enable in silico investigations to elucidate pathophysiological mechanisms, patient arrhythmic risk stratification, and drug safety and efficacy evaluation. Here, this study focuses on evaluating this framework under T2DM conditions, specifically to investigate the relative role of cardiac ionic remodelling vs. cardiac anatomy and whole-organ

electrophysiological properties, and lay the foundations for future in silico studies of the heart of patients with T2DM.

Calibration and validation strategy

The human ventricular cellular ToR-ORd model was calibrated using an extensive set of experimental data, and independently evaluated on separate data [26]. I compared simulated outputs for the T2DM-adapted version of the cellular model to experimental action potential biomarkers reported in the literature [23, 164]. I also verified that this model produced similar outputs using different implementations, namely in MATLAB version 2023b which I used to develop and test the model at single-cell level and to generate steady state outputs, compared to MonoAlg3D where I ran single-cell simulations using the same model before scaling to whole-organ. The MonoAlg3D solver has been verified in a previous study using tissue slab and whole-ventricular benchmark tests [270].

Because of this study's research question and mechanistic nature, I did not consider the real clinical ECG of each case as the "ground truth" for validation, considering that the goal of the study is not to replicate each case's ECG but rather to investigate the relative effect of distinct remodelling factors on ECG morphology in T2DM. To achieve this, for each selected UK Biobank T2DM case, each with its own estimated tissue conduction velocities and root node properties, I simulated an ECG corresponding to two scenarios: 1) baseline "control" without T2DM ionic remodelling and 2) T2DM stage-specific ionic remodelling. I then examined changes arising in the simulated ECGs in scenario 2) compared to 1) to quantify the specific effect of ionic remodelling in each case.

6.2.8 Simulation protocol and software

Biventricular 3D simulations were performed using the MonoAlg3D monodomain solver software. I ran the cellular electrophysiology model for 500 beats in MATLAB to achieve a steady state initial condition for the cells. I incorporated these initial conditions into the MonoAlg3D implementation of the cellular model. I computed and tuned tissue conductivities in the sheet (transmural), fibre (longitudinal) and normal directions based on estimated tissue conduction speeds (Section 6.2.3). For each case, the root nodes

were stimulated at their estimated local activation time (Section 6.2.3), once for the first beat and then 800ms later for the second beat, to simulate the ECG in sinus rhythm with a basic cycle length of 800ms, equivalent to a heart rate of 75 beats per minute. The second beat was considered for analysis.

Simulations were run using a GPU implementation of MonoAlg3D on GPUheart, an Intel(R) Xeon(R) Silver 4214 CPU 2.20GHz system equipped with NVIDIA GPUs. The execution time for one simulation of 2 beats in sinus rhythm took about 12-14 hours on this machine.

6.3 Results

6.3.1 Subject characteristics and generated patient-specific anatomical meshes

For privacy-preserving reasons, the UK Biobank strongly advises against the reporting of individual participant level data that could easily lead to patient re-identification. Thus, instead of displaying patient-specific characteristics, I calculated differences between age, BMI and anatomical characteristics of the selected subjects with T2DM, compared to characteristics of their corresponding "control twin", to quantify the extent of similarity, in as defined earlier, achieved between pairs of T2DM patients and healthy twins (Table 6.2).

Table 6.2: Differences in age, BMI and cardiac anatomical characteristics of selected T2D subjects relative to their respective healthy twins. Age and BMI biomarkers from each healthy twin are subtracted from that of the respective T2D case. ΔLVEDV and ΔLVESV indicate the percent of original T2D LVEDV and LVESV represented by the difference between each T2D case and its respective healthy twin. BMI: body mass index, LVEDV: left ventricular end-diastolic volume, LVESV: left ventricular end-systolic volume, T2D: type 2 diabetes.

Disease group	Sex	Δage (years)	ΔBMI (km/m^2)	ΔLVEDV (%)	ΔLVESV (%)
1	M	0	-0.01	-8.5	-13.6
	F	0	0	5.6	-17.9
2	M	0	0	4.0	-11.9
	F	0	0	-12.0	-14.6
3	M	-1	0.35	-7.2	-17.2
	F	0	0	-5.0	-13.0

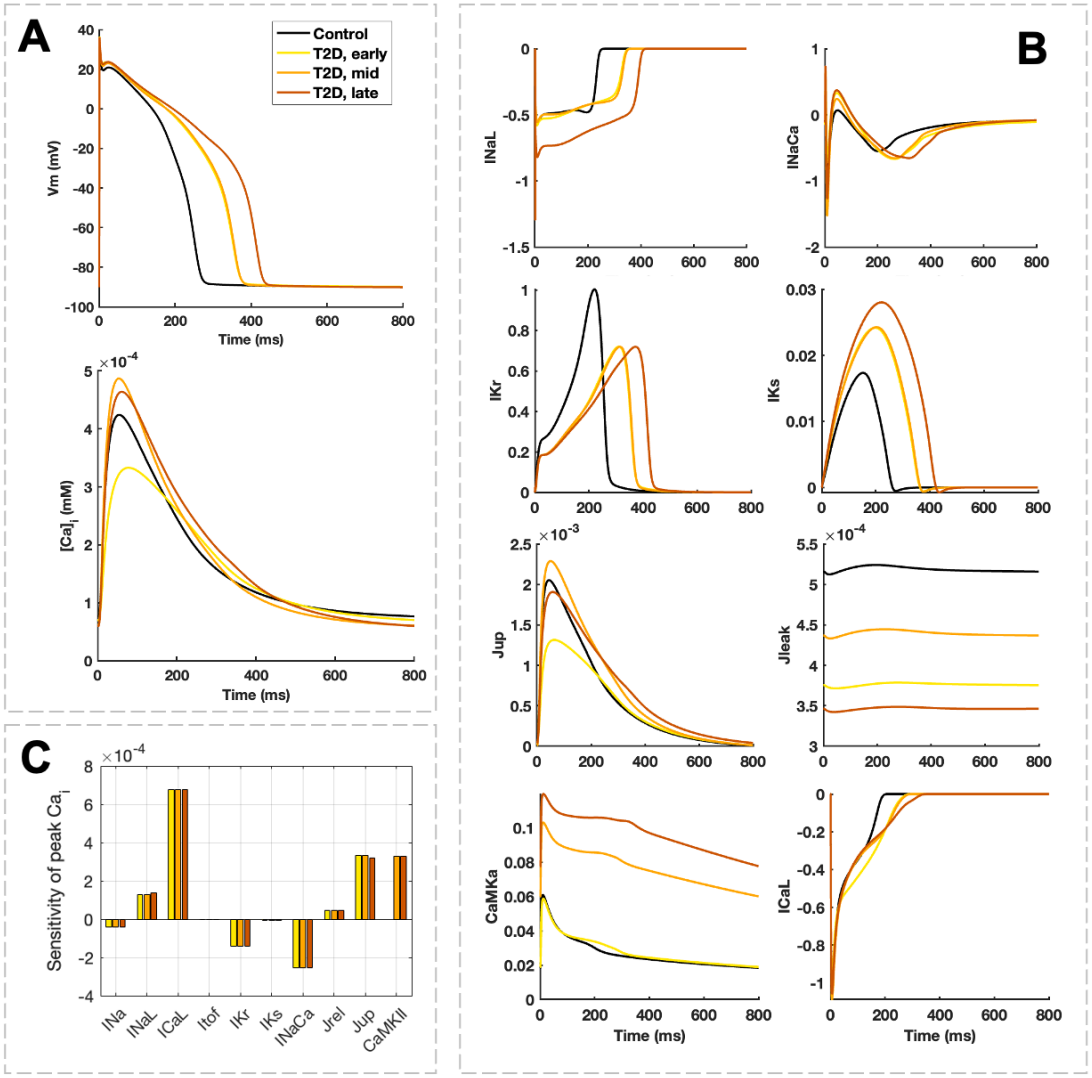

Figure 6.6: Simulated ToR-ORd ventricular cardiomyocyte model outputs for baseline conditions and for the progressive type 2 diabetes ionic remodelling scenarios proposed. Panel A: Action potential and calcium transient. Panel B: individual ionic current and flux traces, and CaMKII activity. I_{Kr} and I_{Ks} are voltage-dependent. Panel C: sensitivity analysis investigating effect of individual ionic currents and CaMKII on peak calcium transient amplitudes. Traces shown are results of an endocardial cell simulation.

6.3.2 Validation, verification, and key findings of the type 2 diabetes ventricular cardiomyocyte model

Figure 6.6 shows the simulated action potential and calcium transient traces generated using the human ventricular cell model adapted to the progressive T2DM ionic remodelling scenarios proposed in Section 6.2.5, as well as simulated individual ionic currents, fluxes, and CaMKII.

Table 6.3: Key action potential biomarkers computed from the simulated action potential traces using the baseline ToR-ORd model and my type 2 diabetes ionic remodelling adaptations of the model. APD: action potential duration, RMP: resting membrane potential, V_{max}: maximum voltage, dV/dt_{max}: maximum rate of change of voltage.

Model	**APD90 (ms)**	**ΔAPD90 from baseline (ms)**	**APD50 (ms)**	**APD40 (ms)**	**V_{max} (mV)**	**RMP (mV)**	**dV/dt_{max} (mV/ms)**
Baseline	259	-	206	183	32	-90	352
T2D, early	360	+101	288	243	36	-90	394
T2D, mid	362	+103	292	248	35	-91	388
T2D, late	420	+161	348	284	35	-91	384

Action potential was prolonged in all three ionic remodelling scenarios; key simulated action potential biomarkers are reported in Table 6.3. Action potential prolongation is consistent with experimental findings of APD90 increase in T2D mice models: Lu et al. reported an average +130ms increase in db/db mice [164] while Hegyi et al. reported a +15ms increase in mouse ventricles under glucose infusion [23].

In addition to action potential prolongation, I found that calcium transients in the T2DM ionic remodelling scenarios were different, suggesting differences in contractile behaviour at the cellular level, as myocardial contractility is directly linked to intracellular Ca^{2+} levels.

Notably, peak calcium transient amplitude $[Ca^{2+}]_i^{max}$ in the baseline model was 4.2×10^{-4} mM, while it was 3.3×10^{-4} mM (-0.9×10^{-4} mM from baseline) in the early-stage T2DM model, 4.9×10^{-4} mM ($+0.7 \times 10^{-4}$ mM) in the mid-stage T2DM model, and 4.6×10^{-4} mM ($+0.4 \times 10^{-4}$ mM) in the late-stage T2DM model.

I hypothesised that these changes were due to CaMKII activation states in the early vs. later T2DM remodelling scenarios. To further examine this, I ran a sensitivity analysis

to investigate the effect of individual ionic currents and CaMKII scaling factors on peak calcium transient amplitude $[Ca^{2+}]_i^{max}$. For each model scaling factor SF, $[Ca^{2+}]_i^{max}$ sensitivity s was computed as follows:

$$s = \frac{[Ca^{2+}]_i^{max} - ([Ca^{2+}]_i^{max})_{baseline}}{SF - 1} \tag{6.1}$$

Results of this analysis suggest that I_{Kr} and I_{NaCa} are the main sources of peak $[Ca^{2+}]_i^{max}$ reduction in the early T2DM ionic remodelling scenario. This decrease aligns with experimental findings in db/db mice, suggesting a decrease of up to -50% in $[Ca^{2+}]_i^{max}$ [163, 168].

However, in my simulations, this effect appears to be counterbalanced in the other two remodelling scenarios (mid-stage and late-stage T2DM) by the increase in CaMKII activation, the key difference between those two scenarios and early-stage T2DM. This is likely due to phosphorylation of ionic targets by CaMKII as explained in Chapter 2. Indeed, CaMKII phosphorylates the Na^+ channel and L-type Ca^{2+} channel, which leads to a direct increase in intracellular Na^+ and Ca^{2+} levels, as well as reduced Ca^{2+} extrusion by the Na^+-Ca^{2+} exchanger.

6.3.3 Calibrated activation properties and fast endocardial layer

Figure 6.7 and Table 6.4 illustrate the results of the calibration of the myocardial tissue activation properties using the inference method described in Section 6.2.3. Between 5 and 7 root nodes were estimated for each case, in heterogenous locations across the endocardium in the left and right ventricles. Local activation times varied between 21ms and 48ms. This reflects the propagation of the electrical signal travelling through the Purkinje fibres descending from the base to the apex of the heart through the septum and back up through the left and right ventricular walls, as well as differences in tissue conduction speeds in orthogonal directions.

The results of the fast endocardial layer calibration step for each case are reported in Table 6.5. Figure 6.8 illustrates the process for one selected case (male sex, disease group 1), highlighting differences in ECG using each fast endocardial layer scaling factor value.

Table 6.4: Conduction speed (CV) and conductivity (σ) values by sex, group, and fibre direction. Anisotropic conduction speeds in the fibre/longitudinal (l) and normal (n) directions are set according to Taggart et al. [412], while conduction speed in the sheet/transmural (t) direction was estimated using the Bayesian inference algorithm described in Section 6.2.3. Tissue conductivities were then tuned to the corresponding conduction speed as described in Section 6.2.3.

Group	Fibre direction	Male		Female	
		CV (cm/s)	σ (µS/µm)	CV (cm/s)	σ (µS/µm)
Group 1	Longitudinal (l)	65	0.260	65	0.260
	Transverse (t)	**38**	**0.120**	**40**	**0.129**
	Normal (n)	48	0.166	48	0.166
Group 2	Longitudinal (l)	65	0.260	65	0.260
	Transverse (t)	**28**	**0.080**	**44**	**0.146**
	Normal (n)	48	0.166	48	0.166
Group 3	Longitudinal (l)	65	0.260	65	0.260
	Transverse (t)	**58**	**0.219**	**60**	**0.230**
	Normal (n)	48	0.166	48	0.166

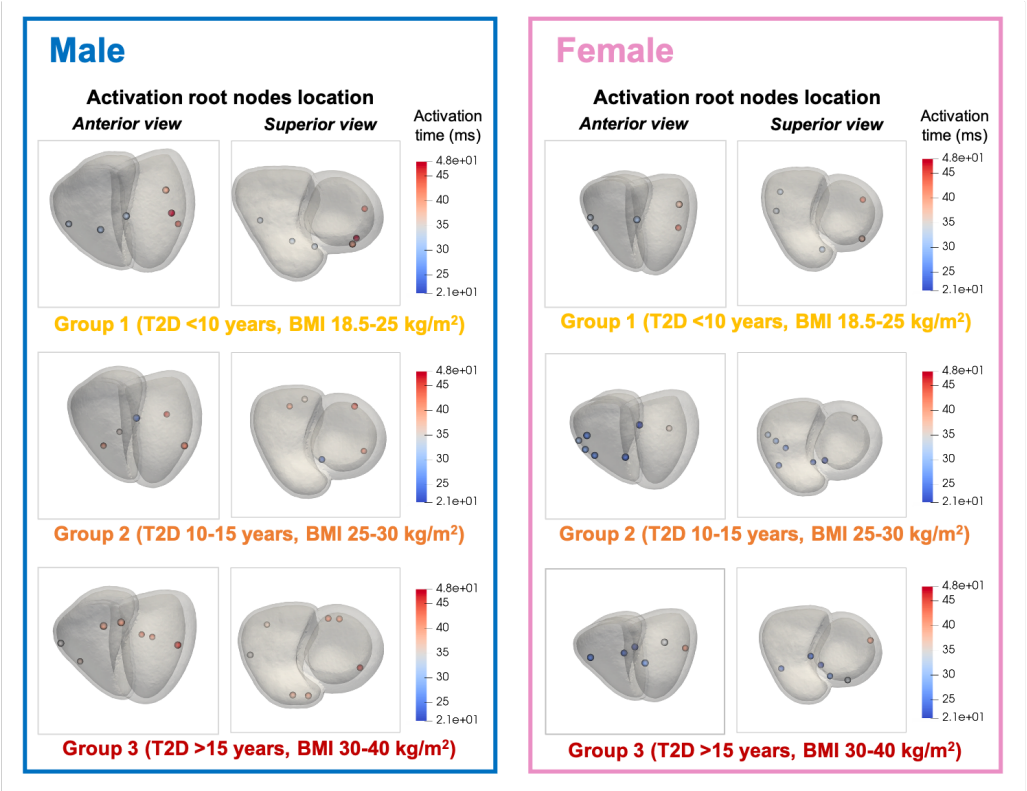

Figure 6.7: Patient-specific estimates of location and time of tissue electrical activation sites in selected type 2 diabetes cases, stratified by sex and disease group. Meshes are visualised to the same scale. Root node positions and local activation times are shown on the corresponding cardiac mesh for each case, from anterior and superior views.

Table 6.5: Pearson's correlation coefficient for fast endocardial layer calibration. Coefficients, reported in the table for each selected case and fast endocardial layer conductivity scaling factor (SF), were calculated between corresponding leads from the ECG simulated using the Bayesian inference algorithm described in Section 6.2.3, and the ECGs simulated in MonoAlg3D with each different scaling factor, respectively. The scaling factor yielding the best match i.e. highest coefficient was selected as the calibrated value and is highlighted in bold for each case.

Disease group	**Sex**	**SF = 1**	**SF = 2**	**SF = 3**	**SF = 4**	**SF = 5**	**SF = 6**	**Selected SF**
1	M	0.36	0.55	**0.64**	0.63	0.54	0.43	SF = 3
	F	0.10	**0.43**	0.39	0.22	0.16	0.10	SF = 2
2	M	0.38	0.40	0.55	**0.58**	0.49	0.33	SF = 4
	F	0.24	0.29	0.31	**0.32**	0.31	0.30	SF = 4
3	M	-0.14	0.05	0.18	0.20	**0.27**	0.15	SF = 5
	F	-0.01	0.34	**0.40**	0.37	0.30	0.20	SF = 3

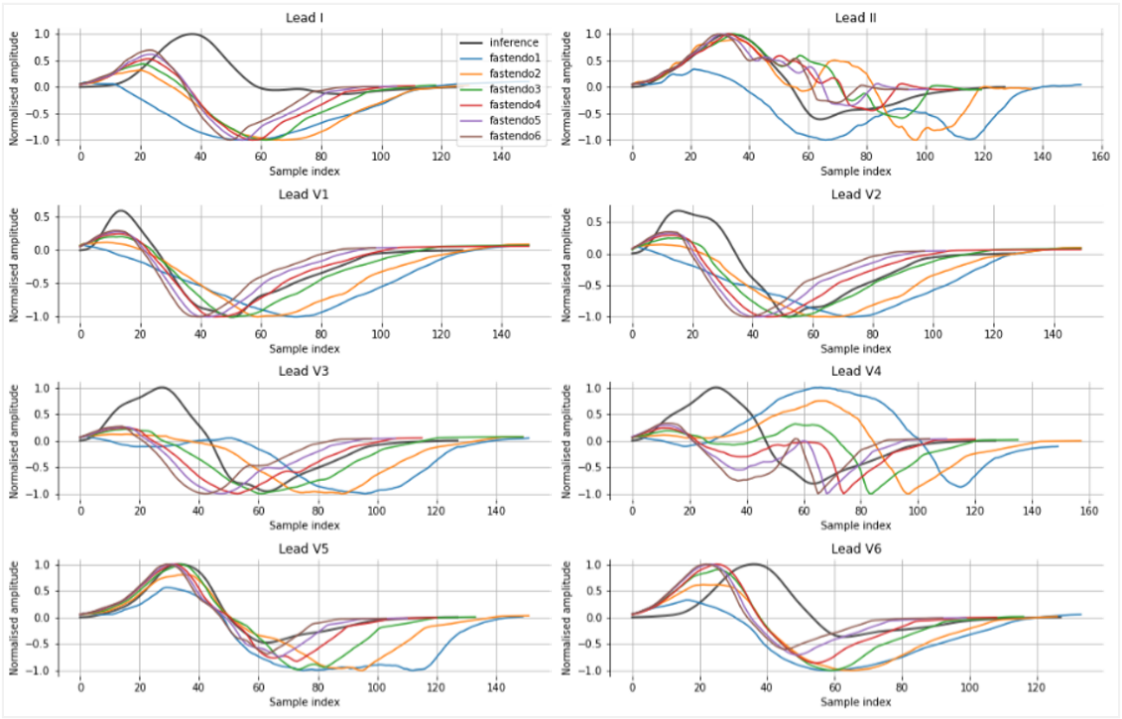

Figure 6.8: Fast endocardial layer calibration step for an example selected case (male sex, disease group 1). Illustrated in the figure are the QRS complex of the ECG generated using the Eikonal-based inference method (black trace) described in Section 6.2.3 along with those of the six ECGs generated using MonoAlg3D with inferred tissue conduction speeds and root node locations, each with a different fast endocardial scaling factor (coloured traces, fastendo1 to fastendo6).

6.3.4 Simulated ECGs in type 2 diabetes using MonoAlg3D

Figures 6.9 and 6.10 present the simulated ECGs for each of the six cases, under baseline "control" conditions and disease stage-specific T2D ionic remodelling. The start and end of the QRS complex, end of the QT interval, and T wave amplitude were automatically identified using the algorithm developed for this purpose (see Appendix C, Section C.1.2; Figure C.2 illustrates the outputs of this algorithm for a single lead per case and scenario). These biomarkers were used to compute the QRS duration, QT interval and T wave amplitude for individual cases (Table 6.6).

Simulated 12-lead ECGs incorporating the T2DM-specific ionic remodelling scenarios revealed stage-dependent alterations in both QRS and QT intervals. The QT interval was markedly prolonged across all T2DM models, consistent with the action potential duration prolongation observed in the cellular simulations presented in Section 6.3.2. In males, QT interval increased from 389.6 ms at baseline to 481.6 ms in Group 1 (+92.0 ms), from 338.4 ms to 465.0 ms in Group 2 (+126.6 ms), and from 333.4 ms to 490.6 ms in Group 3 (+157.2 ms). A similar trend was observed in females, with QT prolongation of +92.4 ms (383.8 to 476.2 ms) in Group 1, +93.5 ms (358.3 to 451.8 ms) in Group 2, and +155.9 ms (374.5 to 530.4 ms) in Group 3. These organ-level changes directly reflect the underlying prolongation of ventricular repolarisation at the single-cell level, driven by reductions in repolarising currents I_{Kr} and I_{Ks}, further exacerbated in later stages of the action potential due to their voltage-dependence, and enhanced I_{NaL} (see Figure 6.6).

Table 6.6: QRS duration, QT interval and T-wave amplitude biomarkers in the simulated ECGs from selected T2D subjects, under baseline control conditions and disease group-specific type 2 diabetes ionic remodelling. Absolute T-wave amplitudes are reported (mV). Values are given as the mean across all leads excluding ECG leads where the QRS and/or QT delineation failed due to small wave amplitudes. The number of ECG leads (max $l = 12$) across which the mean was calculated is indicated for each group. T2D: type 2 diabetes.

Sex	Group (n)	Baseline QRS (ms)	T2D QRS (ms)	Δ QRS (ms)	Baseline QT (ms)	T2D QT (ms)	Δ QT (ms)	Baseline T amp (mV)	T2D T amp (mV)	Δ T amp (mV)
Male	Group 1 (l=10)	129.0	120.5	-8.5	389.6	481.6	+92.0	0.073	0.071	-0.002
Male	Group 2 (l=11)	107.4	100.8	-6.6	338.4	465.0	+126.6	0.003	0.002	-0.001
Male	Group 3 (l=12)	84.4	82.7	-1.7	333.4	490.6	+157.2	0.018	0.016	-0.002
Female	Group 1 (l=12)	125.3	110.5	-14.8	383.8	476.2	+92.4	0.010	0.010	0.000
Female	Group 2 (l=12)	86.3	83.9	-2.3	358.3	451.8	+93.5	0.036	0.035	-0.001
Female	Group 3 (l=11)	119.9	107.9	-12.0	374.5	530.4	+155.9	0.040	0.037	-0.003

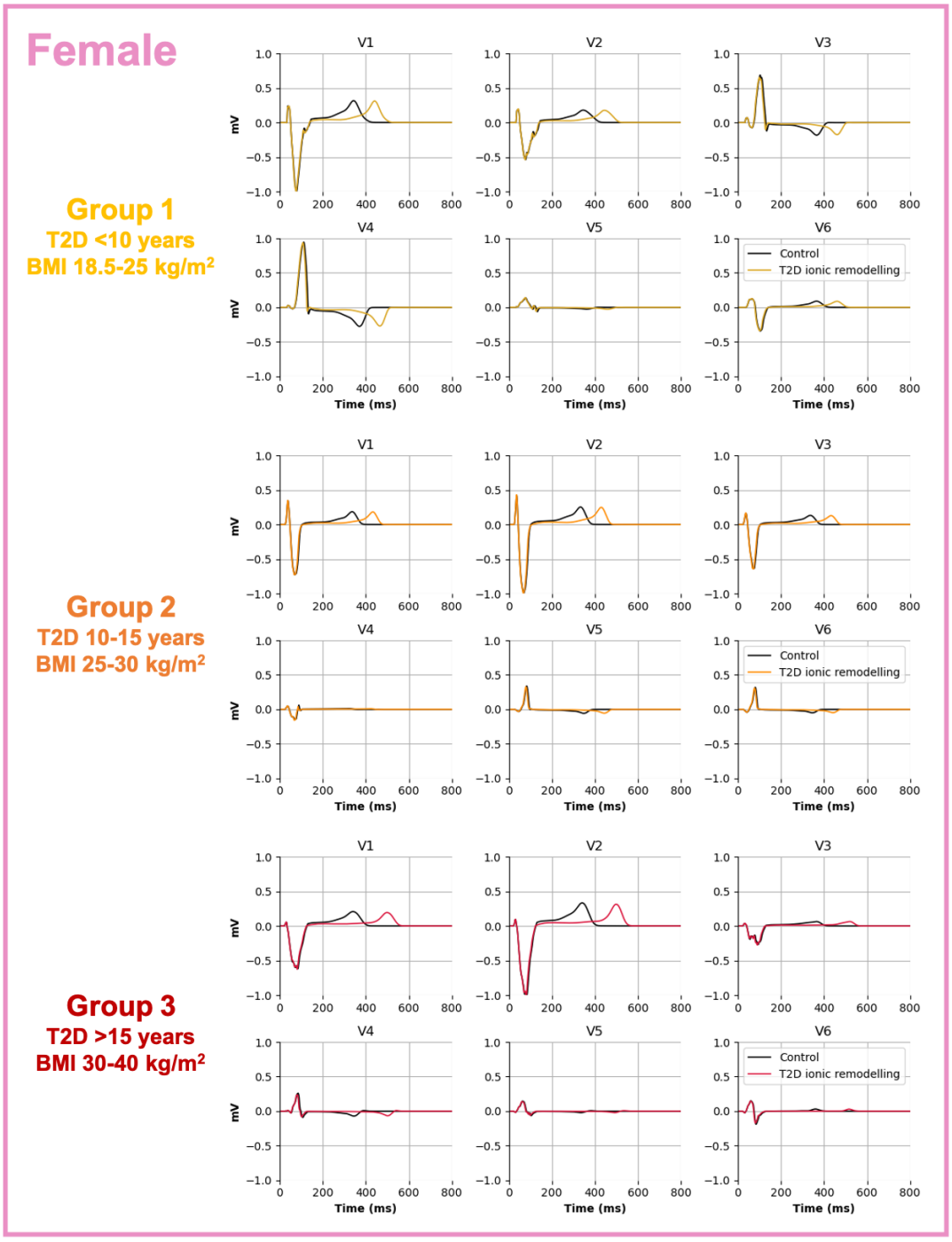

Figure 6.9: Simulated ECGs visualised for the six chest limbs, in the three selected female cases with early- (group 1), mid- (group 2) and late-stage (group 3) type 2 diabetes. For each disease group and sex-specific case, two scenarios were considered, namely 1) baseline control and 2) disease stage-specific type 2 diabetes ionic remodelling.

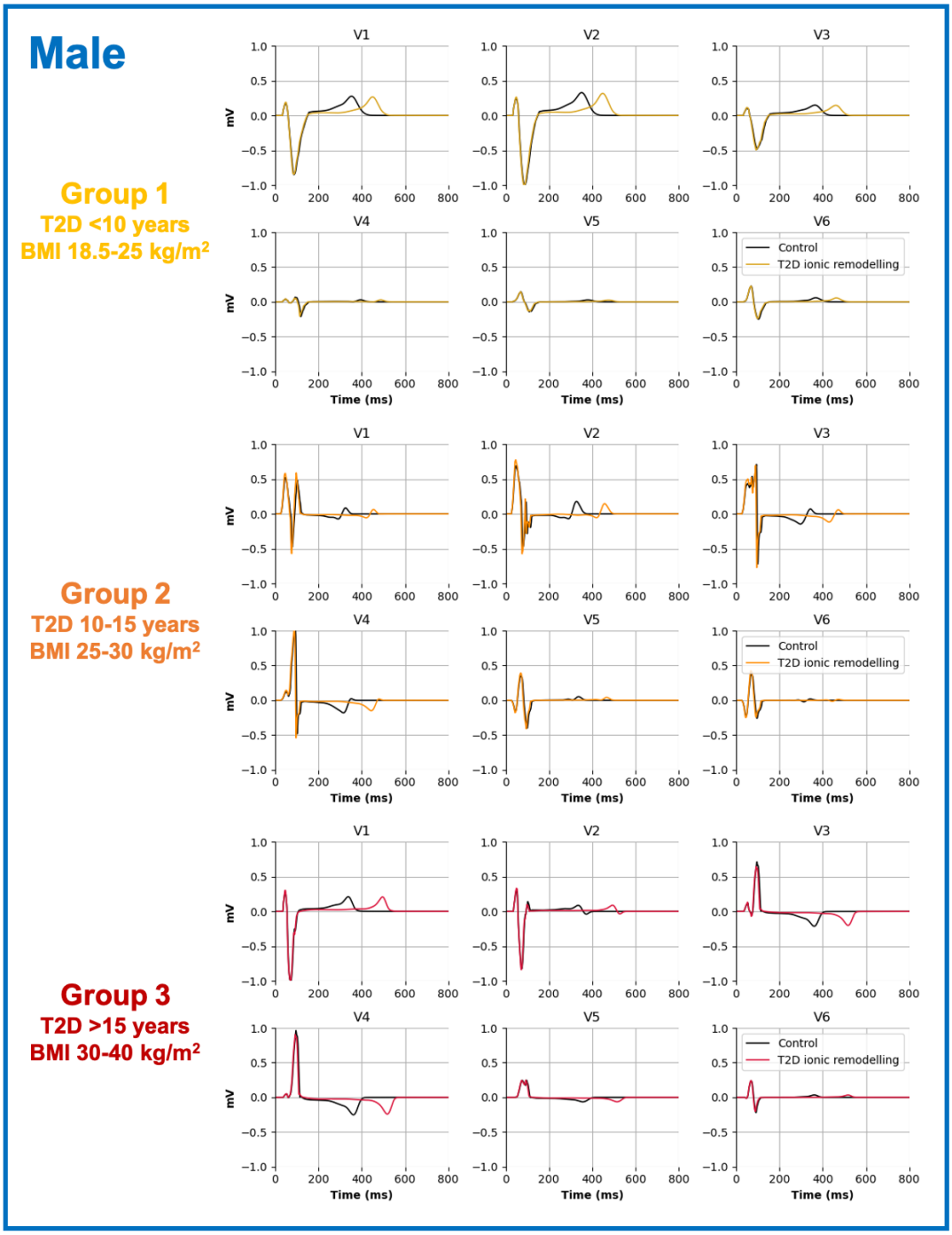

Figure 6.10: Simulated ECGs visualised for the six chest limbs, in the three selected male cases with early- (group 1), mid- (group 2) and late-stage (group 3) type 2 diabetes. For each disease group and sex-specific case, two scenarios were considered, namely 1) baseline control and 2) disease stage-specific type 2 diabetes ionic remodelling.

In contrast, QRS duration was slightly shortened, and to different extents, under T2DM ionic remodelling conditions. This is likely explained by increased I_{Na} in the cellular models, which leads to faster depolarisation and enhanced conduction velocity, resulting in accelerated ventricular activation. In each disease group, under the same T2DM ionic remodelling conditions, slight numerical differences in the extent of QT interval prolongation and QRS reduction arise from differences in torso or cardiac anatomy, tissue conduction velocity, or fast endocardial scaling factor. These findings demonstrate that ionic-level remodelling in T2DM, particularly increased depolarising currents and delayed repolarisation, manifest as detectable alterations in ECG morphology compared to a baseline ionic state given the same organ-level anatomical and electrophysiological (i.e. root node position, local activation time, and tissue conduction velocity) conditions.

6.4 Discussion

In this Chapter, I presented the development of a human-based multi-scale computational modelling and simulation framework for in silico studies of T2DM cardiac electrophysiology, with a focus on the investigation of the relative effects of cardiac ionic remodelling, organ-level electrophysiological parameters and cardiac anatomy on ECG morphology. At ionic level, I showed that T2DM remodelling led to action potential duration prolongation, with more extreme prolongation noted in the model of late-stage T2DM with chronically hyperactivated CaMKII, increased I_{NaL} and reduced SERCA pump flux. The model of early-stage T2DM yielded a decrease in peak calcium transient amplitude stemming from decreased I_{Kr} and increased I_{NaCa}, which appears to be counterbalanced by increased CaMKII in the later-stage T2DM models. This change would have a direct effect on cell-level contractility, reducing excitation-contraction coupling and therefore reducing mechanical function, possibly explaining the reduced stroke volume observed clinically (see Chapter 5). At whole-organ level, simulated ECGs reflect this progression in action potential prolongation through a corresponding increase in QT interval and slight T wave reduction, as well as subtle QRS reduction, recapitulating clinically observed findings (see Chapter 5).

In this section, I discuss the clinical relevance of the main findings from this study, as well as from individual components of this framework nuanced with their respective limitations within this context of use. Throughout, I propose avenues for future improvements and expansion of the framework, and conclude with some limitations, which will later be revisited in Chapter 7.

6.4.1 Cardiac modelling and simulation in type 2 diabetes: research gaps and potential significance

As discussed in Chapter 3, computational cardiac modelling and simulation offers a promising methodological framework to complement pre-clinical experimental studies. Such in silico approaches have been applied to many clinical contexts of use covering both health (e.g. sex-specific differences) and disease (e.g. myocardial ischemia and infarction, atrial fibrillation, heart failure), to investigate multi-scale disease mechanisms and to conduct in silico drug safety and evaluation tests. Although its cardiac effects are becoming increasingly well established, T2DM has not benefitted from the same level of development in the cardiac modelling and simulation field. Clinically, diabetic myocardial disorder remains a highly complex and multi-factorial condition, making it is challenging to study accurately and holistically across different biophysical scales in humans [106]. Clinical guidelines for arrhythmic risk stratification in patients with T2DM remain fairly unspecific [19, 104], likely due to a lack of clinical and pathophysiological understanding of the cardiac changes arising from the condition. Thus, multi-scale modelling and simulation represents a promising methodological approach to study this condition. To the best of my knowledge, the only whole-organ cardiac electrophysiological simulation study in diabetes prior to this study was carried out by Sedova et al., to determine how ventricular repolarisation dispersion and repolarisation gradients across orthogonal directions in the ventricles contribute to the formation of T-wave flattening [414]. However, this study reproduced behaviours observed in a rabbit model of diabetes type 1, not 2, making the context of use presented here novel in terms of clinical area of interest. Furthermore, Sedova et al. modelled ventricular repolarisation a single dipole located in the centre of the ventricles and generating a potential distribution. Recently, Strocchi et al. proposed a multiscale,

whole-heart electromechanics framework for (all-type) diabetes, integrating ventricular and atrial cellular-level models, anatomically detailed whole-heart mechanics and a 0D circulatory system model [415]. They did not explicitly account for diabetes-specific ionic current changes but rather focused on the sensitivity of physiological parameters across scales to identify key determinants of cardiac mechanical dysfunction in diabetes, which, they found, included systemic resistance, venous pressure and atrial stiffness. In contrast, the framework presented in my study enables the simulation of electrical propagation using a truly multi-scale biophysically accurate approach, comprising of individual ionic current representations and tissue-level electrical propagation using the monodomain model, embedded in patient-specific anatomical meshes. This enables relevant mechanisms to be captured and connected across different scales, such as the effect of changes in ionic currents and CaMKII on action potential duration prolongation and calcium transient reduction, and the ability to incorporate cardiac anatomy variability and tissue conduction velocities as modifiable factors of ECG morphology.

6.4.2 The "control twin" approach

To calibrate the model activation properties (root node location, activation time, and tissue conduction speeds), I selected the ECG of the non-diabetic "control twin" corresponding to each T2DM case. This design choice was based on the research question of this study and its fundamentally mechanistic nature. Indeed, the aim was to investigate different types and extents of cardiac remodelling using T2DM patients' cardiac meshes as the anatomical basis of these investigations, rather than, for each case, to fully personalise and optimise the model in order to obtain a single ECG that matches the clinical ECG of each of the six selected UK Biobank T2DM cases. Calibrating the model to produce a "pseudo-control" ECG produces a baseline against which I could then compare ECGs simulated under different remodelling conditions, for the same anatomical case. However, the morphology of the ECG inherently depends, in part, on cardiac anatomy itself, notably heterogeneities in cavity shape and size, and wall thickness. Therefore, the selection of the healthy twin could be further optimised, for example selecting on a wider range of measures of cardiac anatomy similarity rather than age, sex and/or BMI.

6.4.3 Mesh resolution selection and importance of context of use

The spatial resolution of the cardiac anatomical mesh dictates the behaviour of the simulated electrical signal propagation, which determines the accuracy and reliability of the model, depending on its context of use. Here, because the primary outcome we considered is the ECG, we can afford to reduce the mesh resolution (i.e. coarser mesh) to a certain extent in order to provide faster execution times whilst not compromising the degree of accuracy. However, mesh resolution must be carefully considered when modelling more complex phenomena such as ventricular arrhythmias [416]. In a recent in silico study by Berg et al. using the MonoAlg3D solver, authors found that simulated ventricular tachycardia was sustained across mesh resolutions ranging from 250µm to 350µm, with only slight, non-linear differences in the time of ventricular tachycardia onset [417]. They also found that conduction velocity calibration was essential to preserve proper conduction across resolutions. Practically speaking, the selection of an adequate mesh resolution depends on the specific context of use, but these results suggest that a mesh resolution of 350µm is good enough to study ventricular tachycardia in silico as it delivers the best execution time while still providing a physiologically accurate simulation of the phenomenon of interest. These results depend on the solver and numerical methods used, so only apply to MonoAlg3D simulations. If the framework presented in this Chapter were to be extended to study ventricular arrhythmias, the mesh resolution may need to be increased (i.e. by reducing the size of the tetrahedral elements) to capture finer anatomical and electrical detail necessary to reproduce arrhythmic behaviour accurately.

6.4.4 Cardiac ionic remodelling in type 2 diabetes: impaired contractility and arrhythmic risk

In this study, I implemented and validated an adaptation of the ToR-ORd cardiomyocyte model to simulate progressive ionic remodelling in T2DM. Action potential duration was prolonged across all three simulated stages of disease. These results agree with experimental data, where significant APD90 prolongation has been reported in diabetic cardiomyocytes [23, 164]. Importantly, these results also provide a potential mechanistic

explanation for the QTc prolongation observed in the ECG of T2DM patients despite an absence of overt cardiac disease, which we discussed in Chapter 5.

The simulations also revealed distinct calcium transient profiles between early-, mid-, and late-stage T2DM, suggesting stage-dependent alterations in intracellular calcium handling. Notably, I found a reduced peak $[Ca^{2+}]_i$ in early-stage T2DM, potentially explained by reductions in I_{Kr} and I_{NaCa}, identified through sensitivity analysis as key contributors to calcium transient attenuation. This aligns with experimental findings showing depressed calcium transients and contractility [163, 168] (Chapter 2, Section 2.5.4).In contrast, later stages of T2DM showed increased peak $[Ca^{2+}]_i$ amplitude, which I attributed to enhanced CaMKII activation, the primary difference introduced in the mid- and late-stage remodelling scenarios. This finding is consistent with studies showing that hyperglycemia promotes chronic CaMKII activation via O-GlcNAcylation, leading to action potential prolongation and increased $[Ca^{2+}]_i$, (Chapter 2, Section 2.5.4) [170, 171, 174]. These effects may contribute to both early and delayed afterdepolarisations [175, 178], possibly explaining the increased risk of sudden cardiac death in diabetic patients. Overall, these results support the idea that chronic CaMKII hyperactivation in T2DM may play a crucial biphasic role in the heart: initially causing calcium depletion and impaired contractility, and later promoting intracellular calcium overload and arrhythmogenesis.

These simulations also support the hypothesis that ionic remodelling in T2DM shares several features with HFpEF, including altered sodium-calcium exchanger activity, reduced repolarising K^+ currents (I_{Kr}, I_{Ks}) and increased depolarising currents such as I_{NaL} [59]. In heart failure, hyperphosphorylation of the sarcoplasmic reticulum Ca^{2+} release channel, enhanced NCX activity, and reduced SERCA also generate diastolic Ca^{2+} leak and reduced reuptake of Ca^{2+} in the sarcoplasmic reticulum, leading to contractile dysfunction and a higher propensity to cardiac arrhythmias [59]. Thus, considering ionic-level mechanisms only, the electrophysiological substrate studied here presents some analogies to that of heart failure.

However, in a recent cellular electromechanical modelling and simulation study, Zhou et al. demonstrated that failing myocytes with lower diastolic tension and elevated end-diastolic volume were more prone to arrhythmias, suggesting that wall stress and

ventricular filling play a central role in arrhythmia risk under ionic remodelling in heart failure [262]. In contrast, in the previous chapter, the reduced end-diastolic and stroke volumes observed at the organ level in T2DM likely reflect impaired ventricular filling and reduced compliance; early hallmarks of diastolic dysfunction and indeed HFpEF [5]. These changes may represent a distinct electromechanical phenotype from that of heart failure cases with elevated end-diastolic volume.

6.4.5 Modelling glycemic control and electrophysiological disturbances in type 2 diabetes

In this mechanistic study, I modelled the chronic consequences of hyperglycemia by representing CaMKII activation as a proxy for glucotoxic stress. The ventricular cardiomyocyte model was parameterised to reflect progressive disease states: early T2DM with no chronic CaMKII activation, mid-stage disease with sustained CaMKII activation, and late-stage disease additionally incorporating CaMKII-dependent remodelling such as increased I_{NaL} and further impaired SERCA pump function. This framework does not explicitly model glucose fluxes, but rather their downstream signalling consequences via CaMKII, capturing how persistent activation reshapes the action potential and calcium handling, affecting excitation-contraction coupling.

Other computational models that represent the cAMP/PKA pathway enable the simulation of metabolic stressors by modelling chronic hyperglycemia explicitly, incorporating subcellular compartmentalisation to capture effects on excitation-contraction coupling and mitochondrial regulation in cardiomyocytes, allowing for a more biophysically detailed simulation of metabolic perturbations such as impaired glucose regulation [418]. Complementing this, Cortassa's thermokinetic model describes how the Krebs cycle for ATP release, oxidative phosphorylation and mitochondrial Ca^{2+} handling work together to produce ATP, providing a mechanistic explanation of how metabolic stress affects cellular energy [419]. Building on my proposed adaptation of the cellular cardiomyocyte model, which uses CaMKII activation as a proxy for chronic hyperglycemia, integrating these models into the framework could provide more granular explanations into the effects of acute changes in glucose on adverse remodelling in the heart.

While I did not explicitly model glucose dynamics, representing downstream CaMKII effects provides a mechanistic link between chronic metabolic stress and alterations in ionic-level electrophysiological processes and calcium handling. Taken together with clinical findings from Chapter 5, the framework presented in this chapter still supports the view that early and tight glycemic control is critical to prevent long-term functional abnormalities in T2DM, highlighting how chronic CaMKII-mediated signalling may underlie some of the cardiac changes observed in T2DM patients.

6.4.6 ECG simulations: methodology and clinical impact

The ECG is one of the most commonly used indicators of cardiac health in clinical settings, with several quantitative biomarkers derived from its raw signal playing a major role in the diagnosis of cardiac pathologies (see Chapter 2). Compared to echocardiography and blood-based biomarkers such as NT-proBNP, the ECG provides a rapid and low-cost first-line screening tool to identify underlying cardiac rhythm or structural impairments. While NT-proBNP has a higher sensitivity for ruling out heart failure, and echocardiography provides more specific information on the heart's structural, valvular, or pumping abnormalities, the ECG can provide a convenient initial assessment of a range of aspects of cardiac function. This makes it a crucial outcome of in silico cardiac electrophysiology modelling and simulation frameworks. Contingent on sufficient validation, if such a framework can replicate clinical findings, it may become a powerful tool to support personalised patient risk stratification and help inform clinical guidelines.

The simulated QT prolongation observed across all T2DM models reflects underlying ionic changes such as reduced repolarising K^+ currents and increased I_{NaL} and CaMKII activity, consistent with action potential duration prolongation at the cellular level. These findings align with clinical data showing that QTc prolongation is common in T2DM and associated with arrhythmic risk, even in the absence of overt heart disease. The slight QRS shortening in our models can be attributed to increased peak I_{NaL} and faster conduction velocity, suggesting that early ionic changes may initially enhance depolarisation. Overall, these ECG alterations support the use of surface ECG as a sensitive, non-invasive marker of subclinical electromechanical dysfunction in T2DM.

As discussed in Chapter 3 (Section 3.4.3), the ECG can be computed using various simulation approaches, each with its benefits and drawbacks. In this study, I used a fast Eikonal-based approach in one of the model calibration steps to reproduce QRS complexes under baseline healthy conditions. Compared to bidomain or monodomain-based simulations, this approach offered a significant reduction in computation time with minimal sacrifice in accuracy to achieve the intended purpose. For the final simulations, however, I used a monodomain-based approach, to simulate ECGs with a maximal level of biophysical accuracy as well as full flexibility and control of the underlying ionic model. Additionally, the selection of this methodology allows the framework to be developed further to simulate ECGs under arrhythmic conditions. In silico studies of cardiac arrhythmias strictly require biophysically detailed models due to the complexity of both the underlying ionic trigger mechanisms (i.e. early and delayed afterdepolarisations) and the tissue-level substrate (i.e. fibrosis) that can cause complex re-entry patterns, neither of which are reliably achievable yet using Eikonal-based methods.

6.4.7 Further limitations

I considered variability in anatomy, albeit with 6 cases, and this could be expanded further. The number of ionic remodelling scenarios simulated in each anatomical case is also limited to two (with/without disease stage-specific ionic remodelling), given limitations in computational resources. Additional cases could be considered including early-stage T2DM and high BMI, or late-stage T2DM with normal BMI. Previous studies applying multi-scale frameworks in other disease contexts have been validated extensively, and could be used as reference for further validation steps [28, 253, 413]. Finally, the limited availability and therefore reliability of experimental measurements of ionic channel and protein activity in the human diabetic myocardium pose an important limitation. This human data scarcity was addressed through the use of complementary data from rodent models acting as validation, however this does pose the issue of cross-translation between species. In the future, this could be resolved by incorporating further data from experimental cardiac data in humans with T2DM as new studies are published.

As with any modelling approach, computational models of cardiac cells have inherent limitations that must be considered particularly when representing complex conditions such as T2DM. These models can only simulate proteins, pathways and processes explicitly incorporated into their structure. Although current models are increasingly comprehensive and capture a wide spectrum of physiological phenomena, they remain tools including an incomplete representation of biological reality. Many signalling pathways or regulatory mechanisms are not included yet, which is why some mechanisms contributing to the overall disease phenotype may be overlooked. Despite these limitations, computational models are highly valuable as tools to probe pathophysiological systems like T2DM when used synergistically with experimental approaches. Their strength lies in the mechanistic and interpretable nature of their framework: they explicitly represent known biological processes such as ion channel kinetics or intracellular signalling cascades. This structure allows for the direct integration of experimentally observed cellular remodelling into simulations, as well as for systematic validation of computational predictions against experimental and literature data. Thus, when combined with animal or in vitro studies, computer models contribute to a more comprehensive understanding of the mechanisms underlying pathological remodelling in cardiovascular disease than either approach can provide alone.

The same thinking applies to tissue- and organ-level remodelling (e.g. ischemia, fibrosis, epicardial fat), as well as cross-organ interactions and the effect of external but directly related effects such as nervous system remodelling and hypertension. The inclusion or omission of these components should be carefully decided dependent on the specific context of use and the level of biological representation necessary to test the hypotheses proposed.

Ultimately, no model can fully capture the complexity of biological systems. Careful selection and development of a model that is fit for purpose, including representations of the biological remodelling components that are most relevant to the specific research question of interest, while remaining mindful of its assumptions and limitations, is essential for rigorous and meaningful application of computational modelling and simulation in cardiac research.

The framework I presented here serves as a proof-of-concept to highlight the capacity and potential of human-based cardiac modelling and simulation in the context of use of T2DM cardiac electrophysiology, an area with huge clinical potential as discussed throughout the thesis. While I accounted for ionic remodelling, disease severity and anatomical heterogeneity, new modelling components reflecting different pathophysiological phenomena in T2DM, could be incorporated into this framework to expand the current model's scope and potential. I will discuss avenues for future development of this work in the following chapter.

6.5 Conclusion

In this chapter, I presented a framework for the mechanistic computational modelling and simulation of T2DM-specific human cardiac electrophysiology, accounting for the effect of ionic remodelling, organ-level electrophysiological parameters and anatomical variability in T2DM. Simulations of T2DM cardiac remodelling revealed prolonged action potentials, with late-stage effects amplified by elevated CaMKII, increased I_{NaL}, and reduced SERCA pump flux. Early-stage T2DM decreased calcium transients, impairing excitation-contraction coupling, while later-stage chronic CaMKII activation appears to partially compensate this effect at cellular-level. Organ-level ECG simulations mirrored these effects, showing QT prolongation, minor T wave reduction, and modest QRS duration reduction, consistent with clinical observations discussed in the previous chapter.

Overall, this framework paves the way toward future in silico studies of arrhythmogenesis and cardiac pharmacological treatments in T2DM, while contributing to overcoming the barrier to translation represented by the widespread use of small animal models in T2DM-specific cardiac experimental research, offering novel avenues for pre-clinical studies in this field. The framework could also be extended to model complex, more advanced disease states such as T2DM with comorbid HFpEF, hypertension, or ischemic heart disease. Subject to further development and validation, this approach has the potential to inform patient stratification strategies, support the design of personalised therapeutical avenues, and enable in silico evaluation of drug safety and efficacy in

high-risk T2DM populations.

In the next and final chapter, I will summarise the main contributions arising from the work presented in this thesis, and outline potential future directions to further develop this work and its clinical impact.

The important thing about a problem is not its solution,
but the strength we gain in finding the solution.

— Seneca, Stoic philosopher (4BC-65AD)

7 Conclusion

7.1 Summary and key contributions

Following the Introduction in **Chapter 1**, Chapters 2 and 3 established the clinical and methodological foundations this thesis builds upon. In **Chapter 2**, I reviewed key physiological and pathological concepts, with a focus on heart failure with preserved ejection fraction (HFpEF), and type 2 diabetes mellitus (T2DM) and obesity-related cardiac disorders, providing essential context for the analyses in later chapters. **Chapter 3** introduced the UK Biobank dataset and outlined the core methodologies including statistical, machine learning, and computational modelling and simulation methods that are harnessed throughout the thesis to investigate cardiac structure, function, and electrophysiology across biophysical scales.

In **Chapter 4**, I investigated the heterogeneity of cardiometabolic HFpEF clinical profiles in an elderly cohort of participants from the UK Biobank, with a focus on sex-specific differences. HFpEF is a multi-factorial condition driven by overlapping metabolic comorbidities such as hypertension, obesity, and T2DM. Women, particularly post-menopause, are at elevated risk due to age-related vascular and metabolic changes, yet remain underrepresented in clinical studies. To address this gap, I applied latent class analysis to identify phenogroups within the HFpEF population based on cardiometabolic

profiles, identifying three subgroups with distinct clinical, functional, and cardiac characteristics: a mostly male and multimorbid group; a group with a high prevalence of severe obesity, abnormal waist circumference and with the highest relative proportion of females; and a group with an apparently lower comorbidity burden aside from hypertension. Notable sex- and phenogroup-specific differences in imaging-derived markers of myocardial function in HFpEF, beyond the LVEF, included worse contractility in men within the obese subgroup; an attenuation of expected sex differences in contractility in highly comorbid and hypertensive groups; and a higher resting cardiac output in obese subgroup consistent with high-output HFpEF. These characteristics may influence disease trajectory and treatment; while outcome analyses were not possible in this study due to sample size, such analyses in a validation cohort may further strengthen the clinical relevance and translatability of these subgroups.

These results offer a more nuanced understanding of cardiometabolic HFpEF and, importantly, challenge the traditional "one-size-fits-all" approach to the condition, supporting the need for more tailored, sex-specific diagnostic and therapeutic strategies in these patients. Findings also complement the recent hypothesis proposed on the unifying role of adipokines in HFpEF pathophysiology; while dysfunctional visceral adipose tissue may be the fundamental driving force of HFpEF, this may translate differently across individuals and result in different salient clinical phenogroups as highlighted in the study presented in this thesis. Further investigations on visceral adipose tissue function and distribution in these subgroups, and subsequent outcome analysis in a larger validation cohort, would provide an interesting validation of this hypothesis.

In **Chapter 5**, I examined earlier stages of the HFpEF syndrome spectrum by focusing on T2DM, one of its most significant risk factors. Large-scale studies of cardiac function in patients with T2DM prior to overt cardiovascular disease (CVD) are scarce; this gap is critical, as early detection of cardiac abnormalities in these patients could significantly reduce risk and burden of future HFpEF and cardiac arrhythmias. To address this, I carried out the largest study to-date using ECG and cardiac imaging data from over 1,700 T2DM patients and matched controls in the UK Biobank. I hypothesised that,

even in the absence of diagnosed heart disease, patients with T2DM would exhibit early cardiac dysfunction. Using multivariate multiple regression models, I showed that these patients exhibited markers of early cardiac abnormalities including a higher heart rate, QTc prolongation, T wave amplitude reduction, as well as lower stroke volume and increased left ventricular wall thickness. Increased QRS duration, and left ventricular wall thickness and mass, were most strongly associated with future CVD.

These findings provide strong evidence supporting the presence of subclinical cardiac changes in T2DM patients. This highlights the importance of early cardiac monitoring of these high-risk patients, with direct implications to improve early identification of cardiac decline and ultimately to reduce the progression of disease to overt HFpEF and cardiac arrhythmias in patients with T2DM.

In **Chapter 6**, I harnessed a different methodological approach to further understand the mechanisms underlying electrophysiological cardiac changes observed in the previous chapter. I developed a multi-scale, human-based computational modelling framework to simulate electrophysiological behaviour in virtual hearts of selected UK Biobank participants with T2DM and obesity. The framework presented is the first to apply this approach to diabetes-related cardiac electrophysiological dysfunction. I adapted a previously validated 2D-to-3D anatomical reconstruction pipeline and a human ventricular cardiomyocyte electrophysiology model to create a GPU-enabled multi-scale simulation pipeline that can simulate ECG characteristics in patients with varying cardiac anatomy, body shape, sex, and severity of T2DM and obesity. Simulations of T2DM cardiomyocytes revealed disease stage-dependent changes in action potential duration and calcium handling, driven by reduced repolarising K^+ currents, altered NCX and L-type Ca^{2+} current activity, and increased CaMKII signalling. These cellular-level alterations reproduced QT interval prolongation and modest QRS shortening in simulated ECGs, as well as slight T wave amplitude reduction, reflecting early subclinical changes consistent with clinical observations in T2DM at cohort-level. The results also highlight mechanistic parallels between T2DM and HFpEF at the ionic level, while identifying distinct electromechanical

features such as reduced ventricular filling, i.e. diastolic dysfunction, that may differentiate the diabetic heart from other forms of heart failure.

Overall, this framework allows us to explore certain mechanistic underpinnings of cardiac dysfunction in T2DM and obesity in a highly controllable virtual environment that offers the unique ability to quantify the exact effect of selected parameters in isolation (e.g. ionic remodelling, cardiac and body anatomy, electrophysiological properties at organ level) and their combined effect. Considering the complexity of the problem, this task is typically very hard to achieve experimentally or clinically, making such a modelling and simulation framework particularly advantageous in this context of use. Thus, subject to further development and validation, and the inclusion of other disease-specific factors such as tissue-level remodelling (e.g. diffuse myocardial fibrosis or epicardial adipose tissue), this framework represents a promising first step in the direction of cardiac digital twins, personalised arrhythmic risk stratification and in silico drug evaluation in patients with T2DM and obesity.

7.2 Limitations and challenges

Several limitations must be acknowledged, particularly regarding the use of UK Biobank dataset in the context of the studies presented in this thesis. While some were previously discussed in their respective chapters, they are reiterated and expanded below to provide a comprehensive overview of the thesis Chapters' scope and context.

First, the UK Biobank remains relatively "young" in terms of clinical outcomes, particularly for late-onset conditions such as HFpEF. This inevitably affects the observed event rates and the apparent prevalence patterns reported in Chapter 4, where HFpEF was more frequent in men than women, in contrast to established epidemiological trends [5, 13, 15]. Nevertheless, the large scale, deep phenotyping, and longitudinal follow-up of UK Biobank provide a valuable and evolving resource for investigating early cardiometabolic traits associated with HFpEF. As the UK Biobank participants continue to age and follow-up durations increase, future analyses will be well-positioned to validate and extend these findings in more clinically mature data.

Second, at the time of analysis, linkage to primary care records was incomplete, with only 45% of the UK Biobank cohort having accessible general practice data [25]. This primarily impacts the ascertainment of chronic conditions such as T2DM which are often diagnosed and managed in primary care settings. While this limitation may have led to underestimation of T2DM prevalence and restricted cohort size, rigorous inclusion criteria were applied to ensure a robust definition of the condition. Full linkage across the cohort will further enhance future work in this area.

Third, the UK Biobank cohort remains predominantly of White European ancestry (over 94% of participants) [420]. This restricts the immediate generalisability of findings to broader, multi-ethnic populations, particularly relevant given known differences in cardiometabolic disease risk and expression across ancestry groups [421]. However, the well-characterised nature of this cohort provides a crucial reference point for understanding disease mechanisms in a relatively homogeneous population, which could be compared with external datasets such as the China Kadoorie Biobank to explore ancestry-specific differences in cardiometabolic risk [421].

The granularity of pharmacotherapy data in UK Biobank must also be considered. In this thesis, medication information from broad yet well-defined categories was incorporated systematically as model covariates or analysis outcomes (Chapters 4 and 5) and through explicit exclusion of participants taking QT-prolonging medication from a list of selected drugs (Chapter 6), thereby minimising confounding effects on ECG and imaging-derived phenotypes. Although future availability of prescription duration and dosage data would further refine such analyses, the use of well-curated medication categories already offers meaningful clinical context to the observed associations.

Finally, UK Biobank's data sharing policies and participant confidentiality guidelines understandably restrict the sharing of re-identifiable patient-level outputs. While this presents a challenge for full reproducibility and transparency, it reflects the UK Biobank's high ethical and privacy standards. To protect privacy while supporting open science, creative solutions have been proposed, for example a recent study where ~55000 cardiac meshes were shared with information on sex, age and BMI in bins rather than exact values [238]. Similar anonymisation strategies could support future efforts to release

large-scale outputs for re-use throughout the community. Such initiatives demonstrate how UK Biobank-based modelling studies can continue to promote open and reproducible biomedical research while safeguarding participant confidentiality.

The reader is referred to Chapter 6 for limitations of computational modelling and simulation, discussed extensively at the end of that Chapter.

7.3 Future directions

7.3.1 Improvements to the multi-scale modelling and simulation framework

As discussed in Chapter 6, sex-specific ionic remodelling was not considered yet in the framework presented, but could be implemented in the future to enable sex-specific differences to be studied more accurately at cellular-level and downstream in patients with T2DM and/or obesity prior to overt HFpEF [413]. The number of simulations could be expanded such that each anatomical case is used to simulate an ECG with each of the three proposed stage-specific ionic remodelling scenarios, in order to perform a more extensive sensitivity analysis allowing for the individual effect of each component (sex, anatomy, ionic remodelling severity) to be better evaluated in isolation. Finally, the incorporation of fibrotic and/or epicardial fat remodelling at tissue-level in the anatomical models developed would increase the potential scope of investigations enabled by the framework. Using the current dataset, generating a physiologically accurate, patient-specific virtual replica of such tissue-level remodelling would be challenging, as the UK Biobank does not contain sufficient data to obtain ground truth myocardial fibrosis and epicardial fat patterns across the entire ventricles (relevant available data is limited to a single slice of T1 mapping). However, there exist multiple approaches to model these computationally, which could be incorporated into the virtual cardiac geometries developed in Chapter 6 [30, 302] (see Chapter 3). For example, a computational representation of fibrosis could be implemented in the current framework to reflect existing clinical evidence demonstrating the presence of heterogeneously distributed diffuse fibrosis across different regions of the ventricles in T2DM patients [125]. This would enable us to conduct in silico mechanistic studies to generate or test new hypotheses, for example to quantify

the relative role of ionic vs. fibrotic remodelling in arrhythmic risk in patients with T2DM and/or obesity across the HFpEF continuum.

As it stands, the aim of the framework presented in Chapter 6 was not to generate digital twins that produce exact replicas of each T2DM patient's clinical ECG, but rather using T2DM patient anatomies and corresponding ECGs from equivalent "control twins" as a modelling basis for mechanistic investigations. Achieving a "true" digital twin would require better parametrisation including unmeasured factors (e.g. fibrosis, epicardial fat, autonomic dysfunction, possible subclinical ischemia, etc.) and more precise, comprehensive data on pharmacotherapy. These challenges highlight the difficulty of fully capturing a patient's state in a model, as discussed in Chapter 6. Ultimately, the scope and required level of accuracy of a model depend firmly on its context of use; in the future, the framework will have to be adapted and validated for each specific research question it intends to answer.

7.3.2 Autonomous nervous system modelling

A particular novel and promising avenue of future work would be to investigate the interplay of autonomic nervous system remodelling in T2DM, obesity and HFpEF. Computational modelling of cardiac innervation in the ventricles remains unexplored but represents a very promising avenue to explore in the T2DM-obesity-HFpEF inflammatory phenotype, given the relevance of underlying pathophysiological mechanisms discussed throughout this thesis.

One relevant previous computational study modelled the intrinsic cardiac autonomic nervous system, including cardiac ganglia and sympathetic nerve activity in the atria, to investigate its effect on the initiation and maintenance of atrial fibrillation [422]. The cardiac conduction system has also been modelled successfully in silico, accurately replicating the intricate branching patters of fibres observed in macroscopic histological studies of the heart and capturing electrophysiological heterogeneities at tissue-level [423, 424]. These fibre networks can be optimised to fit individual patient anatomies, further contributing to the personalisation of cardiac modelling and simulation [423]. Together, these methods represent exciting avenues to model the electrophysiological abnormalities

and tissue-level heterogeneities that are present in the T2DM myocardium, and to quantify the impact of these abnormalities, relative to ionic remodelling, at organ-level.

7.3.3 In silico studies of pharmacological treatment of type 2 diabetes

To date, most of the information on the effectiveness and safety of drugs in pre-clinical stages is derived from animal models. Although beneficial, there are challenges in the translation of these findings to humans, including mechanistic relevance, as pathophysiological mechanisms may differ between species and sex-specific effects, as experimental models often include male animals only. In addition, although often considered as focusing on single variables, animal models inherently involve feedback and systemic responses that complicate mechanistic interpretation.

In silico trials may help us tackle these issues, by offering an alternative pathway to evaluate drug safety and efficacy: simulating the action of drug in a highly controllable virtual human environment instead of a live animal model [425]. Using a multi-scale, human-based, biophysically detailed framework similar to those presented in Chapter 3 and Chapter 6, the known or hypothesised mechanistic action of a given compound is modelled at ionic level, and its downstream effects are examined by quantifying changes in simulated electrophysiological or electromechanical biomarkers at cellular and/or whole-organ level [26, 246, 426–428]. This novel computational approach may be considered complementary to pre-clinical studies, rather than a replacement; it remains tightly intertwined with experimental studies, and which are crucial in providing data quantifying the effect of a given compound on protein expression and/or ionic current density.

Contingent on the availability of sufficient and reliable experimental data, and further validation of the T2DM modelling and simulation framework I presented in Chapter 6, in silico drug evaluation holds huge potential to investigate the multi-scale cardiac effects of drugs in T2DM (e.g. those presented in Chapter 2, Section 2.5.5). Such an approach would be especially valuable in high-risk clinical scenarios that have been poorly studied so far, especially in humans, such as the complex overlap of T2DM, obesity, HFpEF

and ventricular arrhythmias, or investigating the cardiac mechanisms of promising new therapeutic avenues such as SGLT2, I_{NaL} and CaMKII inhibitors.

7.3.4 Extending image-based characterisation with T1 mapping and radiomics

Radiomics is an emerging field of medicine, where powerful algorithmic approaches are applied to high-dimensional, often multi-modal clinical imaging datasets of large patient cohorts, to uncover new biomarkers of disease and to improve diagnosis and prognosis. This approach holds promise in clinical management of heart failure and other CVDs, subject to large enough datasets and high computational power [429]. One study extracted over 600 radiomics features, reflecting cardiac structure and texture, to characterise groups of patients from the UK Biobank with different cardiovascular risk factors [430]. A similar study combined vascular risk factors, traditional image-derived metrics such as ventricular volume and mass, and radiomics features to quantify tissue shape and texture, in order to build a predictive model for different forms of CVDs including heart failure [431]. The inclusion of radiomics features improved the predictive power of models, especially when predicting atrial fibrillation and heart failure. Multiple studies used radiomics features to predict arrythmias and to stratify patients at risk post-infarction [432–434]. Many studies have investigated the value of T1 as a useful marker to establish diagnosis and/or prognosis of heart failure, especially in HFpEF where traditional imaging parameters such as LVEF are not sufficient to fully characterise the disease [435–438]. A higher global myocardial T1 was also associated with higher risk of incident disease and mortality [439].

Thus, radiomics features may have the ability to detect subtle tissue characteristics of cardiac remodelling that are otherwise impossible to capture using traditional imaging metrics or structural imaging only. Considering additional, functional imaging modalities concurrently in population-level analyses of HFpEF and T2DM would provide a novel image-based approach for disease quantification, for example to study the link between tissue-level metabolism (magnetic resonance spectroscopy, functional magnetic resonance imaging) and diffuse fibrosis (T1 mapping to estimate extracellular volume), and their potential value in assessing and predicting disease progression and outcomes. As the UK

Biobank contains a single slice of T1 mapping for each imaging case, external datasets would be required to harness the advantages offered by other imaging modalities.

7.3.5 Ideal dataset specifications

The obesity-T2DM-HFpEF spectrum is highly complex in terms of etiology, underlying biological mechanisms, and salient clinical profiles. Thus, I propose three key dataset components to consider when selecting or indeed setting up a new cohort study to investigate the interplay of these tightly intertwined conditions, namely: breadth (i.e. number of participants and population heterogeneity), depth (i.e. availability of different data types reflecting different biological and clinical features), and a longitudinal component (i.e. reflecting progression of health/disease states through time, including linkage to medical history). Rarely are all three components fully satisfied. For example, electronic health records (EHRs) such as the UK's Clinical Practice Research Datalink (CPRD) provide extensive coverage of medical history of over 70 million patients (of which 13-18 million are still active) through combined access to primary care data from general practices (GPs) as well as linked secondary care data on hospital admissions through Hospital Episode Statistics (HES), procedure codes from the Office of Population Censuses and Surveys Classification of Interventions and Procedures (OPCS), death registration data from the Office for National Statistics (ONS), and cancer registry data from the National Cancer Registration and Analysis Service (NCRAS). While this dataset provides extensive coverage in terms of cohort size and timeframe, it does not capture molecular or clinical phenotypes, which are essential to characterise disease beyond simple diagnosis codes. In contrast, HFpEF-focused initiatives may prioritise depth over breadth to ensure better phenotypical characterisation, such as the UK HFpEF Collaborative Group and the United States National Heart, Lung and Blood Institute's "HeartShare" consortium [350–352]. These consortia aim to aggregate multiple existing HFpEF data sources, and collect new and diverse data including functional molecular omics data (i.e. blood-based characterisation of the genome, proteome, and metabolome) and multi-modal clinical data, to enable larger-scale HFpEF phenotyping research and trials. Due to the depth of data collected, however, these studies may be limited in

sample size. As extensively discussed throughout this thesis, the UK Biobank aims to tackle these considerations by providing the largest and most "deeply" phenotyped longitudinal study of human health to date, including medical history through EHR linkage and patient-reported outcomes. It does, however, lack previously discussed data types such as different imaging modalities offering complementary information on underlying organ and tissue (patho)physiology; in-depth information on medication dosage and treatment duration; and detailed long-term information on environmental and lifestyle exposures. Thus, to study the obesity-T2DM-HFpEF phenotype at scale, with sufficient depth, and throughout time, selected or new datasets should ideally account for an optimal combination of the aforementioned data components.

7.4 Concluding remarks

This thesis advances our understanding of the cardiac consequences of metabolic disease, an increasingly prevalent health burden that poses major medical and societal challenges. While its metabolic driving factors are becoming increasingly acknowledged, specifically the role of dysfunctional visceral adipose tissue and adipokine secretion, HFpEF remains an extremely heterogenous cardiac condition clinically. T2DM is tightly linked to obesity; affecting over half a billion adults worldwide, it is a known major risk factor for HFpEF, cardiac arrhythmias, and sudden cardiac death, the most current mode of death in HFpEF. However, its early adverse cardiac effects in humans remain understudied.

By integrating large-scale, population-level statistics with mechanistic, patient-specific computational modelling and simulation, I have leveraged multi-modal clinical data from the UK Biobank, firstly to identify and then explain cardiac abnormalities arising in patients with abnormal metabolic profiles. I uncovered sex-specific patterns in HFpEF linked to different cardiometabolic profiles and identified early markers of cardiac dysfunction in patients with T2DM before the onset of HFpEF and other CVDs. I then developed a modelling framework capable of simulating electrophysiological function across biophysical scales in human hearts of patients with varying severity of T2DM and obesity, bridging cellular and organ-level mechanisms. Unlike data-driven approaches, this framework allows direct interrogation of mechanisms underlying the cardiac changes

observed clinically in individuals with T2DM. Simulations showed disease stage-dependent alterations in action potential duration and calcium handling, driven by reduced repolarising K^+ currents, modified NCX and L-type Ca^{2+} current activity, and heightened CaMKII signalling, leading to QT prolongation and mild QRS duration shortening and T wave amplitude reduction, consistent with early subclinical changes identified at cohort-level. Contingent on further validation, this methodological framework represents a significant step for future in silico investigations of patients with T2DM-related cardiac conditions, enabling personalised patient arrhythmic risk stratification and in silico trials.

Together, these contributions offer exciting new directions, aligned with the vision of personalised medicine, for the early identification and improved management of life-threatening cardiac conditions stemming from metabolic diseases.

References

[1] Ramon Luengo-Fernandez et al. “Economic burden of cardiovascular diseases in the European Union: a population-based cost study”. In: *European Heart Journal* 44.45 (Dec. 2023), pp. 4752–4767.

[2] Dhruv S. Kazi et al. “Forecasting the Economic Burden of Cardiovascular Disease and Stroke in the United States Through 2050: A Presidential Advisory From the American Heart Association”. In: *Circulation* 150.4 (July 2024).

[3] World Health Organization. “Global action plan for the prevention and control of noncommunicable diseases 2013-2020”. In: *Global action plan for the prevention and control of noncommunicable diseases 2013-2020*. 2013.

[4] Bryan Chong et al. “Global burden of cardiovascular diseases: projections from 2025 to 2050”. In: *European Journal of Preventive Cardiology* 32.11 (Aug. 2025), pp. 1001–1015.

[5] Margaret M. Redfield and Barry A. Borlaug. “Heart Failure With Preserved Ejection Fraction”. In: *JAMA* 329.10 (Mar. 2023), p. 827.

[6] Gianluigi Savarese et al. “Global burden of heart failure: a comprehensive and updated review of epidemiology”. In: *Cardiovascular Research* 118.17 (Jan. 2023), pp. 3272–3287.

[7] Sanjiv J. Shah et al. “Phenotype-Specific Treatment of Heart Failure With Preserved Ejection Fraction”. In: *Circulation* 134.1 (July 2016), pp. 73–90.

[8] Sanjiv J. Shah et al. “Research Priorities for Heart Failure With Preserved Ejection Fraction”. In: *Circulation* 141.12 (Mar. 2020), pp. 1001–1026.

[9] Stefan D. Anker et al. “Patient phenotype profiling in heart failure with preserved ejection fraction to guide therapeutic decision making. A scientific statement of the Heart Failure Association, the European Heart Rhythm Association of the European Society of Cardiology, and the European Society of Hypertension”. In: *European Journal of Heart Failure* 25.7 (July 2023), pp. 936–955.

[10] Stefan D. Anker et al. “Empagliflozin in Heart Failure with a Preserved Ejection Fraction”. In: *New England Journal of Medicine* 385.16 (Oct. 2021), pp. 1451–1461.

[11] Scott D. Solomon et al. “Dapagliflozin in heart failure with preserved and mildly reduced ejection fraction: rationale and design of the DELIVER trial”. In: *European Journal of Heart Failure* 23.7 (July 2021), pp. 1217–1225.

[12] Anna L. Beale et al. “Sex Differences in Cardiovascular Pathophysiology”. In: *Circulation* 138.2 (July 2018), pp. 198–205.

[13] Annamaria De Bellis et al. “Gender-related differences in heart failure: beyond the “one-size-fits-all” paradigm”. In: *Heart Failure Reviews* 25.2 (Mar. 2020), pp. 245–255.

[14] Carolyn S P Lam et al. “Sex differences in heart failure”. In: *European Heart Journal* 40.47 (Dec. 2019), pp. 3859–3868.

[15] Vera Regitz-Zagrosek. “Sex and Gender Differences in Heart Failure”. In: *International Journal of Heart Failure* 2.3 (2020), p. 157.

[16] Sara Hägg and Juulia Jylhävä. “Sex differences in biological aging with a focus on human studies”. In: *eLife* 10 (May 2021).

[17] Daria V. Zhernakova et al. “Age-dependent sex differences in cardiometabolic risk factors”. In: *Nature Cardiovascular Research* 1.9 (Sept. 2022), pp. 844–854.

[18] Hong Sun et al. “IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045”. In: *Diabetes Research and Clinical Practice* 183 (Jan. 2022), p. 109119.

[19] Nikolaus Marx et al. “2023 ESC Guidelines for the management of cardiovascular disease in patients with diabetes: Developed by the task force on the management of cardiovascular disease in patients with diabetes of the European Society of Cardiology (ESC)”. In: *European Heart Journal* 44.39 (Oct. 2023), pp. 4043–4140. URL: https://doi.org/10.1093/eurheartj/ehad192.

[20] L. J. M. Boonman-de Winter et al. “High prevalence of previously unknown heart failure and left ventricular dysfunction in patients with type 2 diabetes”. In: *Diabetologia* 55.8 (Aug. 2012), pp. 2154–2162.

[21] Petar M. Seferović and Walter J. Paulus. “Clinical diabetic cardiomyopathy: a two-faced disease with restrictive and dilated phenotypes”. In: *European Heart Journal* 36.27 (July 2015), pp. 1718–1727.

[22] Araz Rawshani et al. “Cardiac arrhythmias and conduction abnormalities in patients with type 2 diabetes”. In: *Scientific Reports* 13.1 (Jan. 2023), p. 1192.

[23] Bence Hegyi et al. “CaMKII Serine 280 O-GlcNAcylation Links Diabetic Hyperglycemia to Proarrhythmia”. In: *Circulation Research* 129.1 (June 2021), pp. 98–113.

[24] Konstantinos Makrilakis and Stavros Liatis. “Cardiovascular Screening for the Asymptomatic Patient with Diabetes: More Cons Than Pros”. In: *Journal of Diabetes Research* 2017 (2017), pp. 1–19.

[25] Naomi E. Allen et al. “Prospective study design and data analysis in UK Biobank”. In: *Science Translational Medicine* 16.729 (Jan. 2024).

[26] Jakub Tomek et al. “Development, calibration, and validation of a novel human ventricular myocyte model in health, disease, and drug block”. In: *eLife* 8 (Dec. 2019).

[27] Oliver J. Britton et al. “Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology”. In: *Proceedings of the National Academy of Sciences* 110.23 (June 2013).

[28] Xin Zhou et al. “Clinical phenotypes in acute and chronic infarction explained through human ventricular electromechanical modelling and simulations”. In: *eLife* 13 (Dec. 2024).

[29] Aurore Lyon et al. “Electrocardiogram phenotypes in hypertrophic cardiomyopathy caused by distinct mechanisms: apico-basal repolarization gradients vs. Purkinje-myocardial coupling abnormalities”. In: *EP Europace* 20.suppl_3 (Nov. 2018), pp. iii102–iii112.

[30] James A Coleman et al. “Mechanisms of ischaemia-induced arrhythmias in hypertrophic cardiomyopathy: a large-scale computational study”. In: *Cardiovascular Research* 120.8 (July 2024), pp. 914–926.

[31] E Pierce. *Diagram of the human heart, with labels for chambers, valves and blood vessels.* 2003. URL: https://commons.wikimedia.org/wiki/File:Diagram_of_the_human_heart.svg.

[32] Richard Klabunde. *Normal and Abnormal Electrical Conduction*. 2023. URL: https://cvphysiology.com/arrhythmias/a003.

[33] Sanne de Jong et al. "Fibrosis and Cardiac Arrhythmias". In: *Journal of Cardiovascular Pharmacology* 57.6 (June 2011), pp. 630–638.

[34] Yasar Sattar and Lovely Chhabra. *Electrocardiogram*. 2025.

[35] Daniel R. Messroghli et al. "Clinical recommendations for cardiovascular magnetic resonance mapping of T1, T2, T2* and extracellular volume: A consensus statement by the Society for Cardiovascular Magnetic Resonance (SCMR) endorsed by the European Association for Cardiovascular Imaging (EACVI)". In: *Journal of Cardiovascular Magnetic Resonance* 19.1 (Dec. 2016), p. 75.

[36] Nathalie Conrad et al. "Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals". In: *The Lancet* 391.10120 (Feb. 2018), pp. 572–580.

[37] Theresa A McDonagh et al. "2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure". In: *European Heart Journal* 42.36 (Sept. 2021), pp. 3599–3726.

[38] Theresa A McDonagh et al. "2023 Focused Update of the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure". In: *European Heart Journal* 44.37 (Oct. 2023), pp. 3627–3639.

[39] Walter J. Paulus and Carsten Tschöpe. "A Novel Paradigm for Heart Failure With Preserved Ejection Fraction". In: *Journal of the American College of Cardiology* 62.4 (July 2013), pp. 263–271.

[40] Carly Adamson et al. "Dapagliflozin for heart failure according to body mass index: the DELIVER trial". In: *European Heart Journal* 43.41 (Nov. 2022), pp. 4406–4417.

[41] Salvatore Carbone et al. "Obesity Contributes to Exercise Intolerance in Heart Failure With Preserved Ejection Fraction". In: *Journal of the American College of Cardiology* 68.22 (Dec. 2016), pp. 2487–2488.

[42] Yogesh N.V. Reddy et al. "Characterization of the Obese Phenotype of Heart Failure With Preserved Ejection Fraction: A RELAX Trial Ancillary Study". In: *Mayo Clinic Proceedings* 94.7 (July 2019), pp. 1199–1209.

[43] Milton Packer. "The Adipokine Hypothesis of Heart Failure With a Preserved Ejection Fraction". In: *JACC* 86.16 (Oct. 2025), pp. 1269–1373.

[44] Ting Wang et al. "Estrogen-Related Receptor α (ERRα) and ERRγ Are Essential Coordinators of Cardiac Metabolism and Function". In: *Molecular and Cellular Biology* 35.7 (Apr. 2015), pp. 1281–1298.

[45] Ariane A. Sickinghe et al. "Estrogen Contributions to Microvascular Dysfunction Evolving to Heart Failure With Preserved Ejection Fraction". In: *Frontiers in Endocrinology* 10 (July 2019).

[46] Maria Luisa Barcena de Arellano et al. "Sex differences in the aging human heart: decreased sirtuins, pro-inflammatory shift and reduced anti-oxidative defense". In: *Aging* 11.7 (Apr. 2019), pp. 1918–1933.

[47] Domingo A. Pascual-Figal et al. “Sex Hormone-Binding Globulin: A New Marker of Disease Severity and Prognosis in Men With Chronic Heart Failure”. In: *Revista Española de Cardiología (English Edition)* 62.12 (Dec. 2009), pp. 1381–1387.

[48] Ewa A. Jankowska et al. “Anabolic Deficiency in Men With Chronic Heart Failure”. In: *Circulation* 114.17 (Oct. 2006), pp. 1829–1837.

[49] Akiomi Yoshihisa et al. “Relation of Testosterone Levels to Mortality in Men With Heart Failure”. In: *The American Journal of Cardiology* 121.11 (June 2018), pp. 1321–1327.

[50] Vera Regitz-Zagrosek, Sebastian Brokat, and Carsten Tschope. “Role of Gender in Heart Failure with Normal Left Ventricular Ejection Fraction”. In: *Progress in Cardiovascular Diseases* 49.4 (Jan. 2007), pp. 241–251.

[51] Ankeet S. Bhatt, Andrew P. Ambrosy, and Eric J. Velazquez. “Adverse Remodeling and Reverse Remodeling After Myocardial Infarction”. In: *Current Cardiology Reports* 19.8 (Aug. 2017), p. 71.

[52] Beatrice Musumeci et al. “Left Ventricular Remodeling in Hypertrophic Cardiomyopathy: An Overview of Current Knowledge”. In: *Journal of Clinical Medicine* 10.8 (Apr. 2021), p. 1547.

[53] Stanley Nattel, Brett Burstein, and Dobromir Dobrev. “Atrial Remodeling and Atrial Fibrillation”. In: *Circulation: Arrhythmia and Electrophysiology* 1.1 (Apr. 2008), pp. 62–73.

[54] Nikolaos G. Frangogiannis. “Regulation of the Inflammatory Response in Cardiac Repair”. In: *Circulation Research* 110.1 (Jan. 2012), pp. 159–173.

[55] Sameer Ather et al. “Impact of Noncardiac Comorbidities on Morbidity and Mortality in a Predominantly Male Population With Heart Failure and Preserved Versus Reduced Ejection Fraction”. In: *Journal of the American College of Cardiology* 59.11 (Mar. 2012), pp. 998–1005.

[56] Rudolf A. de Boer et al. “Towards better definition, quantification and treatment of fibrosis in heart failure. A scientific roadmap by the Committee of Translational Research of the Heart Failure Association (HFA) of the European Society of Cardiology”. In: *European Journal of Heart Failure* 21.3 (Mar. 2019), pp. 272–285.

[57] Bilal Aijaz et al. “Abnormal Cardiac Structure and Function in the Metabolic Syndrome: A Population-Based Study”. In: *Mayo Clinic Proceedings* 83.12 (Dec. 2008), pp. 1350–1357.

[58] John William McEvoy et al. “2024 ESC Guidelines for the management of elevated blood pressure and hypertension”. In: *European Heart Journal* 45.38 (Oct. 2024), pp. 3912–4018.

[59] Stanley Nattel et al. “Arrhythmogenic Ion-Channel Remodeling in the Heart: Heart Failure, Myocardial Infarction, and Atrial Fibrillation”. In: *Physiological Reviews* 87.2 (Apr. 2007), pp. 425–456.

[60] Sven-Oliver Tröbs et al. “Association of Global Longitudinal Strain With Clinical Status and Mortality in Patients With Chronic Heart Failure”. In: *JAMA Cardiology* 6.4 (Apr. 2021), p. 448.

[61] Jin Joo Park et al. “Global Longitudinal Strain to Predict Mortality in Patients With Acute Heart Failure”. In: *Journal of the American College of Cardiology* 71.18 (May 2018), pp. 1947–1957.

[62] Steffen Pabel et al. “Empagliflozin directly improves diastolic function in human heart failure”. In: *European Journal of Heart Failure* 20.12 (Dec. 2018), pp. 1690–1700.

[63] Gary D. Lopaschuk and Subodh Verma. “Mechanisms of Cardiovascular Benefits of Sodium Glucose Co-Transporter 2 (SGLT2) Inhibitors”. In: *JACC: Basic to Translational Science* 5.6 (June 2020), pp. 632–644.

[64] Alexey V. Karpushev et al. “SGLT2 Inhibitor Empagliflozin Modulates Ion Channels in Adult Zebrafish Heart”. In: *International Journal of Molecular Sciences* 23.17 (Aug. 2022), p. 9559.

[65] Mikhail N. Kosiborod et al. “Semaglutide in Patients with Heart Failure with Preserved Ejection Fraction and Obesity”. In: *New England Journal of Medicine* 389.12 (Sept. 2023), pp. 1069–1084.

[66] Javed Butler et al. “Semaglutide versus placebo in people with obesity-related heart failure with preserved ejection fraction: a pooled analysis of the STEP-HFpEF and STEP-HFpEF DM randomised trials”. In: *The Lancet* 403.10437 (Apr. 2024), pp. 1635–1648.

[67] Barry A. Borlaug et al. “Semaglutide in HFpEF across obesity class and by body weight reduction: a prespecified analysis of the STEP-HFpEF trial”. In: *Nature Medicine* 29.9 (Sept. 2023), pp. 2358–2365.

[68] Milton Packer et al. “Tirzepatide for Heart Failure with Preserved Ejection Fraction and Obesity”. In: *New England Journal of Medicine* 392.5 (Jan. 2025), pp. 427–437.

[69] Muthiah Vaduganathan. “GLP-1 receptor agonists in heart failure”. In: *The Lancet* 404.10454 (Aug. 2024), pp. 727–729.

[70] Scott D. Solomon et al. “Finerenone in Heart Failure with Mildly Reduced or Preserved Ejection Fraction”. In: *New England Journal of Medicine* 391.16 (Oct. 2024), pp. 1475–1485.

[71] Alberto Aimo et al. “Sex-related differences in chronic heart failure”. In: *International Journal of Cardiology* 255 (Mar. 2018), pp. 145–151.

[72] Raffaele Bugiardini et al. “Prior Beta-Blocker Therapy for Hypertension and Sex-Based Differences in Heart Failure Among Patients With Incident Coronary Heart Disease”. In: *Hypertension* 76.3 (Sept. 2020), pp. 819–826.

[73] Jan B Östergren. “Angiotensin receptor blockade with candesartan in heart failure: findings from the Candesartan in Heart failure — Assessment of Reduction in Mortality and morbidity (CHARM) programme”. In: *Journal of Hypertension* 24.Suppl 1 (Mar. 2006), S3–S7.

[74] John J.V. McMurray et al. “Angiotensin–Neprilysin Inhibition versus Enalapril in Heart Failure”. In: *New England Journal of Medicine* 371.11 (Sept. 2014), pp. 993–1004.

[75] Bertram Pitt et al. “The Effect of Spironolactone on Morbidity and Mortality in Patients with Severe Heart Failure”. In: *New England Journal of Medicine* 341.10 (Sept. 1999), pp. 709–717.

[76] Bertram Pitt et al. “The EPHESUS Trial: Eplerenone in Patients with Heart Failure Due to Systolic Dysfunction Complicating Acute Myocardial Infarction”. In: *Cardiovascular Drugs and Therapy* 15.1 (Jan. 2001), pp. 79–87.

[77] Sanjiv J. Shah et al. “Phenomapping for Novel Classification of Heart Failure With Preserved Ejection Fraction”. In: *Circulation* 131.3 (Jan. 2015), pp. 269–279.

[78] Rebecca J. Woolley et al. “Machine learning based on biomarker profiles identifies distinct subgroups of heart failure with preserved ejection fraction”. In: *European Journal of Heart Failure* 23.6 (June 2021), pp. 983–991.

[79] Tasha Nagamine et al. “Multiscale classification of heart failure phenotypes by unsupervised clustering of unstructured electronic medical record data”. In: *Scientific Reports* 10.1 (Dec. 2020), p. 21340.

[80] Matthew W. Segar et al. “Machine Learning to Predict the Risk of Incident Heart Failure Hospitalization Among Patients With Diabetes: The WATCH-DM Risk Score”. In: *Diabetes Care* 42.12 (Dec. 2019), pp. 2298–2306.

[81] Andreas B. Gevaert et al. “Clinical phenogroups are more effective than left ventricular ejection fraction categories in stratifying heart failure outcomes”. In: *ESC Heart Failure* 8.4 (Aug. 2021), pp. 2741–2754.

[82] Jennifer E. Ho et al. “Predicting Heart Failure With Preserved and Reduced Ejection Fraction”. In: *Circulation: Heart Failure* 9.6 (June 2016).

[83] Dominik Jenča et al. “Heart failure after myocardial infarction: incidence and predictors”. In: *ESC Heart Failure* 8.1 (Feb. 2021), pp. 222–237.

[84] Thomas J Cahill and Rajesh K Kharbanda. “Heart failure after myocardial infarction in the era of primary percutaneous coronary intervention: Mechanisms, incidence and identification of patients at risk”. In: *World Journal of Cardiology* 9.5 (2017), p. 407.

[85] Tripti Rastogi et al. “Comparing and contrasting risk factors for heart failure in patients with and without history of myocardial infarction: data from HOMAGE and the UK Biobank”. In: *European Journal of Heart Failure* 24.6 (June 2022), pp. 976–984.

[86] David P. Kao et al. “Characterization of subgroups of heart failure patients with preserved ejection fraction with possible implications for prognosis and treatment response”. In: *European Journal of Heart Failure* 17.9 (Sept. 2015), pp. 925–935.

[87] Manting Choy et al. “Phenotypes of heart failure with preserved ejection fraction and effect of spironolactone treatment”. In: *ESC Heart Failure* 9.4 (Aug. 2022), pp. 2567–2575.

[88] Jordana B. Cohen et al. “Clinical Phenogroups in Heart Failure With Preserved Ejection Fraction”. In: *JACC: Heart Failure* 8.3 (Mar. 2020), pp. 172–184.

[89] Brian R. Lindman et al. “Cardiovascular Phenotype in HFpEF Patients With or Without Diabetes”. In: *Journal of the American College of Cardiology* 64.6 (Aug. 2014), pp. 541–549.

[90] Selma F. Mohammed et al. “Comorbidity and Ventricular and Vascular Structure and Function in Heart Failure With Preserved Ejection Fraction”. In: *Circulation: Heart Failure* 5.6 (Nov. 2012), pp. 710–719.

[91] Masaru Obokata et al. “Evidence Supporting the Existence of a Distinct Obese Phenotype of Heart Failure With Preserved Ejection Fraction”. In: *Circulation* 136.1 (July 2017), pp. 6–19.

[92] Tonje Amb Aksnes et al. “Impact of New-Onset Diabetes Mellitus on Development of Atrial Fibrillation and Heart Failure in High-Risk Hypertension (from the VALUE Trial)”. In: *The American Journal of Cardiology* 101.5 (Mar. 2008), pp. 634–638.

[93] E J Benjamin et al. “Independent risk factors for atrial fibrillation in a population-based cohort. The Framingham Heart Study.” In: *JAMA* 271.11 (Mar. 1994), pp. 840–4.

[94] Rachel R Huxley et al. “Type 2 diabetes, glucose homeostasis and incident atrial fibrillation: the Atherosclerosis Risk in Communities study”. In: *Heart* 98.2 (Jan. 2012), pp. 133–138.

[95] Praloy Chakraborty et al. “Sudden cardiac death due to ventricular arrhythmia in diabetes mellitus: A bench to bedside review”. In: *Heart Rhythm* 21.10 (Oct. 2024), pp. 1827–1837.

[96] Christiane Jungen et al. “Increased arrhythmia susceptibility in type 2 diabetic mice related to dysregulation of ventricular sympathetic innervation”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 317.6 (Dec. 2019), H1328–H1341.

[97] André J. Scheen. “Glucose-lowering agents and risk of ventricular arrhythmias and sudden cardiac death: A comprehensive review ranging from sulphonylureas to SGLT2 inhibitors”. In: *Diabetes & Metabolism* 48.6 (Nov. 2022), p. 101405.

[98] M. R. MacDonald et al. “Impact of diabetes on outcomes in patients with low and preserved ejection fraction heart failure: An analysis of the Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM) programme”. In: *European Heart Journal* 29.11 (Jan. 2008), pp. 1377–1385.

[99] Robert J. Mentz et al. “Noncardiac Comorbidities in Heart Failure With Reduced Versus Preserved Ejection Fraction”. In: *Journal of the American College of Cardiology* 64.21 (Dec. 2014), pp. 2281–2293.

[100] Martin H. Ruwald et al. “Influence of Diabetes Mellitus on Inappropriate and Appropriate Implantable Cardioverter-Defibrillator Therapy and Mortality in the Multicenter Automatic Defibrillator Implantation Trial–Reduce Inappropriate Therapy (MADIT-RIT) Trial”. In: *Circulation* 128.7 (Aug. 2013), pp. 694–701.

[101] Nisha Arenja et al. “Prevalence, Extent, and Independent Predictors of Silent Myocardial Infarction”. In: *The American Journal of Medicine* 126.6 (June 2013), pp. 515–522.

[102] Matthew J. Singleton et al. “QRS duration is associated with all-cause mortality in type 2 diabetes: The diabetes heart study”. In: *Journal of Electrocardiology* 58 (Jan. 2020), pp. 150–154.

[103] Hirofumi Soejima et al. “One quarter of total myocardial infarctions are silent manifestation in patients with type 2 diabetes mellitus”. In: *Journal of Cardiology* 73.1 (Jan. 2019), pp. 33–37.

[104] Bulent Gorenek et al. “European Heart Rhythm Association (EHRA) position paper on arrhythmia management and device therapies in endocrine disorders, endorsed by Asia Pacific Heart Rhythm Society (APHRS) and Latin American Heart Rhythm Society (LAHRS)”. In: *EP Europace* 20.6 (June 2018), pp. 895–896.

[105] Shirley Rubler et al. “New type of cardiomyopathy associated with diabetic glomerulosclerosis”. In: *The American Journal of Cardiology* 30.6 (Nov. 1972), pp. 595–602.

[106] Petar M. Seferović et al. “Diabetic myocardial disorder. A clinical consensus statement of the Heart Failure Association of the ESC and the ESC Working Group on Myocardial & Pericardial Diseases”. In: *European Journal of Heart Failure* (June 2024).

[107] Richard B. Devereux et al. “Impact of Diabetes on Cardiac Structure and Function”. In: *Circulation* 101.19 (May 2000), pp. 2271–2276.

[108] Kazuo Eguchi et al. “Association Between Diabetes Mellitus and Left Ventricular Hypertrophy in a Multiethnic Population”. In: *The American Journal of Cardiology* 101.12 (June 2008), pp. 1787–1791.

[109] Marshall Lee et al. “Diabetes mellitus and echocardiographic left ventricular function in free-living elderly men and women: The Cardiovascular Health Study”. In: *American Heart Journal* 133.1 (Jan. 1997), pp. 36–43.

[110] Nader Parsa et al. “The prevalence of left ventricular hypertrophy associated with type-2 diabetes in Shiraz, Iran: a cross-sectional study”. In: *BMC Cardiovascular Disorders* 23.1 (Feb. 2023), p. 88.

[111] Eylem Levelt et al. “Relationship Between Left Ventricular Structural and Metabolic Remodeling in Type 2 Diabetes”. In: *Diabetes* 65.1 (Jan. 2016), pp. 44–52.

[112] Stuart A. Cook et al. “Abnormal myocardial insulin signalling in type 2 diabetes and left-ventricular dysfunction”. In: *European Heart Journal* 31.1 (Jan. 2010), pp. 100–111.

[113] K Karason et al. “Impact of blood pressure and insulin on the relationship between body fat and left ventricular structure”. In: *European Heart Journal* 24.16 (Aug. 2003), pp. 1500–1505.

[114] Carla Giordano et al. “Myocardial fibrosis: morphologic patterns and role of imaging in diagnosis and prognostication”. In: *Cardiovascular Pathology* 56 (Jan. 2022), p. 107391.

[115] Doron Aronson. “Cross-linking of glycated collagen in the pathogenesis of arterial and myocardial stiffening of aging and diabetes”. In: *Journal of Hypertension* 21.1 (Jan. 2003), pp. 3–12.

[116] Joseph M Pappachan et al. “Diabetic cardiomyopathy: Pathophysiology, diagnostic evaluation and management”. In: *World Journal of Diabetes* 4.5 (2013), p. 177.

[117] Andrew A. Gibb, Michael P. Lazaropoulos, and John W. Elrod. “Myofibroblasts and Fibrosis”. In: *Circulation Research* 127.3 (July 2020), pp. 427–447.

[118] Guanghong Jia, Michael A. Hill, and James R. Sowers. “Diabetic Cardiomyopathy”. In: *Circulation Research* 122.4 (Feb. 2018), pp. 624–638.

[119] Guanghong Jia, Vincent G. DeMarco, and James R. Sowers. “Insulin resistance and hyperinsulinaemia in diabetic cardiomyopathy”. In: *Nature Reviews Endocrinology* 12.3 (Mar. 2016), pp. 144–153.

[120] M Shimizu et al. “Collagen remodelling in myocardia of patients with diabetes.” In: *Journal of Clinical Pathology* 46.1 (Jan. 1993), pp. 32–36.

[121] Kirsten H.W.J. Ten Tusscher and Alexander V. Panfilov. “Influence of diffuse fibrosis on wave propagation in human ventricular tissue”. In: *EP Europace* 9.suppl_6 (Nov. 2007), pp. vi38–vi45.

[122] M A Rossi. “Connective tissue skeleton in the normal left ventricle and in hypertensive left ventricular hypertrophy and chronic chagasic myocarditis.” In: *Medical science monitor : international medical journal of experimental and clinical research* 7.4 (2001), pp. 820–32.

[123] Craig S. Broberg et al. “Quantification of Diffuse Myocardial Fibrosis and Its Association With Myocardial Dysfunction in Congenital Heart Disease”. In: *Circulation: Cardiovascular Imaging* 3.6 (Nov. 2010), pp. 727–734.

[124] Chee Jian Pua et al. “Impact of Diabetes on Myocardial Fibrosis in Patients With Hypertension: The REMODEL Study”. In: *Circulation: Cardiovascular Imaging* 16.7 (July 2023), pp. 545–553.

[125] A Bojer et al. “Regional distribution of diffuse fibrosis in the basal and mid-ventricular AHA segments in patients with type 2 diabetes”. In: *European Heart Journal* 45.Supplement_1 (Oct. 2024).

[126] Yaron Fridman et al. “Myocardial fibrosis is associated with subsequent death and hospitalization for heart failure in obese adults”. In: *Journal of Cardiovascular Magnetic Resonance* 17 (Feb. 2015), p. M8.

[127] Prathap Kanagala et al. “Relationship Between Focal and Diffuse Fibrosis Assessed by CMR and Clinical Outcomes in Heart Failure With Preserved Ejection Fraction”. In: *JACC: Cardiovascular Imaging* 12.11 (Nov. 2019), pp. 2291–2301.

[128] Mark Dennis et al. “Myocardial fibrosis in Type 2 Diabetes is associated with functional and metabolomic parameters”. In: *International Journal of Cardiology* 363 (Sept. 2022), pp. 179–184.

[129] Izabela Tuleta and Nikolaos G. Frangogiannis. “Fibrosis of the diabetic heart: Clinical significance, molecular mechanisms, and therapeutic opportunities”. In: *Advanced Drug Delivery Reviews* 176 (Sept. 2021), p. 113904.

[130] S. W. Rabkin. “Epicardial fat: properties, function and relationship to obesity”. In: *Obesity Reviews* 8.3 (May 2007), pp. 253–261.

[131] Arnold C.T. Ng et al. “Impact of Epicardial Adipose Tissue, Left Ventricular Myocardial Fat Content, and Interstitial Fibrosis on Myocardial Contractile Function”. In: *Circulation: Cardiovascular Imaging* 11.8 (Aug. 2018).

[132] Auriane C. Ernault, Veronique M.F. Meijborg, and Ruben Coronel. “Modulation of Cardiac Arrhythmogenesis by Epicardial Adipose Tissue”. In: *Journal of the American College of Cardiology* 78.17 (Oct. 2021), pp. 1730–1745.

[133] Yi Tan et al. “Mechanisms of diabetic cardiomyopathy and potential therapeutic strategies: preclinical and clinical evidence”. In: *Nature Reviews Cardiology* 17.9 (Sept. 2020), pp. 585–607.

[134] Michele Miragoli, Giedrius Gaudesius, and Stephan Rohr. “Electrotonic Modulation of Cardiac Impulse Conduction by Myofibroblasts”. In: *Circulation Research* 98.6 (Mar. 2006), pp. 801–810.

[135] A. Nygren et al. “Propagation of the cardiac impulse in the diabetic rat heart: reduced conduction reserve”. In: *The Journal of Physiology* 580.2 (Apr. 2007), pp. 543–560.

[136] Hai Lin et al. “Alterations of Connexin 43 in the Diabetic Rat Heart”. In: *Cardiovascular Gap Junctions*. Basel: KARGER, 2006, pp. 243–254.

[137] Deniz Billur, Yusuf Olgar, and Belma Turan. “Intracellular Redistribution of Left Ventricular Connexin 43 Contributes to the Remodeling of Electrical Properties of the Heart in Insulin-resistant Elderly Rats”. In: *Journal of Histochemistry & Cytochemistry* 70.6 (June 2022), pp. 447–462.

[138] Emmanuel Dupont et al. “Altered Connexin Expression in Human Congestive Heart Failure”. In: *Journal of Molecular and Cellular Cardiology* 33.2 (Feb. 2001), pp. 359–371.

[139] Mohamed Boulaksil et al. “Spatial Heterogeneity of Cx43 is an Arrhythmogenic Substrate of Polymorphic Ventricular Tachycardias during Compensated Cardiac Hypertrophy in Rats”. In: *Frontiers in Cardiovascular Medicine* 3 (Mar. 2016).

[140] N. J. Severs et al. “Remodelling of gap junctions and connexin expression in diseased myocardium”. In: *Cardiovascular Research* 80.1 (Aug. 2008), pp. 9–19.

[141] Jiajie Yan et al. “Molecular remodeling of Cx43, but not structural remodeling, promotes arrhythmias in an arrhythmogenic canine model of nonischemic heart failure”. In: *Journal of Molecular and Cellular Cardiology* 158 (Sept. 2021), pp. 72–81.

[142] Jae Hyung Cho. “Sudden Death and Ventricular Arrhythmias in Heart Failure With Preserved Ejection Fraction”. In: *Korean Circulation Journal* 52.4 (2022), p. 251.

[143] Marcelo F. Di Carli et al. “Effects of Autonomic Neuropathy on Coronary Blood Flow in Patients With Diabetes Mellitus”. In: *Circulation* 100.8 (Aug. 1999), pp. 813–819.

[144] Jeffrey J. Goldberger et al. “Autonomic Nervous System Dysfunction”. In: *Journal of the American College of Cardiology* 73.10 (Mar. 2019), pp. 1189–1206.

[145] Giuseppe Giannino et al. “The Intrinsic Cardiac Nervous System: From Pathophysiology to Therapeutic Implications”. In: *Biology* 13.2 (Feb. 2024), p. 105.

[146] J. Andrew Armour et al. “Gross and microscopic anatomy of the human intrinsic cardiac nervous system”. In: *The Anatomical Record* 247.2 (Feb. 1997), pp. 289–298.

[147] Dainius H. Pauza et al. “Morphology, distribution, and variability of the epicardiac neural ganglionated subplexuses in the human heart”. In: *The Anatomical Record* 259.4 (Aug. 2000), pp. 353–382.

[148] Anthony J. Evans and Yu-Long Li. “Remodeling of the Intracardiac Ganglia During the Development of Cardiovascular Autonomic Dysfunction in Type 2 Diabetes: Molecular Mechanisms and Therapeutics”. In: *International Journal of Molecular Sciences* 25.22 (Nov. 2024), p. 12464.

[149] A. I. Vinik, R. E. Maser, and D. Ziegler. “Autonomic imbalance: prophet of doom or scope for hope?” In: *Diabetic Medicine* 28.6 (June 2011), pp. 643–651.

[150] M S Draman et al. “A silent myocardial infarction in the diabetes outpatient clinic: case report and review of the literature”. In: *Endocrinology, Diabetes & Metabolism Case Reports* 2013 (Oct. 2013).

[151] Aaron I Vinik, Tomris Erbas, and Carolina M Casellini. “Diabetic cardiac autonomic neuropathy, inflammation and cardiovascular disease”. In: *Journal of Diabetes Investigation* 4.1 (Jan. 2013), pp. 4–18.

[152] J. Mustonen et al. “Autonomic nervous function and its relationship to cardiac performance in middle-aged diabetic patients without clinically evident cardiovascular disease”. In: *Journal of Internal Medicine* 232.1 (July 1992), pp. 65–72.

[153] P. Cavallo Perin, S. Maule, and R. Quadri. “SYMPATHETIC NERVOUS SYSTEM, DIABETES, AND HYPERTENSION”. In: *Clinical and Experimental Hypertension* 23.1-2 (Jan. 2001), pp. 45–55.

[154] Peter Eriksson et al. “Bundle-Branch Block in a General Male Population”. In: *Circulation* 98.22 (Dec. 1998), pp. 2494–2500.

[155] Saranda Haxha et al. “Type 2 diabetes mellitus and higher rate of complete atrioventricular block: a Danish Nationwide Registry”. In: *European Heart Journal* 44.9 (Mar. 2023), pp. 752–761.

[156] Mohammad Reza Movahed et al. “Strong independent association between third-degree AV-block and diabetes mellitus using a large database”. In: *Diabetes Research and Clinical Practice* 205 (Nov. 2023), p. 110948.

[157] H. Sophia Chen et al. “Ventricular Arrhythmia Substrate Distribution and Its Relation to Sympathetic Innervation in Nonischemic Cardiomyopathy Patients”. In: *JACC: Clinical Electrophysiology* 8.10 (Oct. 2022), pp. 1234–1245.

[158] Valentina Parisi et al. “Increased Epicardial Adipose Tissue Volume Correlates With Cardiac Sympathetic Denervation in Patients With Heart Failure”. In: *Circulation Research* 118.8 (Apr. 2016), pp. 1244–1253.

[159] Peter Vargovic et al. “Adipocytes as a new source of catecholamine production”. In: *FEBS Letters* 585.14 (July 2011), pp. 2279–2284.

[160] J M Pappachan et al. “Cardiac autonomic neuropathy in diabetes mellitus: prevalence, risk factors and utility of corrected QT interval in the ECG for its diagnosis”. In: *Postgraduate Medical Journal* 84.990 (Apr. 2008), pp. 205–210.

[161] R. Ashrafi et al. “Arrhythmogenic gene remodelling in elderly patients with type 2 diabetes with aortic stenosis and normal left ventricular ejection fraction”. In: *Experimental Physiology* 102.11 (Nov. 2017), pp. 1424–1434.

[162] Mónica Gallego et al. “Electrical Features of the Diabetic Myocardium. Arrhythmic and Cardiovascular Safety Considerations in Diabetes”. In: *Frontiers in Pharmacology* 12 (July 2021).

[163] Laetitia Pereira et al. “Mechanisms of [Ca2+]i Transient Decrease in Cardiomyopathy of db/db Type 2 Diabetic Mice”. In: *Diabetes* 55.3 (Mar. 2006), pp. 608–615.

[164] Zhongju Lu et al. “Increased Persistent Sodium Current Due to Decreased PI3K Signaling Contributes to QT Prolongation in the Diabetic Heart”. In: *Diabetes* 62.12 (Dec. 2013), pp. 4257–4265.

[165] Yuichi Hattori et al. “Diminished function and expression of the cardiac Na+-Ca2+ exchanger in diabetic rats: implication in Ca2+ overload”. In: *The Journal of Physiology* 527.1 (Aug. 2000), pp. 85–94.

[166] Yasushi Teshima et al. “Diminished Expression of Sarcoplasmic Reticulum Ca2+-ATPase and Ryanodine Sensitive Ca2+Channel mRNA in Streptozotocin-induced Diabetic Rat Heart”. In: *Journal of Molecular and Cellular Cardiology* 32.4 (Apr. 2000), pp. 655–664.

[167] Thomas Netticadan et al. “Depressed Levels of Ca2+-Cycling Proteins May Underlie Sarcoplasmic Reticulum Dysfunction in the Diabetic Heart”. In: *Diabetes* 50.9 (Sept. 2001), pp. 2133–2138.

[168] Darrell D. Belke, Eric A. Swanson, and Wolfgang H. Dillmann. “Decreased Sarcoplasmic Reticulum Activity and Contractility in Diabetic db/db Mouse Heart”. In: *Diabetes* 53.12 (Dec. 2004), pp. 3201–3208.

[169] Lin Zhang et al. “Altered Calcium Homeostasis Does Not Explain the Contractile Deficit of Diabetic Cardiomyopathy”. In: *Diabetes* 57.8 (Aug. 2008), pp. 2158–2166.

[170] Lorna J. Daniels et al. “Inhibition of calcium/calmodulin-dependent kinase II restores contraction and relaxation in isolated cardiac muscle from type 2 diabetic rats”. In: *Cardiovascular Diabetology* 17.1 (Dec. 2018), p. 89.

[171] Jeffrey R. Erickson et al. “Diabetic hyperglycaemia activates CaMKII and arrhythmias by O-linked glycosylation”. In: *Nature* 502.7471 (Oct. 2013), pp. 372–376.

[172] Bence Hegyi, Donald M. Bers, and Julie Bossuyt. “CaMKII signaling in heart diseases: Emerging role in diabetic cardiomyopathy”. In: *Journal of Molecular and Cellular Cardiology* 127 (Feb. 2019), pp. 246–259.

[173] Stefan Wagner et al. “Ca2+/calmodulin-dependent protein kinase II regulates cardiac Na+ channels”. In: *Journal of Clinical Investigation* 116.12 (Dec. 2006), pp. 3127–3138.

[174] Sanda Despa and Donald M. Bers. “Na+ transport in the normal and failing heart — Remember the balance”. In: *Journal of Molecular and Cellular Cardiology* 61 (Aug. 2013), pp. 2–10.

[175] Donald M. Bers and Stefano Morotti. “Ca2+ current facilitation is CaMKII-dependent and has arrhythmogenic consequences”. In: *Frontiers in Pharmacology* 5 (June 2014).

[176] W. Yuan and D. M. Bers. “Ca-dependent facilitation of cardiac Ca current is due to Ca-calmodulin-dependent protein kinase”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 267.3 (Sept. 1994), H982–H993.

[177] Tomas O. Stolen et al. “Interval Training Normalizes Cardiomyocyte Function, Diastolic Ca2+ Control, and SR Ca2+ Release Synchronicity in a Mouse Model of Diabetic Cardiomyopathy”. In: *Circulation Research* 105.6 (Sept. 2009), pp. 527–536.

[178] Iuliana Popescu et al. “Lower sarcoplasmic reticulum Ca2+ threshold for triggering afterdepolarizations in diabetic rat hearts”. In: *Heart Rhythm* 16.5 (May 2019), pp. 765–772.

[179] Matthew W. Ebinger, Subramaniam Krishnan, and Claudio D. Schuger. “Mechanisms of ventricular arrhythmias in heart failure”. In: *Current Heart Failure Reports* 2.3 (Sept. 2005), pp. 111–117.

[180] S Pogwizd. “Intracellular Na in animal models of hypertrophy and heart failure: contractile function and arrhythmogenesis”. In: *Cardiovascular Research* 57.4 (Mar. 2003), pp. 887–896.

[181] Zhongju Lu et al. “Restoration of Defective L-type Ca2+ Current in Cardiac Myocytes of Type 2 Diabetic db/db Mice by Akt and PKC-ι”. In: *Journal of Cardiovascular Pharmacology* 58.4 (Oct. 2011), pp. 439–445.

[182] J Magyar et al. “Action potentials and potassium currents in rat ventricular muscle during experimental diabetes”. In: *Journal of Molecular and Cellular Cardiology* 24.8 (Aug. 1992), pp. 841–853.

[183] Tatsuya Sato et al. “Type 2 diabetes induces subendocardium-predominant reduction in transient outward K ⁺ current with downregulation of Kv4.2 and KChIP2”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 306.7 (Apr. 2014), H1054–H1065.

[184] Yakhin Shimoni et al. “Gender-dependent attenuation of cardiac potassium currents in type 2 diabetic db/db mice”. In: *The Journal of Physiology* 555.2 (Mar. 2004), pp. 345–354.

[185] K. Tsuchida, H. Watajima, and S. Otomo. “Calcium current in rat diabetic ventricular myocytes”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 267.6 (Dec. 1994), H2280–H2289.

[186] Katsuharu Tsuchida and Hiroshi Watajima. “Potassium currents in ventricular myocytes from genetically diabetic rats”. In: *American Journal of Physiology-Endocrinology and Metabolism* 273.4 (Oct. 1997), E695–E700.

[187] Y Shimoni, D Severson, and W Giles. “Thyroid status and diabetes modulate regional differences in potassium currents in rat ventricle.” In: *The Journal of Physiology* 488.3 (Nov. 1995), pp. 673–688.

[188] Gilberta Giacchetti et al. “The renin–angiotensin–aldosterone system, glucose metabolism and diabetes”. In: *Trends in Endocrinology & Metabolism* 16.3 (Apr. 2005), pp. 120–126.

[189] R Caballero et al. “Interaction of angiotensin II with the angiotensin type 2 receptor inhibits the cardiac transient outward potassium current”. In: *Cardiovascular Research* 62.1 (Apr. 2004), pp. 86–95.

[190] Ting-Ting Zhang et al. “Independent Regulation of Cardiac Kv4.3 Potassium Channel Expression by Angiotensin II and Phenylephrine”. In: *Circulation Research* 88.5 (Mar. 2001), pp. 476–482.

[191] Yakhin Shimoni. “Inhibition of the formation or action of angiotensin II reverses attenuated K ⁺ currents in type 1 and type 2 diabetes”. In: *The Journal of Physiology* 537.1 (Nov. 2001), pp. 83–92.

[192] Julian Zayas-Arrabal et al. “Molecular and Electrophysiological Role of Diabetes-Associated Circulating Inflammatory Factors in Cardiac Arrhythmia Remodeling in a Metabolic-Induced Model of Type 2 Diabetic Rat”. In: *International Journal of Molecular Sciences* 22.13 (June 2021), p. 6827.

[193] Siyavash Joukar. “A comparative review on heart ion channels, action potentials and electrocardiogram in rodents and human: extrapolation of experimental insights to clinic”. In: *Laboratory Animal Research* 37.1 (Sept. 2021), p. 25.

[194] Thomas Hadberg Lynge et al. “Sudden cardiac death among persons with diabetes aged 1–49 years: a 10-year nationwide study of 14294 deaths in Denmark”. In: *European Heart Journal* 41.28 (July 2020), pp. 2699–2706.

[195] Graham Rena, D. Grahame Hardie, and Ewan R. Pearson. “The mechanisms of action of metformin”. In: *Diabetologia* 60.9 (Sept. 2017), pp. 1577–1585.

[196] Lorraine Mascarenhas et al. “Antiarrhythmic effects of metformin”. In: *Heart Rhythm O2* 5.5 (May 2024), pp. 310–320.

[197] Eunice Cristina da Silva Costa et al. “Os efeitos da metformina sobre a dispersão do intervalo QT e QTc de ratos diabéticos”. In: *Arquivos Brasileiros de Cardiologia* 90.4 (Apr. 2008), pp. 254–260.

[198] Hui Wang et al. “Metformin Shortens Prolonged QT Interval in Diabetic Mice by Inhibiting L-Type Calcium Current: A Possible Therapeutic Approach”. In: *Frontiers in Pharmacology* 11 (June 2020).

[199] Layse Malagueta-Vieira et al. “Metformin Reduces Potassium Currents and Prolongs Repolarization in Non-Diabetic Heart”. In: *International Journal of Molecular Sciences* 23.11 (May 2022), p. 6021.

[200] Hung-Yi Chen et al. “Antihyperglycemic drugs use and new-onset atrial fibrillation in elderly patients”. In: *European Journal of Clinical Investigation* 47.5 (May 2017), pp. 388–393.

[201] Yi-Sheng Liou et al. “Antihyperglycemic drugs use and new-onset atrial fibrillation: A population-based nested case control study”. In: *PLOS ONE* 13.8 (Aug. 2018), e0197245.

[202] David Leak and Paul Starr. “The mechanism of arrhythmias during insulin-induced hypoglycemia”. In: *American Heart Journal* 63.5 (May 1962), pp. 688–691.

[203] Caglar Gök, Alan D. Robertson, and William Fuller. “Insulin-induced palmitoylation regulates the Cardiac Na+/Ca2+ exchanger NCX1”. In: *Cell Calcium* 104 (June 2022), p. 102567.

[204] American Diabetes Association Professional Practice Committee. “9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes—2022”. In: *Diabetes Care* 45.Supplement_1 (Jan. 2022), S125–S143.

[205] James P Curtain et al. “Effect of dapagliflozin on ventricular arrhythmias, resuscitated cardiac arrest, or sudden death in DAPA-HF”. In: *European Heart Journal* 42.36 (Sept. 2021), pp. 3727–3738.

[206] Gilson C. Fernandes et al. “Association of SGLT2 inhibitors with arrhythmias and sudden cardiac death in patients with type 2 diabetes or heart failure: A meta-analysis of 34 randomized controlled trials”. In: *Heart Rhythm* 18.7 (July 2021), pp. 1098–1105.

[207] Jia Liao et al. “Effect of SGLT-2 inhibitors on arrhythmia events: insight from an updated secondary analysis of>80,000 patients (the SGLT2i—Arrhythmias and Sudden Cardiac Death)”. In: *Cardiovascular Diabetology* 23.1 (Feb. 2024), p. 78.

[208] Laween Uthman et al. “Class effects of SGLT2 inhibitors in mouse cardiomyocytes and hearts: inhibition of Na+/H+ exchanger, lowering of cytosolic Na+ and vasodilation”. In: *Diabetologia* 61.3 (Mar. 2018), pp. 722–726.

[209] S. K. Banerjee et al. “SGLT1 is a novel cardiac glucose transporter that is perturbed in disease states”. In: *Cardiovascular Research* 84.1 (Oct. 2009), pp. 111–118.

[210] Kwong-Man Ng et al. “Empagliflozin Ammeliorates High Glucose Induced-Cardiac Dysfuntion in Human iPSC-Derived Cardiomyocytes”. In: *Scientific Reports* 8.1 (Oct. 2018), p. 14872.

[211] Koenraad Philippaert et al. “Cardiac Late Sodium Channel Current Is a Molecular Target for the Sodium/Glucose Cotransporter 2 Inhibitor Empagliflozin”. In: *Circulation* 143.22 (June 2021), pp. 2188–2204.

[212] Bence Hegyi et al. “Empagliflozin Reverses Late Na ⁺ Current Enhancement and Cardiomyocyte Proarrhythmia in a Translational Murine Model of Heart Failure With Preserved Ejection Fraction”. In: *Circulation* 145.13 (Mar. 2022), pp. 1029–1031.

[213] Takahide Kadosaka et al. “Empagliflozin attenuates arrhythmogenesis in diabetic cardiomyopathy by normalizing intracellular Ca ²⁺ handling in ventricular cardiomyocytes”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 324.3 (Mar. 2023), H341–H354.

[214] Sean Kang et al. “Direct Effects of Empagliflozin on Extracellular Matrix Remodelling in Human Cardiac Myofibroblasts: Novel Translational Clues to Explain EMPA-REG OUTCOME Results”. In: *Canadian Journal of Cardiology* 36.4 (Apr. 2020), pp. 543–553.

[215] Hsiang-Chun Lee et al. “The sodium–glucose co-transporter 2 inhibitor empagliflozin attenuates cardiac fibrosis and improves ventricular hemodynamics in hypertensive heart failure rats”. In: *Cardiovascular Diabetology* 18.1 (Dec. 2019), p. 45.

[216] Carlos G. Santos-Gallego et al. “Empagliflozin Ameliorates Adverse Left Ventricular Remodeling in Nondiabetic Heart Failure by Enhancing Myocardial Energetics”. In: *Journal of the American College of Cardiology* 73.15 (Apr. 2019), pp. 1931–1944.

[217] Brian Tomlinson et al. “The role of sulfonylureas in the treatment of type 2 diabetes”. In: *Expert Opinion on Pharmacotherapy* 23.3 (Feb. 2022), pp. 387–403.

[218] Nehal Islam et al. “Sulphonylureas versus metformin and the risk of ventricular arrhythmias among people with type 2 diabetes: A population-based cohort study”. In: *Diabetes, Obesity and Metabolism* 25.6 (June 2023), pp. 1523–1533.

[219] Teddy Tai Loy Lee et al. “Sulfonylurea Is Associated With Higher Risks of Ventricular Arrhythmia or Sudden Cardiac Death Compared With Metformin: A Population-Based Cohort Study”. In: *Journal of the American Heart Association* 11.18 (Sept. 2022).

[220] Jiandong Zhou et al. “Metformin versus sulphonylureas for new onset atrial fibrillation and stroke in type 2 diabetes mellitus: a population-based study”. In: *Acta Diabetologica* 59.5 (Feb. 2022), pp. 697–709.

[221] Syed A Najeed et al. “Differential effect of glyburide (glibenclamide) and metformin on qt dispersion: a potential adenosine triphosphate sensitive k+ channel effect”. In: *The American Journal of Cardiology* 90.10 (Nov. 2002), pp. 1103–1106.

[222] Barbara Rosati et al. “Sulfonylureas blockade of neural and cardiac HERG channels”. In: *FEBS Letters* 440.1-2 (Nov. 1998), pp. 125–130.

[223] Bojan Vrtovec et al. “Atorvastatin Therapy Increases Heart Rate Variability, Decreases QT Variability, and Shortens QTc Interval Duration in Patients With Advanced Chronic Heart Failure”. In: *Journal of Cardiac Failure* 11.9 (Dec. 2005), pp. 684–690.

[224] Rui-qin Xie et al. “Statin therapy shortens QTc, QTcd, and improves cardiac function in patients with chronic heart failure”. In: *International Journal of Cardiology* 140.2 (Apr. 2010), pp. 255–257.

[225] Bojan Vrtovec et al. “Statin-associated QTc Interval Shortening as Prognostic Indicator in Heart Transplant Recipients”. In: *The Journal of Heart and Lung Transplantation* 25.2 (Feb. 2006), pp. 234–236.

[226] Abdullah Tekin et al. “Short- and long-term effect of simvastatin therapy on the heterogeneity of cardiac repolarization in diabetic patients”. In: *Pharmacological Research* 57.5 (May 2008), pp. 393–397.

[227] Yeryung Koo et al. “Evaluation of rosuvastatin-induced QT prolongation risk using real-world data, in vitro cardiomyocyte studies, and mortality assessment”. In: *Scientific Reports* 13.1 (May 2023), p. 8108.

[228] Isabelle Plante et al. “Rosuvastatin Blocks hERG Current and Prolongs Cardiac Repolarization”. In: *Journal of Pharmaceutical Sciences* 101.2 (Feb. 2012), pp. 868–878.

[229] Cathie Sudlow et al. “UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age”. In: *PLOS Medicine* 12.3 (Mar. 2015), e1001779.

[230] Steffen E. Petersen et al. “UK Biobank’s cardiovascular magnetic resonance protocol”. In: *Journal of Cardiovascular Magnetic Resonance* 18.1 (Jan. 2016), p. 8.

[231] Avan Suinesiaputra et al. “Fully-automated left ventricular mass and volume MRI analysis in the UK Biobank population cohort: evaluation of initial results”. In: *The International Journal of Cardiovascular Imaging* 34.2 (Feb. 2018), pp. 281–291.

[232] Wenjia Bai et al. “A population-based phenome-wide association study of cardiac and aortic structure and function”. In: *Nature Medicine* 26.10 (Oct. 2020), pp. 1654–1662.

[233] Steffen E. Petersen et al. “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* 19.1 (Dec. 2016), p. 18.

[234] Luca Biasiolli et al. “Automated localization and quality control of the aorta in cine CMR can significantly accelerate processing of the UK Biobank population data”. In: *PLOS ONE* 14.2 (Feb. 2019), e0212272.

[235] Edward Ferdian et al. “Fully Automated Myocardial Strain Estimation from Cardiovascular MRI–tagged Images Using a Deep Learning Framework in the UK Biobank”. In: *Radiology: Cardiothoracic Imaging* 2.1 (Feb. 2020), e190032.

[236] Andrew Bard et al. “Automated Quality-Controlled Cardiovascular Magnetic Resonance Pericardial Fat Quantification Using a Convolutional Neural Network in the UK Biobank”. In: *Frontiers in Cardiovascular Medicine* 8 (July 2021).

[237] Nay Aung et al. “Genome-wide association analysis reveals insights into the genetic architecture of right ventricular structure and function”. In: *Nature Genetics* 54.6 (June 2022), pp. 783–791.

[238] Devran Ugurlu et al. “Cardiac digital twins at scale from MRI: Open tools and representative models from
55000 UK Biobank participants”. In: *PLOS One* 20.7 (July 2025), e0327158.

[239] Thomas J. Littlejohns et al. “The UK Biobank imaging enhancement of 100,000 participants:rationale, data collection, management and future directions”. In: *Nature Communications* 11.1 (May 2020), p. 2624.

[240] M. Austin Argentieri et al. “Integrating the environmental and genetic architectures of aging and mortality”. In: *Nature Medicine* 31.3 (Mar. 2025), pp. 1016–1025.

[241] Miquel Porta. *A Dictionary of Epidemiology*. Oxford University Press, Jan. 2008.

[242] Kenneth J.. Rothman, Sander. Greenland, and Timothy L.. Lash. *Modern epidemiology*. Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008, p. 758.

[243] Jorge Corral-Acero et al. “The ‘Digital Twin’ to enable the vision of precision cardiology”. In: *European Heart Journal* 41.48 (Dec. 2020), pp. 4556–4564.

[244] Angela W.C. Lee et al. “Non-invasive simulated electrical and measured mechanical indices predict response to cardiac resynchronization therapy”. In: *Computers in Biology and Medicine* 138 (Nov. 2021), p. 104872.

[245] Matteo Salvador et al. “Whole-heart electromechanical simulations using Latent Neural Ordinary Differential Equations”. In: *npj Digital Medicine* 7.1 (Apr. 2024), p. 90.

[246] Albert Dasí et al. “In-silico drug trials for precision medicine in atrial fibrillation: From ionic mechanisms to electrocardiogram-based predictions in structurally-healthy human atria”. In: *Frontiers in Physiology* 13 (Sept. 2022).

[247] Jordi Heijman et al. “Computational models of atrial fibrillation: achievements, challenges, and perspectives for improving clinical care”. In: *Cardiovascular Research* 117.7 (June 2021), pp. 1682–1699.

[248] Hermenegild J. Arevalo et al. “Arrhythmia risk stratification of patients after myocardial infarction using personalized heart models”. In: *Nature Communications* 7.1 (May 2016), p. 11437.

[249] Hector Martinez-Navarro et al. “High arrhythmic risk in antero-septal acute myocardial ischemia is explained by increased transmural reentry occurrence”. In: *Scientific Reports* 9.1 (Nov. 2019), p. 16803.

[250] Steven A. Niederer, Joost Lumens, and Natalia A. Trayanova. “Computational models in cardiology”. In: *Nature Reviews Cardiology* 16.2 (Feb. 2019), pp. 100–111.

[251] Patrick M. Boyle et al. “Computationally guided personalized targeted ablation of persistent atrial fibrillation”. In: *Nature Biomedical Engineering* 3.11 (Aug. 2019), pp. 870–879.

[252] Albert Dasí et al. “In Silico TRials guide optimal stratification of ATrIal FIbrillation patients to Catheter Ablation and pharmacological medicaTION: the i-STRATIFICATION study”. In: *Europace* 26.6 (June 2024).

[253] Leto Luana Riebel et al. “In silico evaluation of cell therapy in acute versus chronic infarction: role of automaticity, heterogeneity and Purkinje in human”. In: *Scientific Reports* 14.1 (Sept. 2024), p. 21584.

[254] Marina Strocchi et al. “An in silico guide for ventriculo-ventricular delay programming for left bundle branch-optimized cardiac resynchronization therapy”. In: *Europace* 27.5 (May 2025).

[255] A L Hodgkin and A F Huxley. “A quantitative description of membrane current and its application to conduction and excitation in nerve.” In: *The Journal of physiology* 117.4 (Aug. 1952), pp. 500–44.

[256] D. Noble. “A modification of the Hodgkin—Huxley equations applicable to Purkinje fibre action and pacemaker potentials”. In: *The Journal of Physiology* 160.2 (Feb. 1962), pp. 317–352.

[257] Thomas O’Hara et al. “Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation”. In: *PLoS Computational Biology* 7.5 (May 2011), e1002061.

[258] Eleonora Grandi, Francesco S. Pasqualini, and Donald M. Bers. “A novel computational model of the human ventricular action potential and Ca transient”. In: *Journal of Molecular and Cellular Cardiology* 48.1 (Jan. 2010), pp. 112–121.

[259] K. H. W. J. ten Tusscher et al. “A model for human ventricular tissue”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 286.4 (Apr. 2004), H1573–H1589.

[260] Sander Land et al. “A model of cardiac contraction based on novel measurements of tension development in human cardiomyocytes”. In: *Journal of Molecular and Cellular Cardiology* 106 (May 2017), pp. 68–83.

[261] Francesca Margara et al. “In-silico human electro-mechanical ventricular modelling and simulation for drug-induced pro-arrhythmia and inotropic risk assessment”. In: *Progress in Biophysics and Molecular Biology* 159 (Jan. 2021), pp. 58–74.

[262] Xin Zhou et al. “Lower diastolic tension may be indicative of higher proarrhythmic propensity in failing human cardiomyocytes”. In: *Scientific Reports* 14.1 (July 2024), p. 17351.

[263] Michael Klausen. *Cable theory Neuron RC circuit*. 2017. URL: `https://commons.wikimedia.org/wiki/File:Cable_theory_Neuron_RC_circuit_v3.svg`.

[264] Leslie Tung. “A bi-domain model for describing ischemic myocardial d-c potentials”. PhD thesis. Boston: Massachusetts Institute of Technology, 1978.

[265] James Keener and James Sneyd. *Mathematical Physiology*. Vol. 8. New York, NY: Springer New York, 1998.

[266] M. Potse et al. “A Comparison of Monodomain and Bidomain Reaction-Diffusion Models for Action Potential Propagation in the Human Heart”. In: *IEEE Transactions on Biomedical Engineering* 53.12 (Dec. 2006), pp. 2425–2435.

[267] Bjørn Fredrik Nielsen et al. “Optimal monodomain approximations of the bidomain equations”. In: *Applied Mathematics and Computation* 184.2 (Jan. 2007), pp. 276–290.

[268] Christian Vergara et al. “A coupled 3D–1D numerical monodomain solver for cardiac electrical activation in the myocardium with detailed Purkinje network”. In: *Journal of Computational Physics* 308 (Mar. 2016), pp. 218–238.

[269] Steven Niederer et al. “Simulating Human Cardiac Electrophysiology on Clinical Time-Scales”. In: *Frontiers in Physiology* 2 (2011).

[270] Rafael Sachetto Oliveira et al. “Performance evaluation of GPU parallelization, space-time adaptive algorithms, and their combination for simulating cardiac electrophysiology”. In: *International Journal for Numerical Methods in Biomedical Engineering* 34.2 (Feb. 2018).

[271] Rafael Sachetto Oliveira et al. “Comparing CUDA, OpenCL and OpenGL Implementations of the Cardiac Monodomain Equations”. In: 2012, pp. 111–120.

[272] Guillermo Vigueras et al. “Toward GPGPU accelerated human electromechanical cardiac simulations”. In: *International Journal for Numerical Methods in Biomedical Engineering* 30.1 (Jan. 2014), pp. 117–134.

[273] Andres Mena, Jose M. Ferrero, and Jose F. Rodriguez Matas. “GPU accelerated solver for nonlinear reaction–diffusion systems. Application to the electrophysiology problem”. In: *Computer Physics Communications* 196 (Nov. 2015), pp. 280–289.

[274] Piero Colli Franzone and Luciano Guerri. “Spreading of excitation in 3-d models of the anisotropic cardiac tissue. I. validation of the eikonal model”. In: *Mathematical Biosciences* 113.2 (Feb. 1993), pp. 145–209.

[275] Piero Colli Franzone et al. “Spread of excitation in 3-D models of the anisotropic cardiac tissue. II. Effects of fiber architecture and ventricular geometry”. In: *Mathematical Biosciences* 147.2 (Jan. 1998), pp. 131–171.

[276] Piero Colli Franzone et al. “Spread of excitation in 3-D models of the anisotropic cardiac tissue. III. Effects of ventricular geometry and fiber structure on the potential distribution”. In: *Mathematical Biosciences* 151.1 (July 1998), pp. 51–98.

[277] A. Capozzoli et al. “A comparison of Fast Marching, Fast Sweeping and Fast Iterative Methods for the solution of the eikonal equation”. In: *2013 21st Telecommunications Forum Telfor (TELFOR)*. IEEE, Nov. 2013, pp. 685–688.

[278] M. Sermesant et al. “A Fast-Marching Approach to Cardiac Electrophysiology Simulation for XMR Interventional Imaging”. In: 2005, pp. 607–615.

[279] Gaëtan Desrues et al. “Personal-by-Design: A 3D Electromechanical Model of the Heart Tailored for Personalisation”. In: 2021, pp. 447–457.

[280] Julia Camps et al. “Inference of ventricular activation properties from non-invasive electrocardiography”. In: *Medical Image Analysis* 73 (Oct. 2021), p. 102143.

[281] Mikael Wallman, Nicolas P. Smith, and Blanca Rodriguez. “A Comparative Study of Graph-Based, Eikonal, and Monodomain Simulations for the Estimation of Cardiac Activation Times”. In: *IEEE Transactions on Biomedical Engineering* 59.6 (June 2012), pp. 1739–1748.

[282] Aurel Neic et al. “Efficient computation of electrograms and ECGs in human whole heart simulations using a reaction-eikonal model”. In: *Journal of Computational Physics* 346 (Oct. 2017), pp. 191–211.

[283] Ender Konukoglu et al. “Efficient probabilistic model personalization integrating uncertainty on data and parameters: Application to Eikonal-Diffusion models in cardiac electrophysiology”. In: *Progress in Biophysics and Molecular Biology* 107.1 (Oct. 2011), pp. 134–146.

[284] Megan E. Farquhar et al. “Graph-based homogenisation for modelling cardiac fibrosis”. In: *Journal of Computational Physics* 459 (June 2022), p. 111126.

[285] Kazutaka Gima and Yoram Rudy. “Ionic Current Basis of Electrocardiographic Waveforms”. In: *Circulation Research* 90.8 (May 2002), pp. 889–896.

[286] Claudia Nagel et al. “Comparison of Propagation Models and Forward Calculation Methods on Cellular, Tissue and Organ Scale Atrial Electrophysiology”. In: *IEEE Transactions on Biomedical Engineering* 70.2 (Feb. 2023), pp. 511–522.

[287] Gernot Plank et al. “Generation of histo-anatomically representative models of the individual heart: tools and application”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367.1896 (June 2009), pp. 2257–2292.

[288] M. J. Bishop and G. Plank. “Bidomain ECG Simulations Using an Augmented Monodomain Model for the Cardiac Source”. In: *IEEE Transactions on Biomedical Engineering* 58.8 (Aug. 2011), pp. 2297–2307.

[289] Wenjia Bai et al. “Automated cardiovascular magnetic resonance image analysis with fully convolutional networks”. In: *Journal of Cardiovascular Magnetic Resonance* 20.1 (Feb. 2018), p. 65.

[290] Abhirup Banerjee et al. “A completely automated pipeline for 3D reconstruction of human heart from 2D cine magnetic resonance slices”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2212 (Dec. 2021).

[291] Chen Chen et al. “Improving the Generalizability of Convolutional Neural Network-Based Segmentation on CMR Images”. In: *Frontiers in Cardiovascular Medicine* 7 (June 2020).

[292] Jörg Sander et al. “Reconstruction and completion of high-resolution 3D cardiac shapes using anisotropic CMRI segmentations and continuous implicit neural representations”. In: *Computers in Biology and Medicine* 164 (Sept. 2023), p. 107266.

[293] Marina Strocchi et al. “A Semi-automatic Pipeline for Generation of Large Cohorts of Four-Chamber Heart Meshes”. In: 2024, pp. 117–127.

[294] Benjamin Villard, Vicente Grau, and Ernesto Zacur. “Surface Mesh Reconstruction from Cardiac MRI Contours”. In: *Journal of Imaging* 4.1 (Jan. 2018), p. 16.

[295] Ruben Doste et al. “A rule-based method to model myocardial fiber orientation in cardiac biventricular geometries with outflow tracts”. In: *International Journal for Numerical Methods in Biomedical Engineering* 35.4 (Apr. 2019).

[296] S. Labarthe et al. “A bilayer model of human atria: mathematical background, construction, and assessment”. In: *Europace* 16.suppl 4 (Nov. 2014), pp. iv21–iv29.

[297] Arnold C.T. Ng et al. “Association Between Diffuse Myocardial Fibrosis by Cardiac Magnetic Resonance Contrast-Enhanced T ₁ Mapping and Subclinical Myocardial Dysfunction in Diabetic Patients”. In: *Circulation: Cardiovascular Imaging* 5.1 (Jan. 2012), pp. 51–59.

[298] Yu Zhang et al. “MR extracellular volume mapping and non-contrast T1ρ mapping allow early detection of myocardial fibrosis in diabetic monkeys”. In: *European Radiology* 29.6 (June 2019), pp. 3006–3016.

[299] Philip Haaf et al. “Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: a comprehensive review”. In: *Journal of Cardiovascular Magnetic Resonance* 18.1 (Jan. 2016), p. 89.

[300] Ryan P O’Hara et al. “Personalized computational heart models with T1-mapped fibrotic remodeling predict sudden death risk in patients with hypertrophic cardiomyopathy”. In: *eLife* 11 (Jan. 2022).

[301] Tim De Coster et al. “Arrhythmogenicity of fibro-fatty infiltrations”. In: *Scientific Reports* 8.1 (Feb. 2018), p. 2050.

[302] Brodie A.J. Lawson et al. “Perlin noise generation of physiologically realistic cardiac fibrosis”. In: *Medical Image Analysis* 98 (Dec. 2024), p. 103240.

[303] Walter J. Paulus and Michael R. Zile. “From Systemic Inflammation to Myocardial Fibrosis”. In: *Circulation Research* 128.10 (May 2021), pp. 1451–1467.

[304] Gabriele G. Schiattarella et al. “Immunometabolic mechanisms of heart failure with preserved ejection fraction”. In: *Nature Cardiovascular Research* 1.3 (Mar. 2022), pp. 211–222.

[305] Helene von Bibra, Walter Paulus, and Martin St. John Sutton. “Cardiometabolic Syndrome and Increased Risk of Heart Failure”. In: *Current Heart Failure Reports* 13.5 (Oct. 2016), pp. 219–229.

[306] Matthew J. Gurka et al. “Progression of Metabolic Syndrome Severity During the Menopausal Transition”. In: *Journal of the American Heart Association* 5.8 (Aug. 2016).

[307] Antonella Meloni et al. “Gender Differences and Cardiometabolic Risk: The Importance of the Risk Factors”. In: *International Journal of Molecular Sciences* 24.2 (Jan. 2023), p. 1588.

[308] Gurleen Kaur and Emily Lau. “Sex differences in heart failure with preserved ejection fraction: From traditional risk factors to sex-specific risk factors”. In: *Women’s Health* 18 (Jan. 2022).

[309] Giuseppe M.C. Rosano et al. “Differences in presentation, diagnosis and management of heart failure in women. A scientific statement of the Heart Failure Association of the ESC”. In: *European Journal of Heart Failure* 26.8 (Aug. 2024), pp. 1669–1686.

[310] Kanako Teramoto et al. “Epidemiology and Clinical Features of Heart Failure with Preserved Ejection Fraction”. In: *Cardiac Failure Review* 8 (Aug. 2022).

[311] Åsa K Hedman et al. “Identification of novel pheno-groups in heart failure with preserved ejection fraction using machine learning”. In: *Heart* 106.5 (Mar. 2020), pp. 342–349.

[312] Ryo Nakamaru et al. “Phenotyping of Elderly Patients With Heart Failure Focused on Noncardiac Conditions: A Latent Class Analysis From a Multicenter Registry of Patients Hospitalized With Heart Failure”. In: *Journal of the American Heart Association* 12.3 (Feb. 2023).

[313] Morgane Pierre-Jean et al. “Phenotyping of heart failure with preserved ejection faction using electronic health records and echocardiography”. In: *European Heart Journal Open* 4.1 (Dec. 2023).

[314] Florian Schrub et al. “Heart failure with preserved ejection fraction: A clustering approach to a heterogenous syndrome”. In: *Archives of Cardiovascular Diseases* 113.6-7 (June 2020), pp. 381–390.

[315] Matthew W. Segar et al. “Phenomapping of patients with heart failure with preserved ejection fraction using machine learning-based unsupervised cluster analysis”. In: *European Journal of Heart Failure* 22.1 (Jan. 2020), pp. 148–158.

[316] Yohei Sotomi et al. “Phenotyping of acute decompensated heart failure with preserved ejection fraction”. In: *Heart* 108.19 (Oct. 2022), pp. 1553–1561.

[317] Alicia Uijl et al. “Identification of distinct phenotypic clusters in heart failure with preserved ejection fraction”. In: *European Journal of Heart Failure* 23.6 (June 2021), pp. 973–982.

[318] Lourdes Vicent et al. “Profiling heart failure with preserved or mildly reduced ejection fraction by cluster analysis”. In: *European Heart Journal - Quality of Care and Clinical Outcomes* (Aug. 2024).

[319] Alison Brann et al. “Global longitudinal strain predicts clinical outcomes in patients with heart failure with preserved ejection fraction”. In: *European Journal of Heart Failure* 25.10 (Oct. 2023), pp. 1755–1765.

[320] Adam D. DeVore et al. “Impaired left ventricular global longitudinal strain in patients with heart failure with preserved ejection fraction: insights from the RELAX trial”. In: *European Journal of Heart Failure* 19.7 (July 2017), pp. 893–900.

[321] Franz Duca et al. “Gender-related differences in heart failure with preserved ejection fraction”. In: *Scientific Reports* 8.1 (Jan. 2018), p. 1080.

[322] Mauro Gori et al. “Sex-specific cardiovascular structure and function in heart failure with preserved ejection fraction”. In: *European Journal of Heart Failure* 16.5 (May 2014), pp. 535–542.

[323] Federico Capone, Roberto Vettor, and Gabriele G. Schiattarella. “Cardiometabolic HFpEF: NASH of the Heart”. In: *Circulation* 147.6 (Feb. 2023), pp. 451–453.

[324] John W. Ostrominski et al. “Cardiovascular-Kidney-Metabolic Overlap in Heart Failure With Mildly Reduced or Preserved Ejection Fraction”. In: *Journal of the American College of Cardiology* 84.2 (July 2024), pp. 223–228.

[325] Leah B. Kosyakovsky et al. “Uncovering Unrecognized Heart Failure With Preserved Ejection Fraction Among Individuals With Obesity and Dyspnea”. In: *Circulation: Heart Failure* 17.5 (May 2024).

[326] Sadi Loai and Hai-Ling Margaret Cheng. “Heart failure with preserved ejection fraction: the missing pieces in diagnostic imaging”. In: *Heart Failure Reviews* 25.2 (Mar. 2020), pp. 305–319.

[327] Luca Busetto et al. “A new framework for the diagnosis, staging and management of obesity in adults”. In: *Nature Medicine* 30.9 (Sept. 2024), pp. 2395–2399.

[328] Paul L. Huang. “A comprehensive definition for metabolic syndrome”. In: *Disease Models & Mechanisms* 2.5-6 (Apr. 2009), pp. 231–237.

[329] Pratik Sinha, Carolyn S. Calfee, and Kevin L. Delucchi. “Practitioner’s Guide to Latent Class Analysis: Methodological Considerations and Common Pitfalls”. In: *Critical Care Medicine* 49.1 (Jan. 2021), e63–e79.

[330] Dalane W. Kitzman and Sanjiv J. Shah. “The HFpEF Obesity Phenotype”. In: *Journal of the American College of Cardiology* 68.2 (July 2016), pp. 200–203.

[331] Yogesh N.V. Reddy et al. “Quality of life in heart failure with preserved ejection fraction: importance of obesity, functional capacity, and physical inactivity”. In: *European Journal of Heart Failure* 22.6 (June 2020), pp. 1009–1018.

[332] Anna Fry et al. “Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population”. In: *American Journal of Epidemiology* 186.9 (Nov. 2017), pp. 1026–1034.

[333] Emily S. Lau et al. “Sex Differences in Cardiometabolic Traits and Determinants of Exercise Capacity in Heart Failure With Preserved Ejection Fraction”. In: *JAMA Cardiology* 5.1 (Jan. 2020), p. 30.

[334] Yohei Sotomi et al. “Sex Differences in Heart Failure With Preserved Ejection Fraction”. In: *Journal of the American Heart Association* 10.5 (Mar. 2021).

[335] Anna G. Hoek et al. “Epidemiology of heart failure in diabetes: a disease in disguise”. In: *Diabetologia* 67.4 (Apr. 2024), pp. 574–601.

[336] Jan Malik et al. “Heart failure with preserved ejection fraction is the most frequent but commonly overlooked phenotype in patients on chronic hemodialysis”. In: *Frontiers in Cardiovascular Medicine* 10 (June 2023).

[337] Rajat Kalra et al. “Cardiac Function and Sudden Cardiac Death in Heart Failure With Preserved Ejection Fraction (from the TOPCAT Trial)”. In: *The American Journal of Cardiology* 129 (Aug. 2020), pp. 46–52.

[338] Peter Blomstrand et al. “Overweight and obesity impair left ventricular systolic function as measured by left ventricular ejection fraction and global longitudinal strain”. In: *Cardiovascular Diabetology* 17.1 (Dec. 2018), p. 113.

[339] Wilson Cañon-Montañez et al. “Central Obesity is the Key Component in the Association of Metabolic Syndrome With Left Ventricular Global Longitudinal Strain Impairment”. In: *Revista Española de Cardiología (English Edition)* 71.7 (July 2018), pp. 524–530.

[340] Emily S. Lau et al. “Sex Differences in Left Ventricular Function and Cardiac Mechanics”. In: *Journal of the American Heart Association* 13.12 (June 2024).

[341] Yong-Huai Wang et al. “Left ventricular global longitudinal strain using a novel fully automated method: A head-to-head comparison with a manual layer-specific strain and establishment of normal reference ranges”. In: *International Journal of Cardiology* 403 (May 2024), p. 131886.

[342] Kristoffer Grundtvig Skaarup et al. “The impact of cardiovascular risk factors on global longitudinal strain over a decade in the general population: the copenhagen city heart study”. In: *The International Journal of Cardiovascular Imaging* 36.10 (Oct. 2020), pp. 1907–1916.

[343] Muaz M. Abudiab et al. “Cardiac output response to exercise in relation to metabolic demand in heart failure with preserved ejection fraction”. In: *European Journal of Heart Failure* 15.7 (July 2013), pp. 776–785.

[344] Matthew K. Burrage et al. “Energetic Basis for Exercise-Induced Pulmonary Congestion in Heart Failure With Preserved Ejection Fraction”. In: *Circulation* 144.21 (Nov. 2021), pp. 1664–1678.

[345] Kazunori Omote, Frederik H. Verbrugge, and Barry A. Borlaug. “Heart Failure with Preserved Ejection Fraction: Mechanisms and Treatment Strategies”. In: *Annual Review of Medicine* 73.1 (Jan. 2022), pp. 321–337.

[346] Yogesh N.V. Reddy et al. “High-Output Heart Failure”. In: *Journal of the American College of Cardiology* 68.5 (Aug. 2016), pp. 473–482.

[347] Sheldon E. Litwin et al. “Obesity in heart failure with preserved ejection fraction: Insights from the <scp>REDUCE LAP-HF II</scp> trial”. In: *European Journal of Heart Failure* 26.1 (Jan. 2024), pp. 177–189.

[348] Peter R. Hyson and David P. Kao. *Biomarkers Correspond with Echocardiographic Phenotypes in Heart Failure with Preserved Ejection Fraction: A Secondary Analysis of the RELAX Trial*. May 2024.

[349] Francesco Rubino et al. “Definition and diagnostic criteria of clinical obesity”. In: *The Lancet Diabetes & Endocrinology* (Jan. 2025).

[350] Sanjiv Shah et al. *HeartShare Deep Phenotyping Study Protocol v. 2.3*. 2024.

[351] Sanjiv J. Shah et al. “Accelerating therapeutic discoveries for heart failure: a new public–private partnership”. In: *Nature Reviews Drug Discovery* 21.11 (Nov. 2022), pp. 781–782.

[352] UK HFpEF Collaborative Group. “Rationale and design of the United Kingdom Heart Failure with Preserved Ejection Fraction Registry”. In: *Heart* 110.5 (Mar. 2024), pp. 359–365.

[353] Magnus T. Jensen et al. “Changes in Cardiac Morphology and Function in Individuals With Diabetes Mellitus”. In: *Circulation: Cardiovascular Imaging* 12.9 (Sept. 2019).

[354] Shlomo Stern and Samuel Sclarowsky. “The ECG in Diabetes Mellitus”. In: *Circulation* 120.16 (Oct. 2009), pp. 1633–1636.

[355] Alptug Tokatli et al. “Prolonged Tp-e Interval, Tp-e/QT Ratio and Tp-e/QTc Ratio in Patients with Type 2 Diabetes Mellitus”. In: *Endocrinology and Metabolism* 31.1 (2016), p. 105.

[356] Steven J. Mould et al. “Association of T-wave abnormalities with major cardiovascular events in diabetes: the ACCORD trial”. In: *Diabetologia* 64.3 (Mar. 2021), pp. 504–511.

[357] Sophie V Eastwood et al. “Algorithms for the Capture and Adjudication of Prevalent and Incident Diabetes in UK Biobank”. In: *PLOS ONE* 11.9 (Sept. 2016), e0162388.

[358] Oliver J Rider et al. “Improvements in ECG accuracy for diagnosis of left ventricular hypertrophy in obesity”. In: *Heart* 102.19 (Oct. 2016), pp. 1566–1572.

[359] Christiane Spiel et al. “A Euclidean distance-based matching procedure for nonrandomized comparison studies.” In: *European Psychologist* 13.3 (2008), pp. 180–187.

[360] Andrew S. Levey et al. “A New Equation to Estimate Glomerular Filtration Rate”. In: *Annals of Internal Medicine* 150.9 (May 2009), p. 604.

[361] F. Locatelli et al. “Foreword”. In: *Nephrology Dialysis Transplantation* 19.suppl_2 (May 2004), pp. ii1–ii1.

[362] Maurice Sokolow and Thomas P. Lyon. “The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads”. In: *American Heart Journal* 37.2 (Feb. 1949), pp. 161–186.

[363] Oliver I. Brown et al. “Relationship Among Diabetes, Obesity, and Cardiovascular Disease Phenotypes: A UK Biobank Cohort Study”. In: *Diabetes Care* 46.8 (Aug. 2023), pp. 1531–1540.

[364] Y H Chan. “Biostatistics 104: correlational analysis.” In: *Singapore medical journal* 44.12 (Dec. 2003), pp. 614–9.

[365] N.P. Johnson. “Metformin use in women with polycystic ovary syndrome”. In: *Ann Transl Med* 2.6 (2014), p. 56.

[366] D.J. Ewing, I.W. Campbell, and B.F. Clarke. “Heart rate changes in diabetes mellitus”. In: *The Lancet* 317.8213 (Jan. 1981).

[367] Paolo Palatini et al. “Relationship of Tachycardia With High Blood Pressure and Metabolic Abnormalities”. In: *Hypertension* 30.5 (Nov. 1997), pp. 1267–1273.

[368] Aaron I. Vinik and Dan Ziegler. “Diabetic Cardiovascular Autonomic Neuropathy”. In: *Circulation* 115.3 (Jan. 2007), pp. 387–397.

[369] Bu B Yeap et al. “Hyperglycemia affects cardiovascular autonomic nerve function in normal subjects”. In: *Diabetes Care* 19.8 (Aug. 1996), pp. 880–882.

[370] Stephen J. Synowski et al. “Effects of glucose ingestion on autonomic and cardiovascular measures during rest and mental challenge”. In: *Journal of Psychosomatic Research* 74.2 (Feb. 2013), pp. 149–154.

[371] Yun-Ru Lai et al. “HbA1C Variability Is Strongly Associated With the Severity of Cardiovascular Autonomic Neuropathy in Patients With Type 2 Diabetes After Longer Diabetes Duration”. In: *Frontiers in Neuroscience* 13 (May 2019).

[372] Pentti M. Rautaharju, Jay W. Mason, and Toshio Akiyama. “New age- and sex-specific criteria for QT prolongation based on rate correction formulas that minimize bias at the upper normal limits”. In: *International Journal of Cardiology* 174.3 (July 2014), pp. 535–540.

[373] Jonas L. Isaksen et al. “Electrocardiographic markers in patients with type 2 diabetes and the role of diabetes duration”. In: *Journal of Electrocardiology* 84 (May 2024), pp. 129–136.

[374] Deborah B Diercks et al. “Electrocardiographic manifestations: electrolyte abnormalities”. In: *The Journal of Emergency Medicine* 27.2 (Aug. 2004), pp. 153–160.

[375] Tomi Laitinen et al. “Electrocardiographic Alterations during Hyperinsulinemic Hypoglycemia in Healthy Subjects”. In: *Annals of Noninvasive Electrocardiology* 13.2 (Apr. 2008), pp. 97–105.

[376] Vittorio Palmieri et al. “Effect of Type 2 Diabetes Mellitus on Left Ventricular Geometry and Systolic Function in Hypertensive Subjects”. In: *Circulation* 103.1 (Jan. 2001), pp. 102–107.

[377] K Eguchi et al. “Type 2 diabetes is associated with left ventricular concentric remodeling in hypertensive patients”. In: *American Journal of Hypertension* 18.1 (Jan. 2005), pp. 23–29.

[378] Gerard P. Aurigemma, Michael R. Zile, and William H. Gaasch. “Contractile Behavior of the Left Ventricle in Diastolic Heart Failure”. In: *Circulation* 113.2 (Jan. 2006), pp. 296–304.

[379] Authors/Task Force members et al. “2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: The Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC)”. In: *European Heart Journal* 35.39 (Oct. 2014), pp. 2733–2779. URL: https://doi.org/10.1093/eurheartj/ehu284.

[380] SCORE2-Diabetes Working Group and the ESC Cardiovascular Risk Collaboration. “SCORE2-Diabetes: 10-year cardiovascular risk estimation in type 2 diabetes in Europe”. In: *European Heart Journal* 44.28 (July 2023), pp. 2544–2556. URL: https://doi.org/10.1093/eurheartj/ehad260.

[381] Shota Ikeda et al. “A higher resting heart rate is associated with cardiovascular event risk in patients with type 2 diabetes mellitus without known cardiovascular disease”. In: *Hypertension Research* 46.5 (May 2023), pp. 1090–1099.

[382] Xiaojing Chen et al. “Incremental changes in QRS duration as predictor for cardiovascular disease: a 21-year follow-up of a randomly selected general population”. In: *Scientific Reports* 11.1 (July 2021), p. 13652.

[383] Daniel Levy et al. “Prognostic Implications of Echocardiographically Determined Left Ventricular Mass in the Framingham Heart Study”. In: *New England Journal of Medicine* 322.22 (May 1990), pp. 1561–1566.

[384] Richard S. Cooper et al. “Left ventricular hypertrophy is associated with worse survival independent of ventricular function and number of coronary arteries severely narrowed”. In: *The American Journal of Cardiology* 65.7 (Feb. 1990), pp. 441–445.

[385] Michael J. Koren et al. “Relation of Left Ventricular Mass and Geometry to Morbidity and Mortality in Uncomplicated Essential Hypertension”. In: *Annals of Internal Medicine* 114.5 (Mar. 1991), pp. 345–352.

[386] Wilbert S. Aronow et al. “Congestive heart failure, coronary events and atherothrombotic brain infarction in elderly blacks and whites with systemic hypertension and with and without echocardiographic and electrocardiographic evidence of left ventricular hypertrophy”. In: *The American Journal of Cardiology* 67.4 (Feb. 1991), pp. 295–299.

[387] Paolo Verdecchia et al. “Prognostic value of left ventricular mass and geometry in systemic hypertension with left ventricular hypertrophy”. In: *The American Journal of Cardiology* 78.2 (July 1996), pp. 197–202.

[388] Melanie J Davies et al. “Management of hyperglycaemia in type 2 diabetes, 2022. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD)”. In: *Diabetologia* 65.12 (2022), pp. 1925–1966. URL: https://doi.org/10.1007/s00125-022-05787-2.

[389] Jeffrey R. Erickson et al. “CaMKII in the Cardiovascular System: Sensing Redox States”. In: *Physiological Reviews* 91.3 (July 2011), pp. 889–915.

[390] Mark E. Anderson. “Oxidant stress promotes disease by activating CaMKII”. In: *Journal of Molecular and Cellular Cardiology* 89 (Dec. 2015), pp. 160–167.

[391] Shan Lu et al. “Hyperglycemia Acutely Increases Cytosolic Reactive Oxygen Species via O-linked GlcNAcylation and CaMKII Activation in Mouse Ventricular Myocytes”. In: *Circulation Research* 126.10 (May 2020).

[392] Andrew F. James, Stéphanie C.M. Choisy, and Jules C. Hancox. “Recent advances in understanding sex differences in cardiac repolarization”. In: *Progress in Biophysics and Molecular Biology* 94.3 (July 2007), pp. 265–319.

[393] Peter W. Macfarlane. “The Influence of Age and Sex on the Electrocardiogram”. In: 2018, pp. 93–106.

[394] Sarah R. St. Pierre, Mathias Peirlinck, and Ellen Kuhl. “Sex Matters: A Comprehensive Comparison of Female and Male Hearts”. In: *Frontiers in Physiology* 13 (Mar. 2022).

[395] Jaya M. Mehta and JoAnn E. Manson. “The menopausal transition period and cardiovascular risk”. In: *Nature Reviews Cardiology* 21.3 (Mar. 2024), pp. 203–211.

[396] Yilin Yoshida et al. “Early Menopause and Cardiovascular Disease Risk in Women With or Without Type 2 Diabetes: A Pooled Analysis of 9,374 Postmenopausal Women”. In: *Diabetes Care* 44.11 (Nov. 2021), pp. 2564–2572.

[397] Irfan Khan et al. “Surrogate Adiposity Markers and Mortality”. In: *JAMA Network Open* 6.9 (Sept. 2023), e2334836.

[398] Michelle M.Y. Chan and Carolyn S.P. Lam. “How do patients with heart failure with preserved ejection fraction die?” In: *European Journal of Heart Failure* 15.6 (June 2013), pp. 604–613.

[399] Muthiah Vaduganathan et al. “Mode of Death in Heart Failure With Preserved Ejection Fraction”. In: *Journal of the American College of Cardiology* 69.5 (Feb. 2017), pp. 556–569.

[400] Heikki V. Huikuri, Agustin Castellanos, and Robert J. Myerburg. “Sudden Death Due to Cardiac Arrhythmias”. In: *New England Journal of Medicine* 345.20 (Nov. 2001), pp. 1473–1482.

[401] Jerome Anthony E. Alvarez, Mohsin Saleet Jafri, and Aman Ullah. “Using a Failing Human Ventricular Cardiomyocyte Model to Re-Evaluate Ca2+ Cycling, Voltage Dependence, and Spark Characteristics”. In: *Biomolecules* 14.11 (Oct. 2024), p. 1371.

[402] Leo Priebe and Dirk J. Beuckelmann. “Simulation Study of Cellular Electric Properties in Heart Failure”. In: *Circulation Research* 82.11 (June 1998), pp. 1206–1223.

[403] Juan F. Gomez, Karen Cardona, and Beatriz Trenor. “Lessons learned from multi-scale modeling of the failing heart”. In: *Journal of Molecular and Cellular Cardiology* 89 (Dec. 2015), pp. 146–159.

[404] Shuang Qian et al. “Developing cardiac digital twin populations powered by machine learning provides electrophysiological insights in conduction and repolarization”. In: *Nature Cardiovascular Research* 4.5 (May 2025), pp. 624–636.

[405] Elham Kayvanpour et al. “Towards Personalized Cardiology: Multi-Scale Modeling of the Failing Heart”. In: *PLOS ONE* 10.7 (July 2015), e0134869.

[406] Marcel Beetz et al. “Multi-class point cloud completion networks for 3D cardiac anatomy reconstruction from cine magnetic resonance images”. In: *Medical Image Analysis* 90 (Dec. 2023), p. 102975.

[407] Hannah J Smith et al. “Anatomical basis of sex differences in the electrocardiogram identified by three-dimensional torso-heart imaging reconstruction pipeline”. In: *eLife* (Sept. 2025).

[408] Ruben Doste et al. “An Automated Computational Pipeline for Generating Large-Scale Cohorts of Patient-Specific Ventricular Models in Electromechanical In Silico Trials”. In: *arXiv* (Mar. 2025).

[409] Yeda Wu et al. “Genome-wide association study of medication-use and associated disease in the UK Biobank”. In: *Nature Communications* 10.1 (Apr. 2019), p. 1891.

[410] Carl J. Lavie and Franz H. Messerli. “Cardiovascular Adaptation to Obesity and Hypertension”. In: *Chest* 90.2 (Aug. 1986), pp. 275–279.

[411] Martin A. Alpert. “Obesity Cardiomyopathy: Pathophysiology and Evolution of the Clinical Syndrome”. In: *The American Journal of the Medical Sciences* 321.4 (Apr. 2001), pp. 225–236.

[412] Peter Taggart et al. “Inhomogeneous Transmural Conduction During Early Ischaemia in Patients with Coronary Artery Disease”. In: *Journal of Molecular and Cellular Cardiology* 32.4 (Apr. 2000), pp. 621–630.

[413] Maxx Holmes et al. “Sex-specific human electromechanical multiscale in-silico models for virtual therapy evaluation”. In: *Journal of Molecular and Cellular Cardiology Plus* 13 (Sept. 2025), p. 100479.

[414] K. A. Sedova et al. “Mechanism of Electrocardiographic T-Wave Flattening in Diabetes Mellitus: Experimental and Simulation Study”. In: *Physiological Research* (Oct. 2017), pp. 781–789.

[415] Marina Strocchi et al. “Multi-scale Whole-Heart Electromechanics Modeling to Link Cellular, Tissue and Systemic Properties to Cardiac Biomarkers in the Diabetic Male and Female Heart”. In: 2025, pp. 218–230.

[416] Martin J. Bishop and Gernot Plank. “Stochastic virtual heart model predictions”. In: *Nature Cardiovascular Research* 4.5 (Apr. 2025), pp. 539–542.

[417] Lucas Arantes Berg et al. “Impact of Mesh Resolution on Re-entry Patterns in Healing Infarction using a Biventricular Digital Twin Model”. In: *2025 Computing in Cardiology (CinC)*. 2025. URL: `https://cinc.org/2025/Program/accepted/270_Preprint.pdf`.

[418] Kimberly J. McCabe and Padmini Rangamani. “Computational modeling approaches to cAMP/PKA signaling in cardiomyocytes”. In: *Journal of Molecular and Cellular Cardiology* 154 (May 2021), pp. 32–40.

[419] Sonia Cortassa et al. “An Integrated Model of Cardiac Mitochondrial Energy Metabolism and Calcium Dynamics”. In: *Biophysical Journal* 84.4 (Apr. 2003), pp. 2734–2755.

[420] Clare Bycroft et al. “The UK Biobank resource with deep phenotyping and genomic data”. In: *Nature* 562.7726 (Oct. 2018), pp. 203–209.

[421] Sophie V. Eastwood et al. “Ancestry, ethnicity, and race: explaining inequalities in cardiometabolic disease”. In: *Trends in Molecular Medicine* 30.6 (June 2024), pp. 541–551.

[422] Minki Hwang et al. “Ganglionated plexi stimulation induces pulmonary vein triggers and promotes atrial arrhythmogenecity: In silico modeling study”. In: *PLOS ONE* 12.2 (Feb. 2017), e0172931.

[423] Lucas Arantes Berg et al. “Enhanced optimization-based method for the generation of patient-specific models of Purkinje networks”. In: *Scientific Reports* 13.1 (July 2023), p. 11788.

[424] R. Sebastian et al. “Characterization and Modeling of the Peripheral Cardiac Conduction System”. In: *IEEE Transactions on Medical Imaging* 32.1 (Jan. 2013), pp. 45–55.

[425] Marco Viceconti et al. “In silico trials: Verification, validation and uncertainty quantification of predictive models used in the regulatory evaluation of biomedical products”. In: *Methods* 185 (Jan. 2021), pp. 120–127.

[426] Bernardo L. de Oliveira and Steven Niederer. “A Biophysical Systems Approach to Identifying the Pathways of Acute and Chronic Doxorubicin Mitochondrial Cardiotoxicity”. In: *PLOS Computational Biology* 12.11 (Nov. 2016), e1005214.

[427] Gary R. Mirams et al. “Prediction of Thorough QT study results using action potential simulations based on ion channel screens”. In: *Journal of Pharmacological and Toxicological Methods* 70.3 (Nov. 2014), pp. 246–254.

[428] Elisa Passini et al. “Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity”. In: *Frontiers in Physiology* 8 (Sept. 2017).

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

[430] Irem Cetin et al. “Radiomics Signatures of Cardiovascular Risk Factors in Cardiac MRI: Results From the UK Biobank”. In: *Frontiers in Cardiovascular Medicine* 7 (Nov. 2020).

[431] Esmeralda Ruiz Pujadas et al. “Prediction of incident cardiovascular events using machine learning and CMR radiomics”. In: *European Radiology* 33.5 (Dec. 2022), pp. 3488–3500.

[432] Lasya Priya Kotu et al. “Cardiac magnetic resonance image-based classification of the risk of arrhythmias in post-myocardial infarction patients”. In: *Artificial Intelligence in Medicine* 64.3 (July 2015), pp. 205–215.

[433] T. Gibbs et al. “Quantitative assessment of myocardial scar heterogeneity using cardiovascular magnetic resonance texture analysis to risk stratify patients post-myocardial infarction”. In: *Clinical Radiology* 73.12 (Dec. 2018), pp. 17–1059.

[434] Emine Sebnem Durmaz et al. “Radiomics-based machine learning models in STEMI: a promising tool for the prediction of major adverse cardiac events”. In: *European Radiology* 33.7 (Jan. 2023), pp. 4611–4620.

[435] Mao-Yuan M. Su et al. “CMR-Verified Diffuse Myocardial Fibrosis Is Associated With Diastolic Dysfunction in HFpEF”. In: *JACC: Cardiovascular Imaging* 7.10 (Oct. 2014), pp. 991–997.

[436] Julia Mascherbauer et al. “Cardiac Magnetic Resonance Postcontrast T1 Time Is Associated With Outcome in Patients With Heart Failure and Preserved Ejection Fraction”. In: *Circulation: Cardiovascular Imaging* 6.6 (Nov. 2013), pp. 1056–1065.

[437] Taku Omori et al. “Myocardial Native T1 Predicts Load-Independent Left Ventricular Chamber Stiffness In Patients With HFpEF”. In: *JACC: Cardiovascular Imaging* 13.10 (Oct. 2020), pp. 2117–2128.

[438] Karl-Philipp Rommel et al. “Extracellular Volume Fraction for Characterization of Patients With Heart Failure and Preserved Ejection Fraction”. In: *Journal of the American College of Cardiology* 67.15 (Apr. 2016), pp. 1815–1825.

[439] Zahra Raisi-Estabragh et al. “Incident Clinical and Mortality Associations of Myocardial Native T1 in the UK Biobank”. In: *JACC: Cardiovascular Imaging* 16.4 (Apr. 2023), pp. 450–460.

[440] Steffen Schuler et al. “Cobiveco: Consistent biventricular coordinates for precise and intuitive description of position in the heart – with MATLAB implementation”. In: *Medical Image Analysis* 74 (Dec. 2021), p. 102247.

A

Appendix for Chapter 4

Table A.1: Cardiometabolic diseases included in the study, as defined by the International Classification of Diseases (ICD). Each corresponding UK Biobank field refers to the date of first report of respective ICD-10 codes. MASLD: metabolic dysfunction-associated fatty liver disease.

Disease Category	**ICD-10 Code**	**Description**	**UK Biobank Field**
Ischemic heart disease	I11	Hypertensive heart disease	131288
	I13	Hypertensive heart and renal disease	131292
	I20	Angina pectoris	131296
	I21	Acute myocardial infarction	131298
	I22	Subsequent myocardial infarction	131300
	I24	Other acute ischemic heart disease	131304
	I25	Chronic ischemic heart disease	131306
Hypertension	I10	Essential primary hypertension	131286
	I15	Secondary hypertension	131294
	O16	Unspecified maternal hypertension	132192
Diabetes	E10	Insulin dependent diabetes mellitus	130706
	E11	Non-insulin dependent diabetes mellitus	130708
	E14	Unspecified diabetes	130714
Obesity	E66	Obesity	130792
Chronic kidney disease	N18	Chronic renal failure	132032
Liver disease	K76	Other diseases of liver (primarily including MASLD)	131670
Stroke	I61	Intracerebral haemorrhage	131362
	I62	Other nontraumatic intracranial haemorrhage	131364
	I63	Cerebral infarction	131366
	I64	Stroke, not specified as haemorrhage or infarction	131368

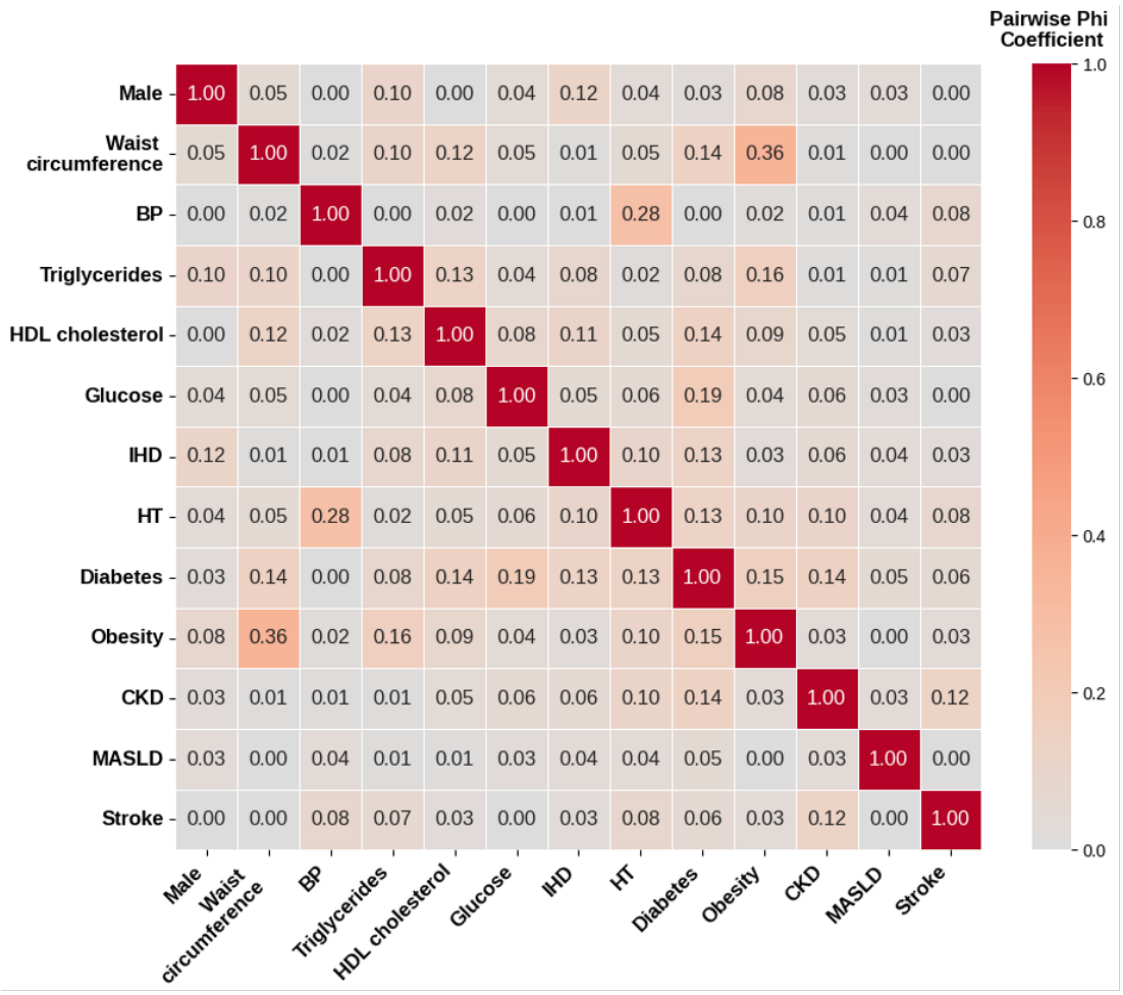

Figure A.1: Variable selection for latent class analysis. Pairwise Phi coefficients computed to assess the association between candidate indicator variables to include in the latent class analysis models. BP: blood pressure, HDL: high-density lipoprotein, IHD: ischemic heart disease, HT: hypertension, CKD: chronic kidney disease, MASLD: metabolic dysfunction-associated fatty liver disease.

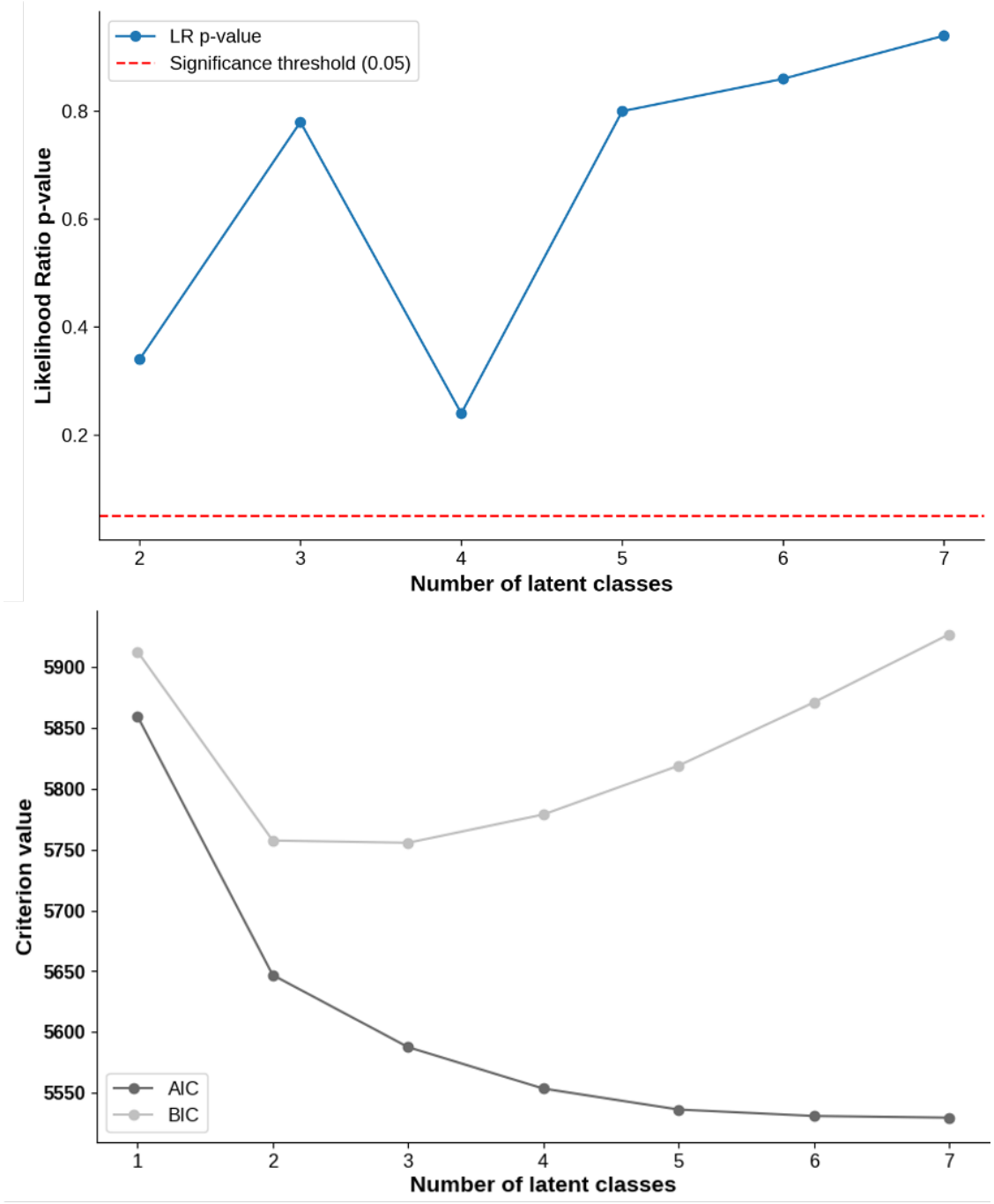

Figure A.2: Latent class analysis model optimisation. Top panel: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) computed for candidate models with an increasing number of latent classes, ranging from 1 to 7. The optimal model was found to have n=3 classes, which minimises BIC and is further supported by AIC(n=3) < AIC(n=2). Bottom panel: likelihood ratio test comparing model fit for candidate models with an increasing number of latent classes, ranging from 2 to 7. The likelihood ratio test did not provide further significant results to consider additionally to the AIC and BIC.

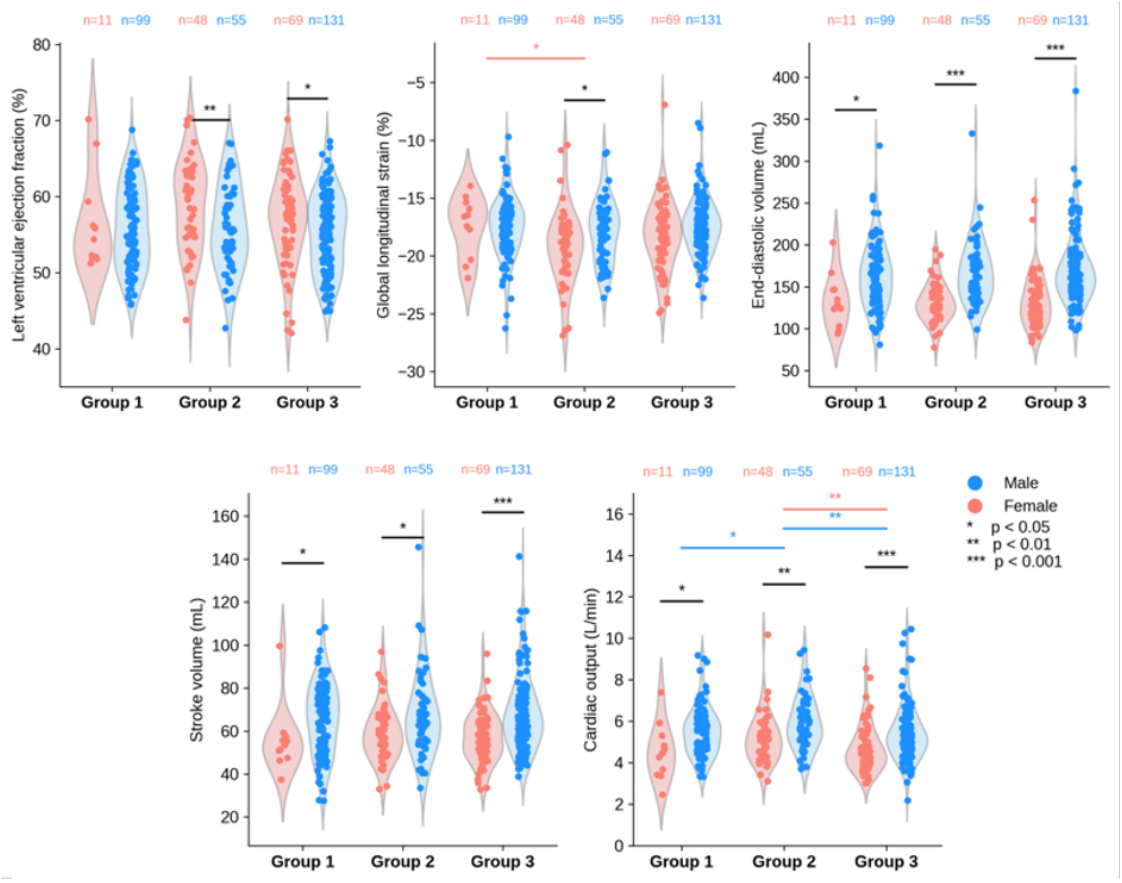

Figure A.3: Distribution of CMR image-derived left ventricular function parameters in the HFpEF cardiometabolic phenogroups identified, excluding cases without a full set of imaging parameters. CMR: cardiac magnetic resonance; HFpEF: heart failure with preserved ejection fraction.

Table A.2: Cardiometabolic comorbidity burden and metabolic syndrome in the study cohort, stratified by HFpEF status (incident vs prevalent). The Chi-square test was used to compare categorical variables between male and female subgroups. The p-values reported refer to the individual results of those respective tests. For sex-specific columns, percentage for each count is reported as the percent of that sex subgroup. The asterisk refers to the number of additional comorbidities refers to the number of cases with *n* cardiometabolic diseases as defined in main text, additionally to HFpEF, and excluding MetS which I consider separately. HFpEF: heart failure with preserved ejection fraction, F: female, M: male, MetS: metabolic syndrome. HFpEF: heart failure with preserved ejection fraction, F: female, M: male, MetS: metabolic syndrome.

	Incident HFpEF (n = 287)	**Prevalent HFpEF (n = 158)**	**p-value**
MetS components and clinically-defined syndrome			
Waist circumference >102 cm (M) or >89 cm (F)	107 (37.3%)	73 (46.2%)	0.083
Blood pressure >130/85 mmHg	227 (79.1%)	108 (68.4%)	0.016
Triglycerides >150 mg/dL	127 (44.3%)	70 (44.3%)	1
HDL-C <40 mg/dL (M) or <50 mg/dL (F)	71 (24.7%)	49 (31.0%)	0.188
Glucose >100 mg/dL	57 (19.9%)	29 (18.4%)	0.795
Clinically-defined MetS (≥ 3 components)	87 (30.3%)	59 (37.3%)	0.16
Cardiometabolic comorbidities			
Ischemic heart disease	109 (38.0%)	75 (47.5%)	0.065
Hypertension	226 (78.7%)	113 (71.5%)	0.11
Diabetes	51 (17.8%)	17 (10.8%)	0.067
Obesity	197 (68.6%)	107 (67.7%)	0.926
CKD	37 (12.9%)	9 (5.7%)	0.026
MASLD	5 (1.7%)	2 (1.3%)	1
Stroke	23 (8.0%)	5 (3.2%)	0.07
Number of additional comorbidities*			0.57
0 (HFpEF only)	19 (6.6%)	12 (7.6%)	-
1	64 (22.3%)	38 (24.1%)	-
2	91 (31.7%)	55 (34.8%)	-
3	67 (23.3%)	35 (22.2%)	-
4	32 (11.1%)	16 (10.1%)	-
5	11 (3.8%)	1 (0.6%)	-
6	3 (1.0%)	1 (0.6%)	-

B

Appendix for Chapter 5

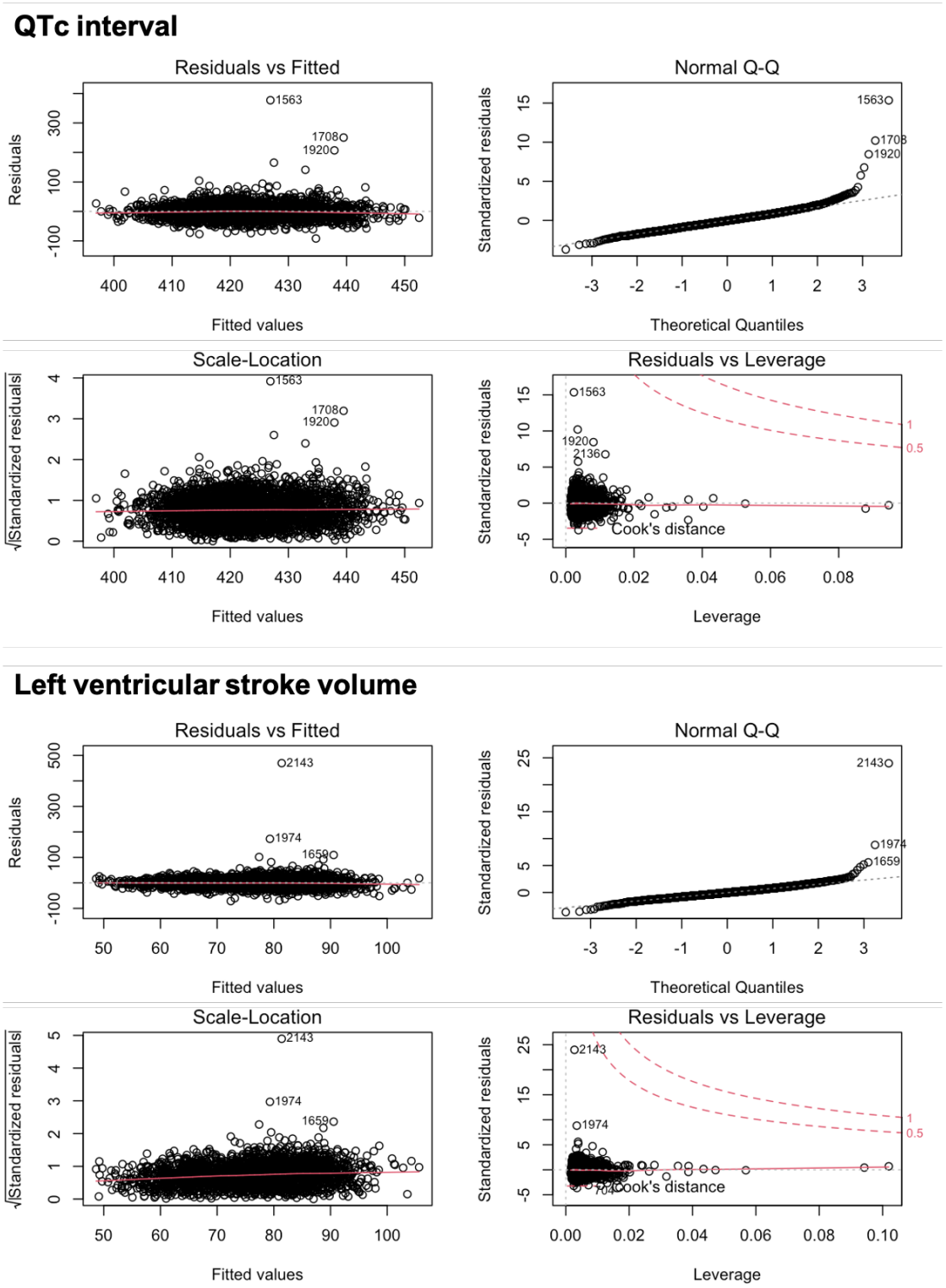

Figure B.1: Residuals vs. fitted values, Q–Q plot, scale–location plot, and residuals vs. leverage plot for QTc interval and left ventricular stroke volume in the combined cohort (combined cohorts with and without type 2 diabetes) for the main linear regression modelling analysis assessing the association between type 2 diabetes and individual ECG and CMR biomarkers. Visual inspection suggests no major violations of model assumptions.

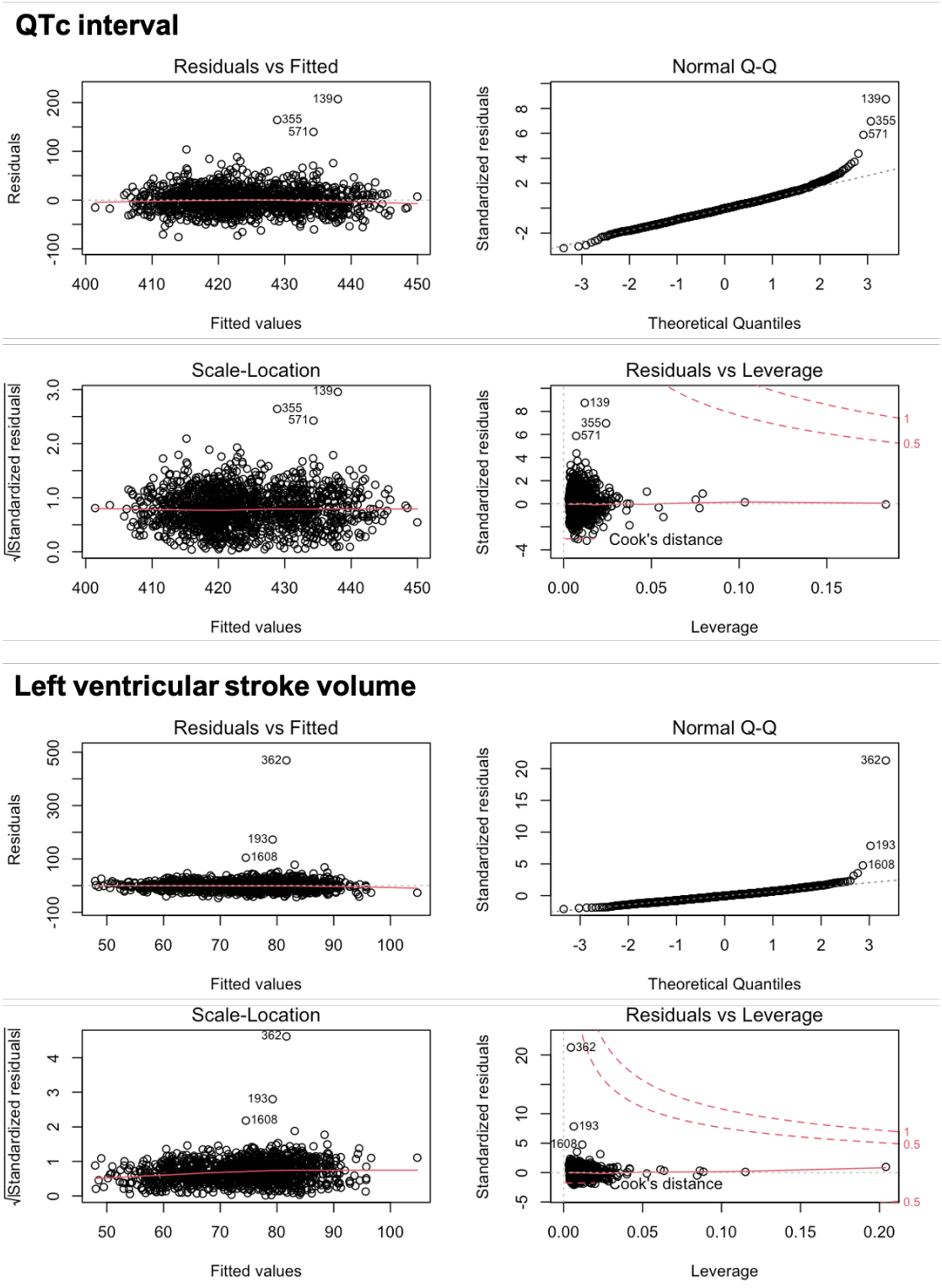

Figure B.2: Residuals vs. fitted values, Q–Q plot, scale–location plot, and residuals vs. leverage plot for QTc interval and left ventricular stroke volume in the type 2 diabetes cohort, for the subgroup analysis assessing the association between HbA1c and individual ECG and CMR biomarkers. Visual inspection suggests no major violations of model assumptions.

Table B.1: Selected ICD-9 and ICD-10 codes for type 2 diabetes and cardiovascular disease used to define the study cohorts. UK Biobank fields corresponding to the first reported date of ICD-10 code are indicated in the right column. No equivalent fields exist for ICD-9 codes.

ICD-X Code	Code	Description	UK Biobank Field
Type 2 Diabetes			
ICD-9	250.x0	Diabetes mellitus, type II or unspecified type, not stated as uncontrolled	-
	250.x2	Diabetes mellitus, type II or unspecified type, uncontrolled	-
ICD-10	E11	Type 2 diabetes mellitus	130708
	E14	Unspecified diabetes mellitus	130714
Cardiovascular Disease			
ICD-9	410	Acute myocardial infarction	-
	411	Other acute and subacute forms of ischaemic heart disease	-
	412	Old myocardial infarction	-
	413	Angina pectoris	-
	414	Other forms of chronic ischaemic heart disease	-
	425	Cardiomyopathy	-
	426	Conduction disorders	-
	427	Cardiac dysrhythmias	-
	428	Heart failure	-
ICD-10	I20	Angina pectoris	131296
	I21	Acute myocardial infarction	131298
	I22	Subsequent myocardial infarction	131300
	I24	Other acute ischaemic heart diseases	131304
	I25	Chronic ischaemic heart disease	131306
	I42	Cardiomyopathy	131338
	I44	Atrioventricular and left bundle-branch block	131342
	I46	Cardiac arrest	131346
	I47	Paroxysmal tachycardia	131348
	I48	Atrial fibrillation and flutter	131350
	I49	Other cardiac arrhythmias	131352
	I50	Heart failure	131354

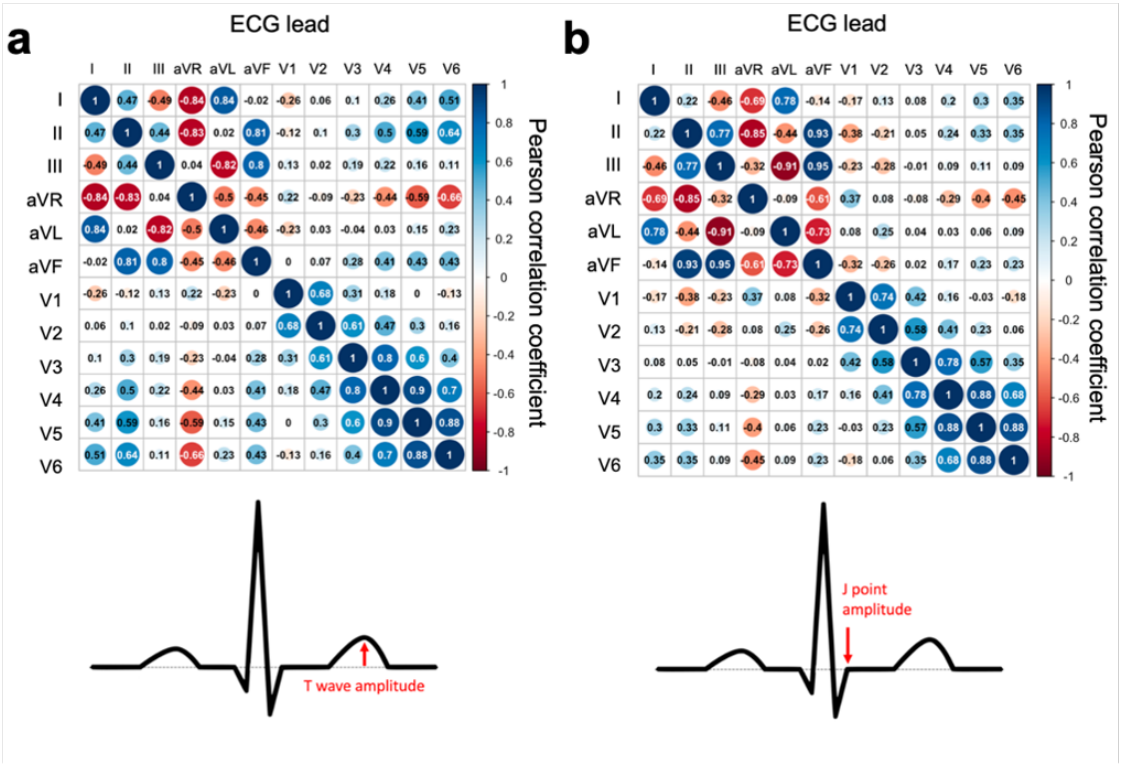

Figure B.3: Matrix of Pearson correlation coefficients for 12-lead ECG amplitude measurements of T wave amplitude (a) and J point amplitude (b). Coefficients were computed on 2702 (a) and 3562 cases (b), respectively. These cases combine participants from both the type 2 diabetes and control cohorts, and contained no missing data among the features of interest.

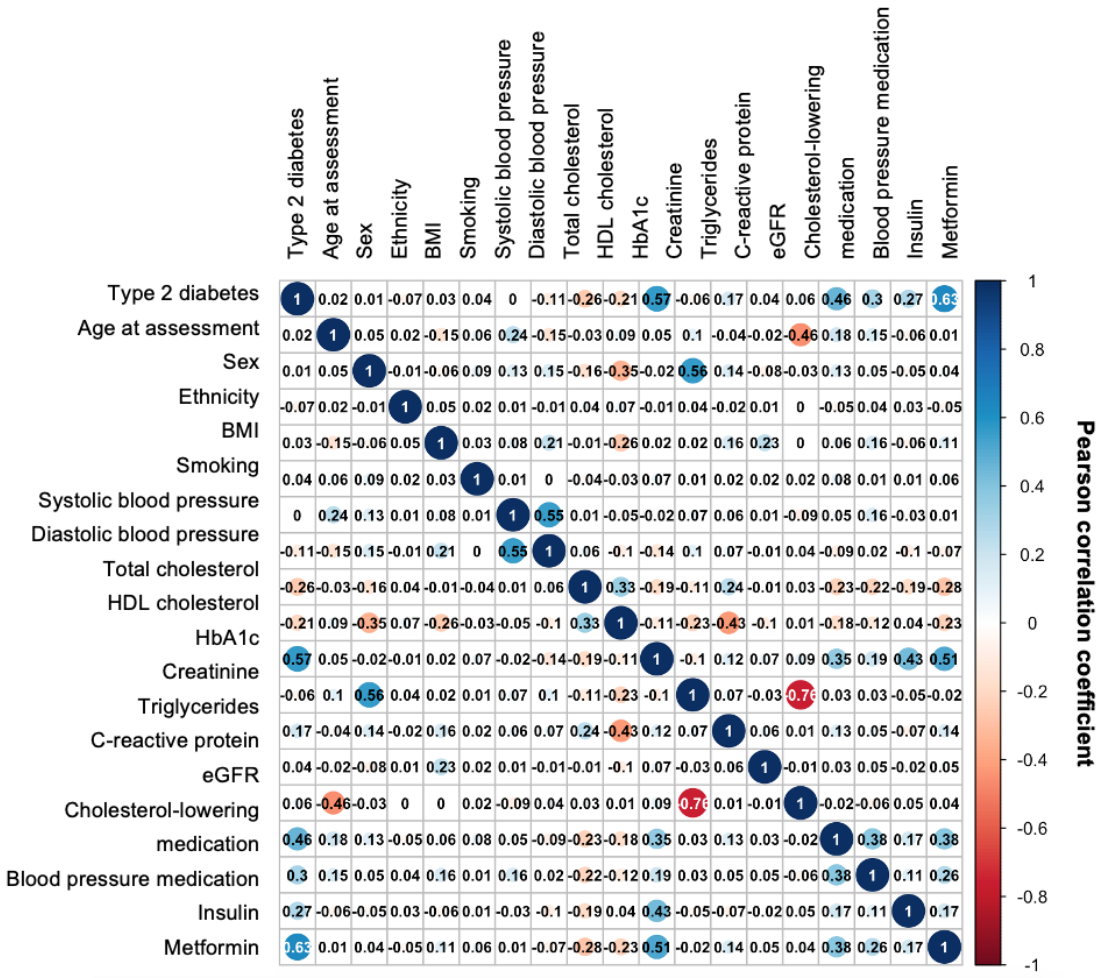

Figure B.4: Matrix of Pearson correlation coefficients between baseline socio-demographic, lifestyle, and clinical measurements. Coefficients were computed on 2620 cases. These cases combine participants from both the type 2 diabetes and control cohorts, and contained no missing data among the features of interest. BMI: body mass index, HDL: high density lipoprotein, eGFR: estimated glomerular filtration rate.

Table B.2: Biomarker differences in the cohort without type 2 diabetes for participants with and without incident cardiovascular disease. CMR: cardiac magnetic resonance, IQR: inter-quartile range, LV: left ventricular, EF: ejection fraction, ED: end-diastolic, ES: end-systolic.

Outcome	**No type 2 diabetes + no incident CVD**		**No type 2 diabetes + incident CVD**		**p-value**
	Median (IQR)	**N (%)**	**Median (IQR)**	**N (%)**	
ECG					
Ventricular rate, bpm	61 (55-68)	1683 (100)	62 (55.2-72.2)	98 (100)	0.218
QRS duration, ms	88 (82-96)	1683 (100)	90 (82-98)	98 (100)	0.078
QTc interval, ms	420 (404-436)	1683 (100)	422 (413-444)	98 (100)	0.100
T-wave offset, ms	854 (834-874)	1683 (100)	848 (824-872)	98 (100)	0.140
J-point amplitude (V3), mV	-0.015 (-0.044-0.019)	1683 (100)	0 (-0.0387-0.034)	98 (100)	0.031
T-wave amplitude (V3), mV	0.366 (0.229-0.533)	1668 (99.1)	0.407 (0.257-0.59)	98 (100)	0.076
J-point amplitude (aVL), mV	0.009 (-0.01-0.029)	1683 (100)	0.009 (-0.015-0.029)	98 (100)	0.570
T-wave amplitude (aVL), mV	0.112 (0.058-0.166)	1623 (96.4)	0.092 (0.019-0.156)	95 (96.9)	0.049
CMR					
Sokolow-Lyon index, mm	20.1 (16.2-24.6)	1564 (92.9)	23.4 (17.7-27.1)	86 (87.8)	0.005
LVEF, %	56 (52-59)	1445 (85.9)	56 (50-60.2)	88 (89.8)	0.682
LV ED volume, ml	140 (119-162)	1445 (85.9)	142 (120-167)	88 (89.8)	0.437
LV ES volume, ml	61 (51-73)	1445 (85.9)	61 (50-78.2)	88 (89.8)	0.447
LV stroke volume, ml	78 (65-90)	1445 (85.9)	79 (65.8-88.2)	88 (89.8)	0.953
Cardiac output, L/min	4.7 (4-5.6)	1445 (85.9)	4.85 (4.17-5.4)	88 (89.8)	0.483
LV mass, g	91.9 (76.1-108)	1425 (84.7)	101 (82.6-118)	93 (94.9)	0.001
LV mass index, g/m^2	46.7 (41.1-52.8)	1425 (84.7)	50.3 (43.0-57.5)	93 (94.9)	0.001
LV wall thickness, mm	5.92 (5.39-6.44)	1423 (84.6)	6.29 (5.55-6.87)	93 (94.9)	<0.001

Table B.3: Biomarker differences in the cohort with type 2 diabetes for participants with and without incident cardiovascular disease. CMR: cardiac magnetic resonance, IQR: inter-quartile range, LV: left ventricular, EF: ejection fraction, ED: end-diastolic, ES: end-systolic.

Outcome	**Type 2 diabetes + no incident CVD**		**Type 2 diabetes + incident CVD**		**p-value**
	Median (IQR)	**N (%)**	**Median (IQR)**	**N (%)**	
ECG					
Ventricular rate, bpm	66 (59-74)	1683 (100)	66 (57-75.8)	98 (100)	0.913
QRS duration, ms	86 (80-94)	1683 (100)	88 (82-98)	98 (100)	0.032
QTc interval, ms	424 (408-440)	1683 (100)	424 (402-441)	98 (100)	0.682
T-wave offset, ms	842 (820-864)	1683 (100)	839 (816-862)	98 (100)	0.325
J-point amplitude (V3), mV	-0.015 (-0.044-0.019)	1683 (100)	-0.005 (-0.044-0.028)	98 (100)	0.472
T-wave amplitude (V3), mV	0.332 (0.205-0.488)	1667 (99)	0.395 (0.191-0.501)	98 (100)	0.343
J-point amplitude (aVL), mV	0.004 (-0.015-0.029)	1683 (100)	0.0065 (-0.01-0.034)	98 (100)	0.395
T-wave amplitude (aVL), mV	0.097 (0.043-0.146)	1582 (94)	0.117 (0.022-0.185)	87 (88.8)	0.294
CMR					
Sokolow-Lyon index, mm	19.1 (15.2-23.4)	1526 (90.7)	20.5 (16.2-25.3)	86 (87.8)	0.094
LVEF, %	56 (51-59)	1412 (83.9)	55 (50-59)	89 (90.8)	0.118
LV ED volume, ml	130 (109-155)	1412 (83.9)	136 (116-155)	89 (90.8)	0.296
LV ES volume, ml	58 (47-71)	1412 (83.9)	62 (51-74)	89 (90.8)	0.106
LV stroke volume, ml	72 (60-85)	1412 (83.9)	72 (58-83)	89 (90.8)	0.808
Cardiac output, L/min	4.7 (4-5.5)	1412 (83.9)	4.8 (4-5.5)	89 (90.8)	0.869
LV mass, g	88.6 (73.9-104)	1398 (83.1)	93.6 (77.1-109)	89 (90.8)	0.139
LV mass index, g/m^2	45.3 (40.4-50.9)	1398 (83.1)	47.4 (42.1-53.4)	89 (90.8)	0.046
LV wall thickness, mm	5.83 (5.32-6.37)	1398 (83.1)	6.01 (5.44-6.45)	89 (90.8)	0.144

Table B.4: Subgroup analysis: biomarker differences by type 2 diabetes status in the female cohort. CMR: cardiac magnetic resonance, IQR: inter-quartile range, LV: left ventricular, EF: ejection fraction, ED: end-diastolic, ES: end-systolic.

Outcome	Female + no type 2 diabetes		Female + type 2 diabetes		p-value
	Median (IQR)	N (%)	Median (IQR)	N (%)	
ECG					
Ventricular rate, bpm	62 (57–70)	648 (100%)	67 (60–74)	648 (100%)	<0.001
QRS duration, ms	82 (76–90)	648 (100%)	82 (76–88)	648 (100%)	0.048
QTc interval, ms	429 (413–443)	648 (100%)	433 (417–447)	648 (100%)	0.001
T-wave offset, ms	858 (838–876)	648 (100%)	850 (826–870)	648 (100%)	<0.001
J-point amplitude (V3), mV	-0.03 (-0.054–0.000)	648 (100%)	-0.025 (-0.054–0.004)	648 (100%)	0.208
T-wave amplitude (V3), mV	0.263 (0.156–0.375)	638 (98.5%)	0.239 (0.141–0.341)	635 (98%)	0.040
J-point amplitude (aVL), mV	0.004 (-0.015–0.029)	648 (100%)	0.004 (-0.015–0.029)	648 (100%)	0.381
T-wave amplitude (aVL), mV	0.097 (0.043–0.156)	615 (94.9%)	0.087 (0.007–0.132)	600 (92.6%)	0.006
CMR					
Sokolow-Lyon index, mm	18.1 (14.8–22.1)	609 (94%)	17.4 (14.2–21.5)	584 (90.1%)	0.019
LVEF, %	57 (54–61)	564 (87%)	57 (53–61)	551 (85%)	0.526
LV ED volume, ml	120 (104–138)	564 (87%)	113 (97–130)	551 (85%)	<0.001
LV ES volume, ml	51 (43–60)	564 (87%)	48 (41–57.5)	551 (85%)	<0.001
LV stroke volume, ml	68 (58–79)	564 (87%)	64 (54–74)	551 (85%)	<0.001
Cardiac output, L/min	4.2 (3.7–4.8)	564 (87%)	4.3 (3.6–4.9)	551 (85%)	0.852
LV mass, g	72.8 (63.9–81.2)	562 (86.7%)	73.2 (65.4–84.2)	556 (85.8%)	0.062
LV mass index, g/m^2	40.1 (36.9–43.6)	562 (86.7%)	41.1 (36.4–45.2)	556 (85.8%)	0.089
LV global average wall thickness, mm	5.31 (4.98–5.67)	561 (86.6%)	5.55 (5.17–5.94)	556 (85.8%)	<0.001

Table B.5: Subgroup analysis: biomarker differences by type 2 diabetes status in the male cohort. CMR: cardiac magnetic resonance, IQR: inter-quartile range, LV: left ventricular, EF: ejection fraction, ED: end-diastolic, ES: end-systolic.

Outcome	Male + no type 2 diabetes		Male + type 2 diabetes		p-value
	Median (IQR)	N (%)	Median (IQR)	N (%)	
ECG					
Ventricular rate, bpm	66 (59–74)	1683 (100%)	66 (57–75.8)	98 (100%)	0.913
QRS duration, ms	86 (80–94)	1683 (100%)	88 (82–98)	98 (100%)	0.032
QTc interval, ms	424 (408–440)	1683 (100%)	424 (402–441)	98 (100%)	0.682
T-wave offset, ms	842 (820–864)	1683 (100%)	839 (816–862)	98 (100%)	0.325
J-point amplitude (V3), mV	-0.015 (-0.044–0.019)	1683 (100%)	-0.005 (-0.044–0.028)	98 (100%)	0.472
T-wave amplitude (V3), mV	0.332 (0.205–0.488)	1667 (99%)	0.395 (0.191–0.501)	98 (100%)	0.343
J-point amplitude (aVL), mV	0.004 (-0.015–0.029)	1683 (100%)	0.0065 (-0.010–0.034)	98 (100%)	0.395
T-wave amplitude (aVL), mV	0.097 (0.043–0.146)	1582 (94%)	0.117 (0.022–0.185)	87 (88.8%)	0.294
CMR					
Sokolow-Lyon index, mm	19.1 (15.2–23.4)	1526 (90.7%)	20.5 (16.2–25.3)	86 (87.8%)	0.094
LVEF, %	56 (51–59)	1412 (83.9%)	55 (50–59)	89 (90.8%)	0.118
LV ED volume, ml	130 (109–155)	1412 (83.9%)	136 (116–155)	89 (90.8%)	0.296
LV ES volume, ml	58 (47–71)	1412 (83.9%)	62 (51–74)	89 (90.8%)	0.106
LV stroke volume, ml	72 (60–85)	1412 (83.9%)	72 (58–83)	89 (90.8%)	0.808
Cardiac output, L/min	4.7 (4–5.5)	1412 (83.9%)	4.8 (4–5.4)	89 (90.8%)	0.665
LV mass, g	90.9 (76.2–108)	1426 (84.7%)	99.6 (85.8–107)	92 (93.9%)	0.006
LV mass index, g/m^2	46.4 (41.2–52.1)	1426 (84.7%)	48.7 (45.3–54.0)	92 (93.9%)	0.001
LV global average wall thickness, mm	6.08 (5.56–6.60)	1424 (84.6%)	6.28 (5.86–6.81)	92 (93.9%)	0.006

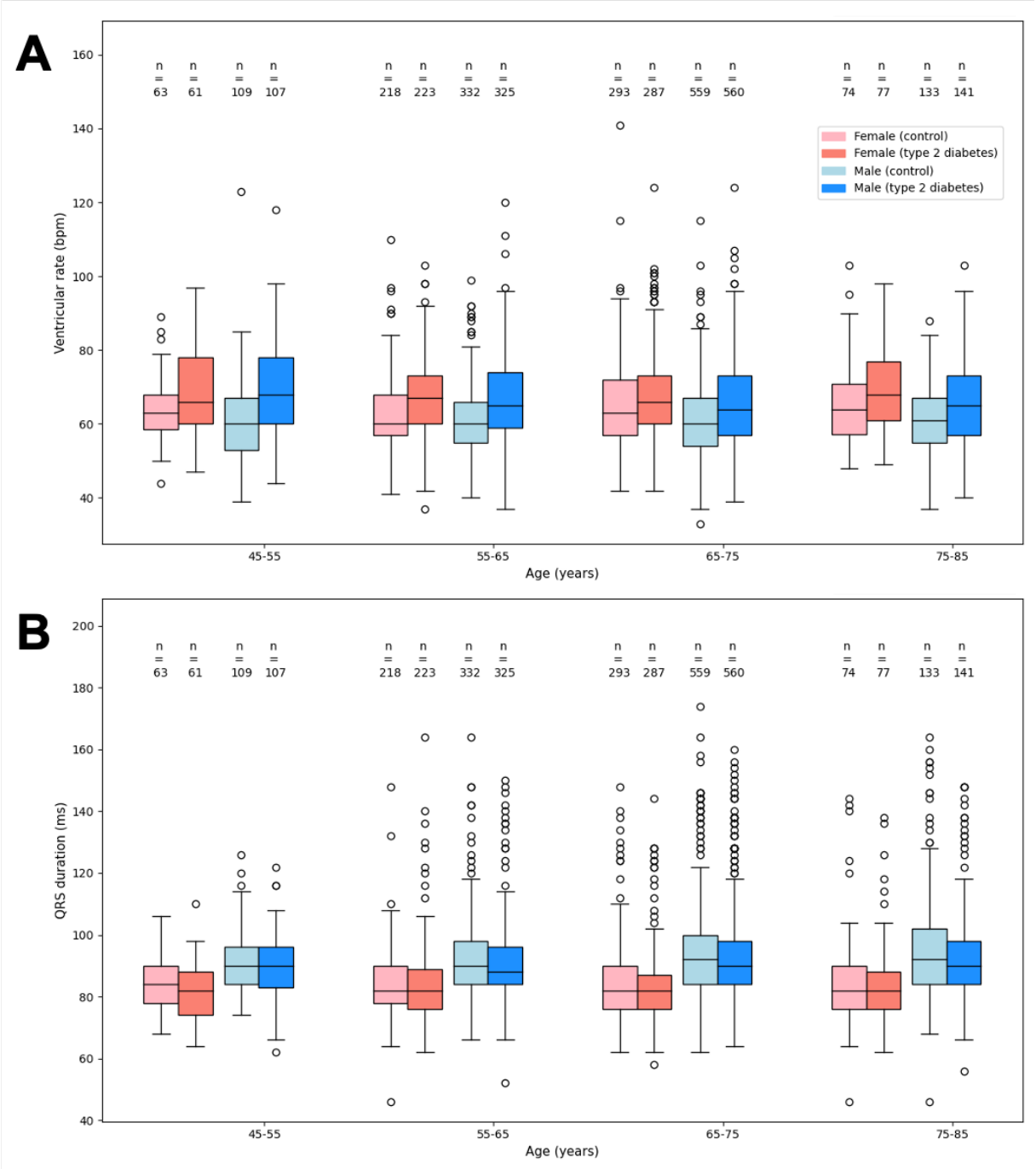

Figure B.5: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by age. Panels A and B: ventricular rate and QRS duration. CMR: cardiac magnetic resonance.

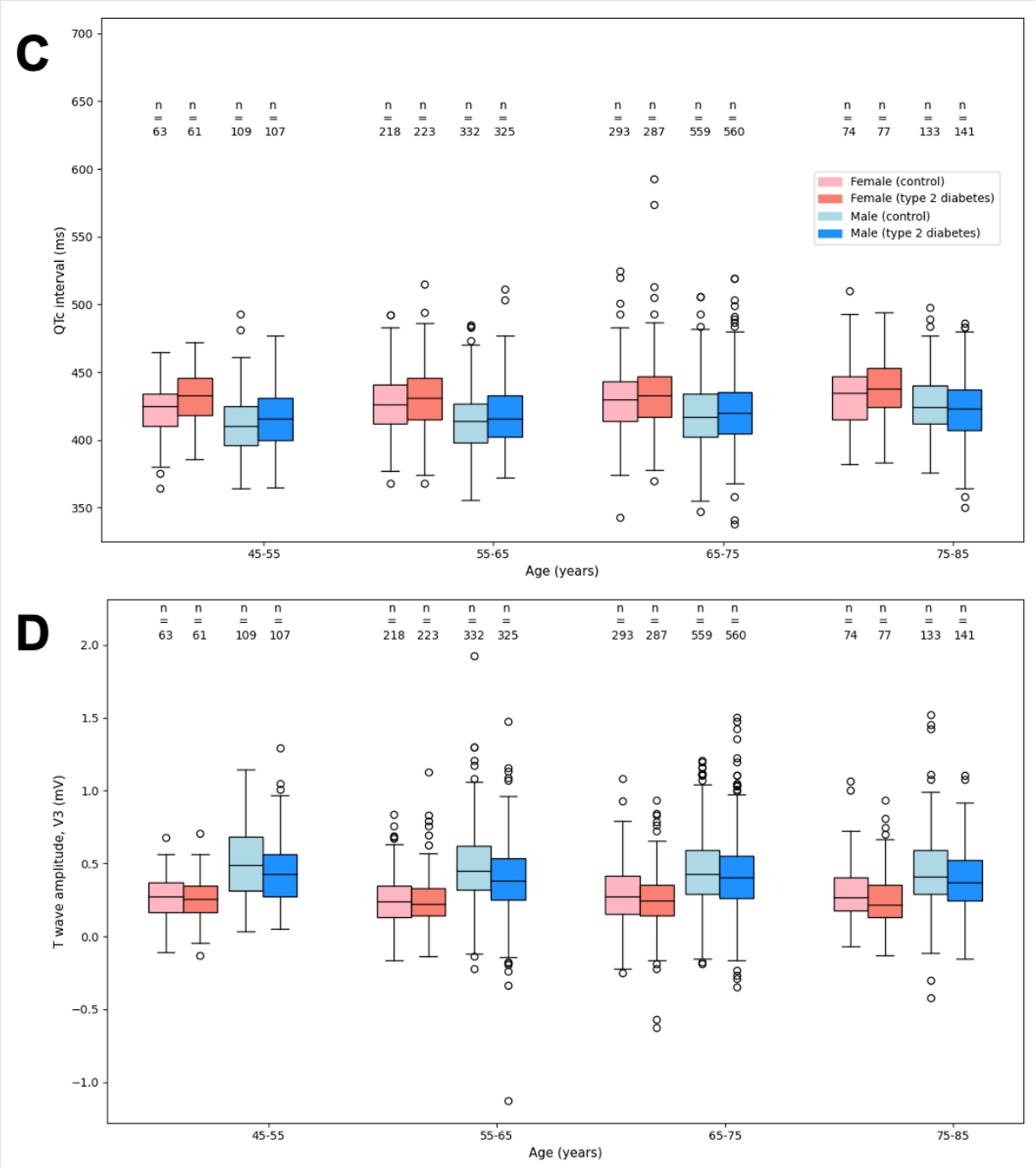

Figure B.5: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by age. Panels C and D: QTc interval and T wave amplitude. Cases with QTc interval > 600 ms. CMR: cardiac magnetic resonance, LV: left ventricular.

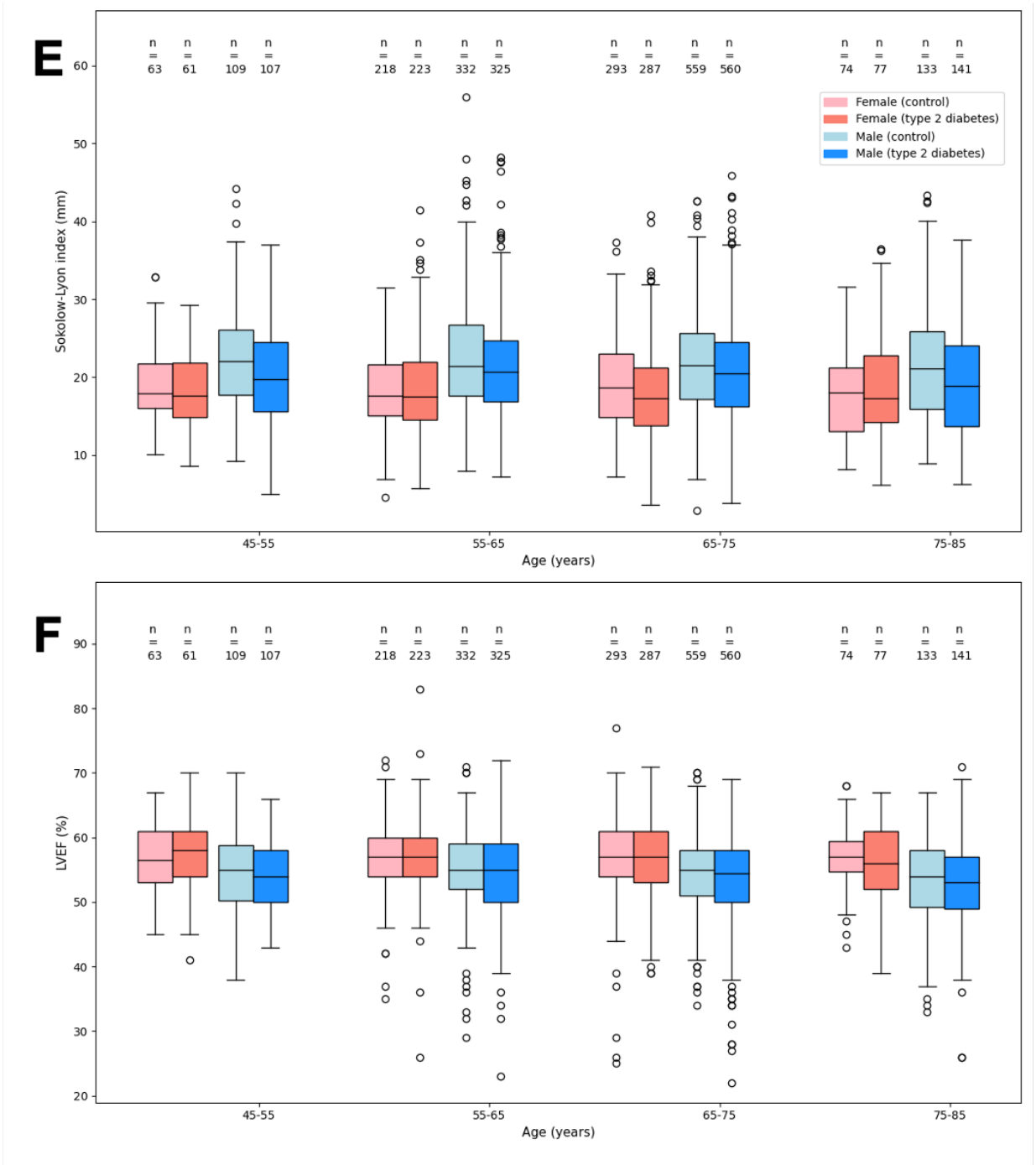

Figure B.5: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by age. Panels E and F: Sokolow-Lyon index and LVEF. Cases with LVEF < 20% are considered outliers and are not shown. CMR: cardiac magnetic resonance, LV: left ventricular, EF: ejection fraction.

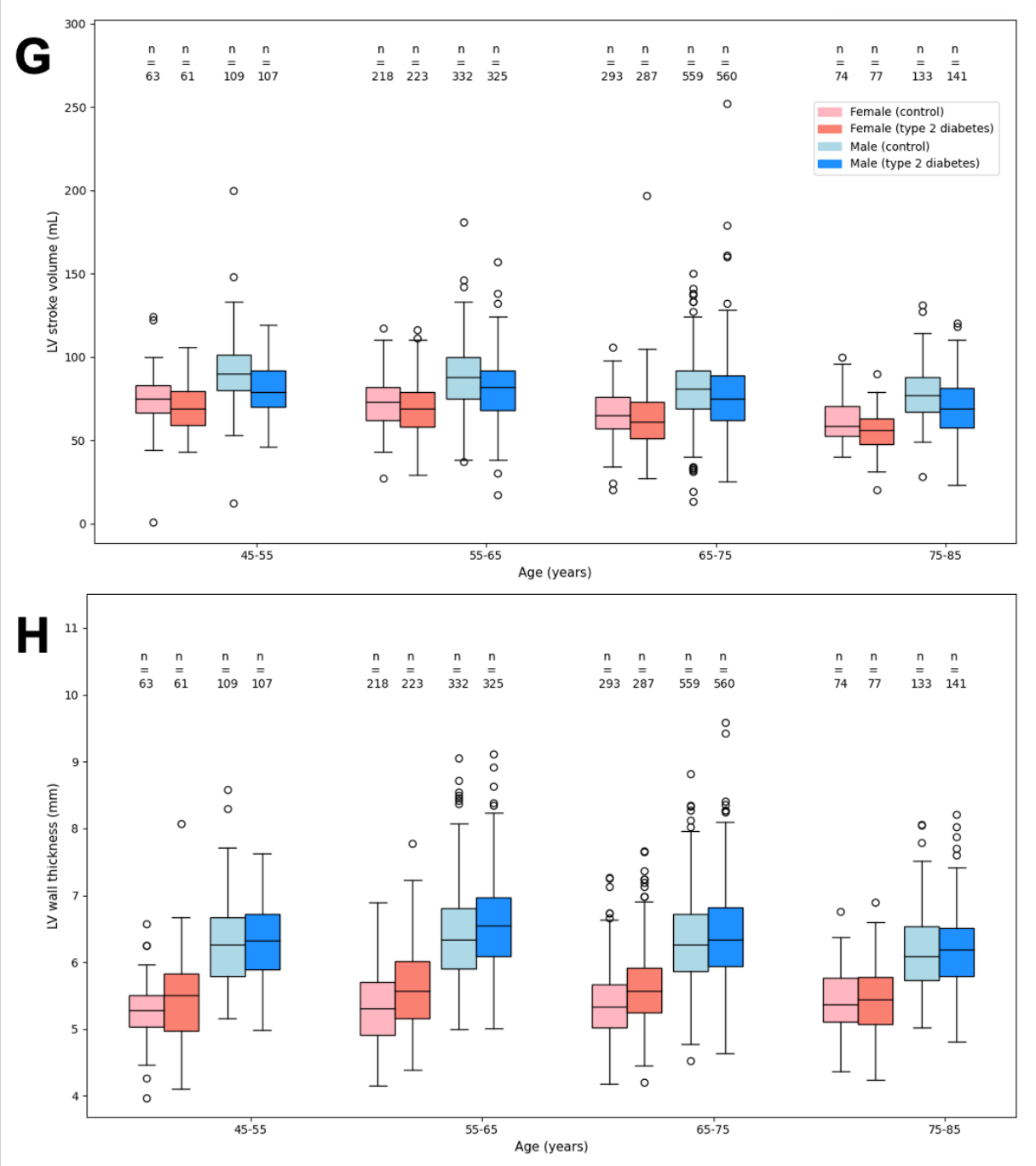

Figure B.5: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by age. Panels G and H: LV stroke volume and LV wall thickness. Cases with LV stroke volume > 300 mL are considered outliers and are not shown. CMR: cardiac magnetic resonance, LV: left ventricular.

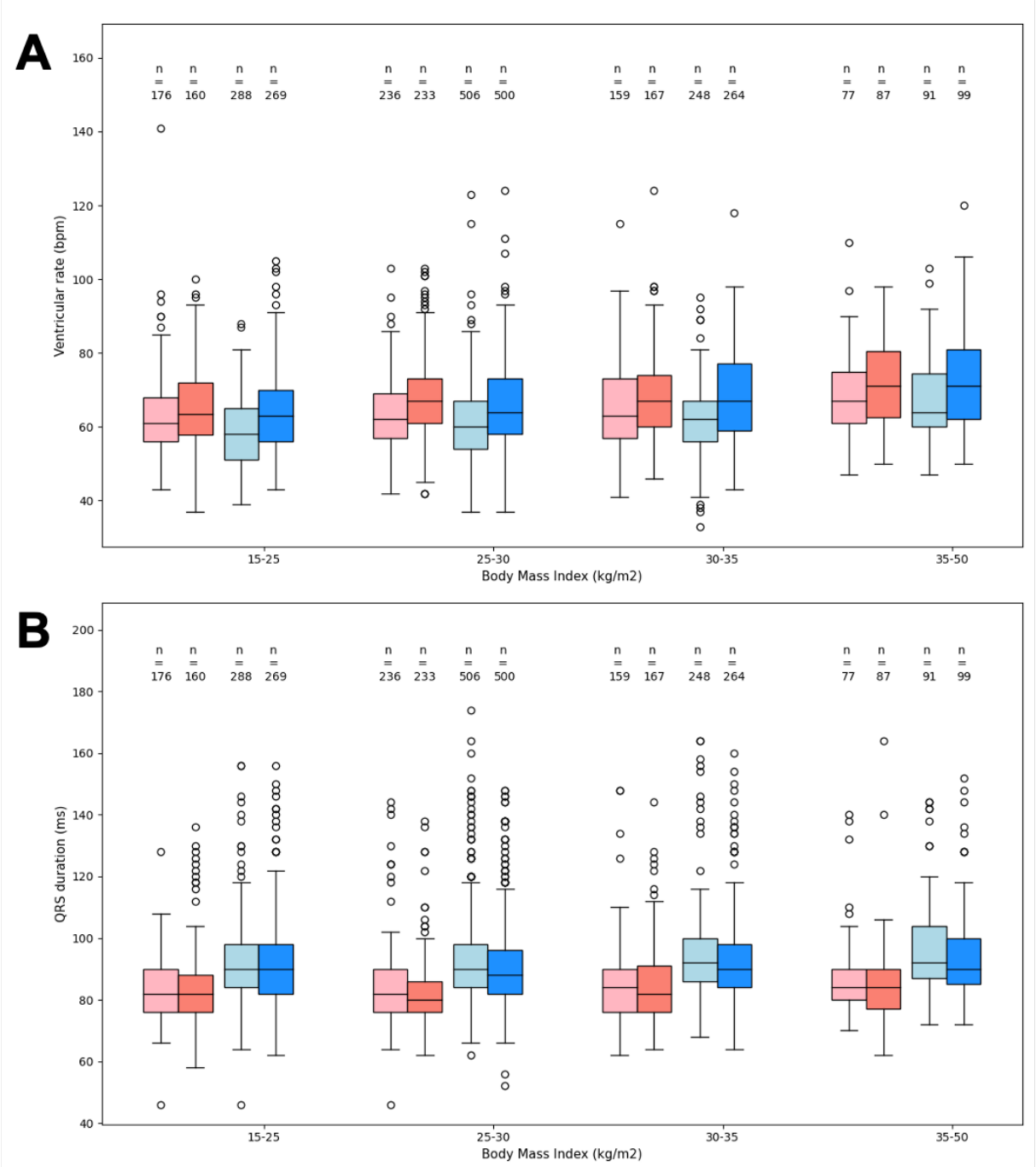

Figure B.6: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by BMI. Panels A and B: ventricular rate and QRS duration. BMI: body mass index, CMR: cardiac magnetic resonance.

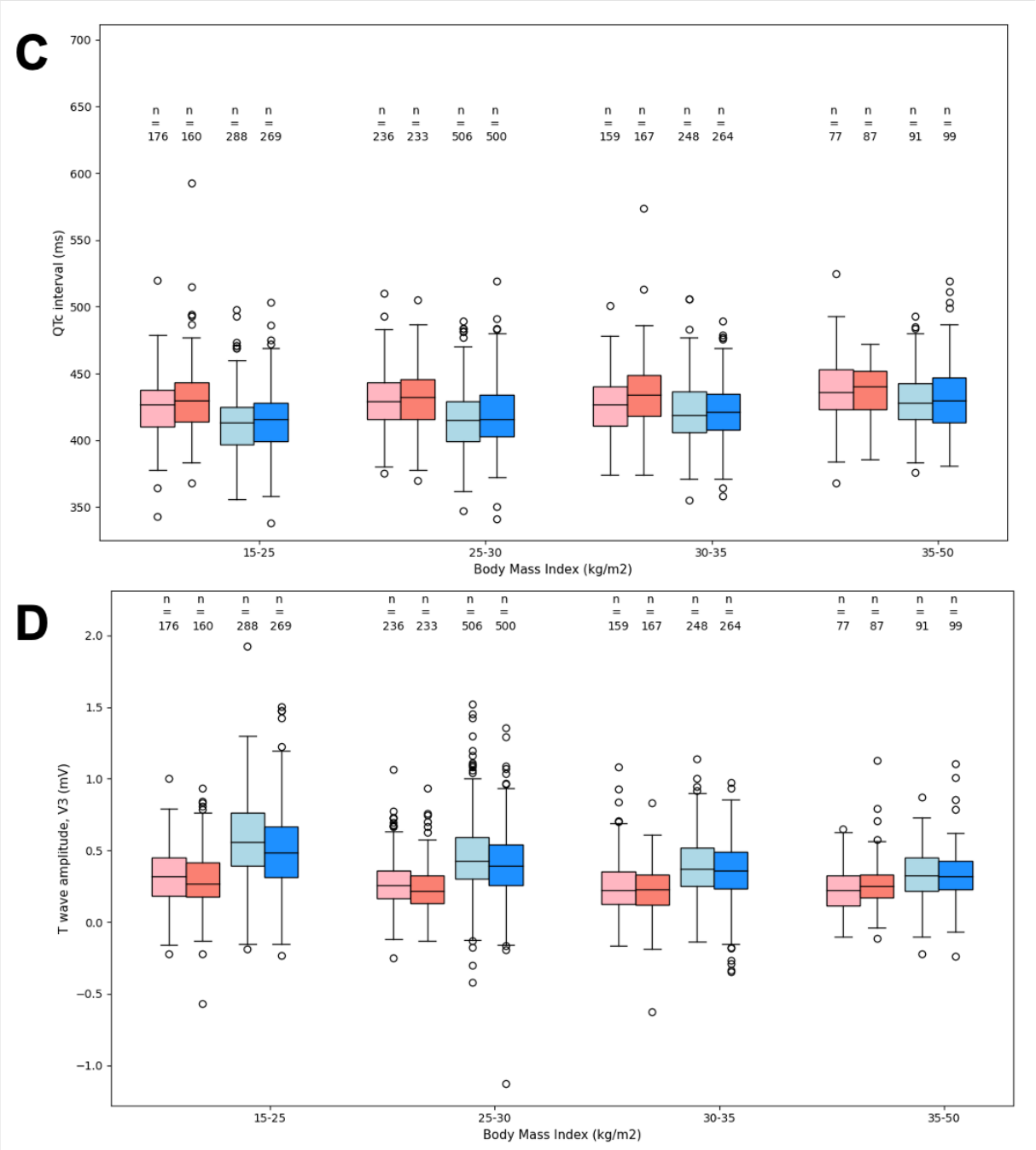

Figure B.6: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by BMI. Panels C and D: QTc interval and T wave amplitude. Cases with QTc interval > 600 ms. BMI: body mass index, CMR: cardiac magnetic resonance, LV: left ventricular.

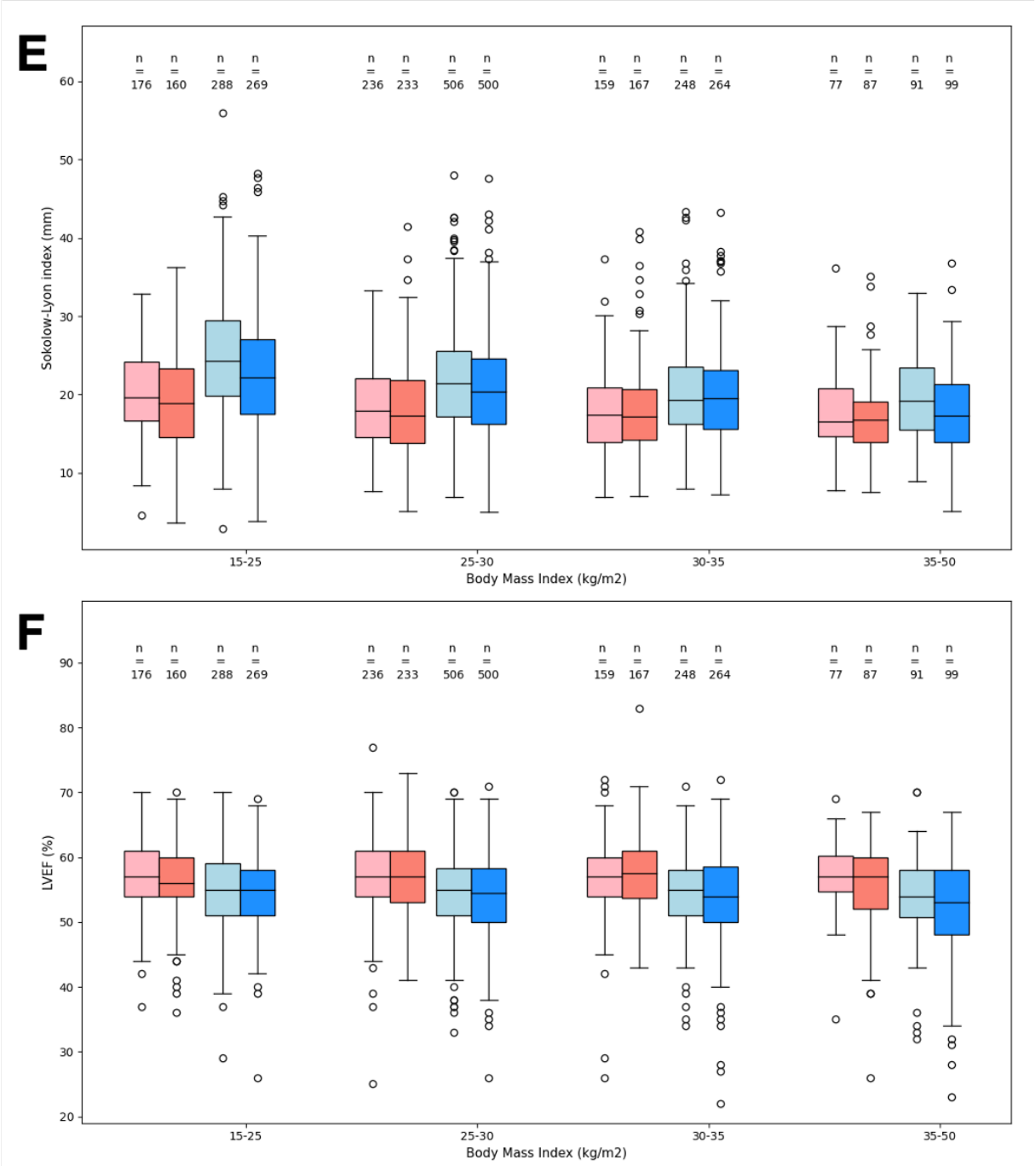

Figure B.6: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by BMI. Panels E and F: Sokolow-Lyon index and LVEF. Cases with LVEF < 20% are considered outliers and are not shown. BMI: body mass index, CMR: cardiac magnetic resonance, LV: left ventricular, EF: ejection fraction.

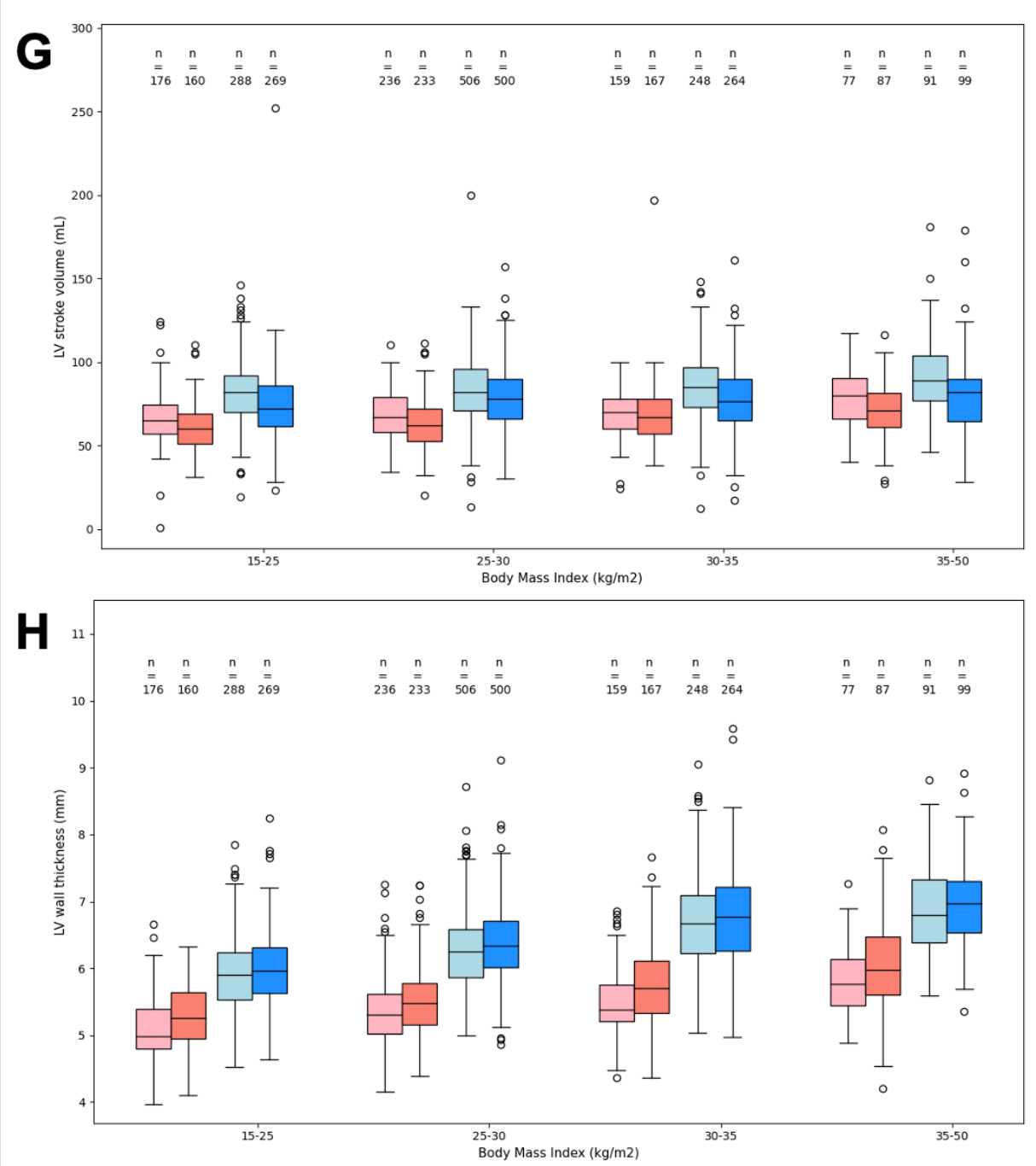

Figure B.6: Distribution of sex-specific ECG and CMR-derived biomarkers stratified by BMI. Panels G and H: LV stroke volume and LV wall thickness. Cases with LV stroke volume > 300 mL are considered outliers and are not shown. BMI: body mass index, CMR: cardiac magnetic resonance, LV: left ventricular.

Table B.6: Subgroup analysis: biomarker differences by type 2 diabetes status in the white ethnic background cohort. CMR: cardiac magnetic resonance, IQR: inter-quartile range, LV: left ventricular, EF: ejection fraction, ED: end-diastolic, ES: end-systolic.

Outcome	White + no type 2 diabetes		White + type 2 diabetes		p-value
	Median (IQR)	N (%)	Median (IQR)	N (%)	
ECG					
Ventricular rate, bpm	61 (55–68)	1733 (100%)	66 (59–74)	1640 (100%)	<0.001
QRS duration, ms	88 (82–96)	1733 (100%)	86 (80–96)	1640 (100%)	0.013
QTc interval, ms	421 (405–436)	1733 (100%)	424 (408–440)	1640 (100%)	<0.001
T-wave offset, ms	854 (834–874)	1733 (100%)	842 (820–864)	1640 (100%)	<0.001
J-point amplitude (V3), mV	-0.015 (-0.044–0.019)	1733 (100%)	-0.015 (-0.044–0.019)	1640 (100%)	0.699
T-wave amplitude (V3), mV	0.366 (0.229–0.537)	1719 (99.2%)	0.332 (0.205–0.488)	1625 (99.1%)	<0.001
J-point amplitude (aVL), mV	0.009 (-0.010–0.029)	1733 (100%)	0.004 (-0.015–0.029)	1640 (100%)	0.021
T-wave amplitude (aVL), mV	0.107 (0.053–0.166)	1671 (96.4%)	0.097 (0.039–0.146)	1534 (93.5%)	<0.001
CMR					
Sokolow-Lyon index, mm	20.1 (16.2–24.7)	1610 (92.9%)	19.1 (15.2–23.4)	1488 (90.7%)	<0.001
LVEF, %	56 (52–59)	1495 (86.3%)	55 (51–59)	1389 (84.7%)	0.041
LV ED volume, ml	141 (119–163)	1495 (86.3%)	131 (110–156)	1389 (84.7%)	<0.001
LV ES volume, ml	61 (51–74)	1495 (86.3%)	58 (48–72)	1389 (84.7%)	<0.001
LV stroke volume, ml	78 (65–90)	1495 (86.3%)	72 (60–85)	1389 (84.7%)	<0.001
Cardiac output, L/min	4.7 (4.1–5.6)	1495 (86.3%)	4.7 (4.0–5.5)	1389 (84.7%)	0.730
LV mass, g	92.5 (76.3–109)	1480 (85.4%)	91.6 (77.2–108)	1408 (85.9%)	0.555
LV mass index, g/m^2	46.9 (41.1–53.0)	1480 (85.4%)	46.8 (41.4–52.3)	1408 (85.9%)	0.700
LV global average wall thickness, mm	5.92 (5.4–6.46)	1478 (85.3%)	6.09 (5.6–6.62)	1406 (85.7%)	<0.001

Table B.7: Subgroup analysis: biomarker differences by type 2 diabetes status in the non-white ethnic background cohort. CMR: cardiac magnetic resonance, IQR: inter-quartile range, LV: left ventricular, EF: ejection fraction, ED: end-diastolic, ES: end-systolic.

Outcome	Non-white + no type 2 diabetes		Non-white + type 2 diabetes		p-value
	Median (IQR)	N (%)	Median (IQR)	N (%)	
ECG					
Ventricular rate, bpm	59.5 (56.5–66)	48 (100%)	67 (59–76)	141 (100%)	<0.001
QRS duration, ms	84 (80–90.5)	48 (100%)	84 (78–92)	141 (100%)	0.871
QTc interval, ms	415 (397–432)	48 (100%)	420 (406–440)	141 (100%)	0.141
T-wave offset, ms	851 (833–867)	48 (100%)	836 (812–860)	141 (100%)	0.004
J-point amplitude (V3), mV	-0.008 (-0.045–0.043)	48 (100%)	0 (-0.040–0.048)	141 (100%)	0.406
T-wave amplitude (V3), mV	0.375 (0.229–0.556)	47 (97.9%)	0.312 (0.194–0.463)	140 (99.3%)	0.188
J-point amplitude (aVL), mV	0.024 (0.003–0.048)	48 (100%)	0.009 (-0.015–0.034)	141 (100%)	0.027
T-wave amplitude (aVL), mV	0.117 (0.085–0.188)	47 (97.9%)	0.112 (0.063–0.170)	135 (95.7%)	0.130
CMR					
Sokolow-Lyon index, mm	20.9 (17.4–25.9)	40 (83.3%)	20.4 (15.5–25.7)	124 (87.9%)	0.368
LVEF, %	56 (51–60)	38 (79.2%)	56 (53–60)	112 (79.4%)	0.898
LV ED volume, ml	124 (115–146)	38 (79.2%)	119 (97.8–139)	112 (79.4%)	0.146
LV ES volume, ml	56.5 (47–71)	38 (79.2%)	51 (43–66)	112 (79.4%)	0.156
LV stroke volume, ml	68.5 (62–83.5)	38 (79.2%)	66 (54–78)	112 (79.4%)	0.143
Cardiac output, L/min	4.45 (3.73–5.08)	38 (79.2%)	4.5 (3.6–5.3)	112 (79.4%)	0.782
LV mass, g	93.3 (81.2–103)	38 (79.2%)	86.5 (70.3–103)	110 (78%)	0.285
LV mass index, g/m^2	46.9 (42.4–52.7)	38 (79.2%)	45.8 (39.8–52.5)	110 (78%)	0.189
LV global average wall thickness, mm	6.26 (5.51–6.68)	38 (79.2%)	6.01 (5.51–6.55)	110 (78%)	0.484

Table B.8: Subgroup analysis: changes in biomarkers within the glycemic spectrum. Multivariate multiple linear regression models to quantify the association of HbA1c with selected ECG and CMR-derived biomarkers. Models are adjusted sequentially for different types of confounding factors. Socio-demographic factors include age, sex, ethnicity; lifestyle factors include body mass index (BMI), smoking; clinical factors include diastolic blood pressure, total cholesterol, triglycerides, C-reactive protein, anti-hypertensive medication and insulin. CMR: cardiac magnetic resonance imaging.

	Model 0 Unadjusted N = 1170		**Model 1 Adjusted for socio-demographic factors N = 1170**		**Model 2 Additionally adjusted for lifestyle factors N = 1154**		**Model 3 Additionally adjusted for clinical factors N = 1014**	
Outcome	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**	**Coefficient (95% CI)**	**p-value**
ECG								
Ventricular rate, bpm	0.095 (0.037–0.153)	0.001	0.094 (0.036–0.151)	0.001	0.097 (0.040–0.154)	0.001	0.12 (0.054–0.187)	0.000
QRS duration, ms	-0.020 (-0.088–0.047)	0.557	-0.009 (-0.073–0.056)	0.790	-0.0114 (-0.076–0.054)	0.731	0.011 (-0.063–0.086)	0.768
QTc interval, ms	0.012 (-0.102–0.126)	0.835	0.004 (-0.105–0.114)	0.937	0.009 (-0.100–0.118)	0.871	0.026 (-0.106–0.157)	0.702
T-wave offset, ms	-0.26 (-0.425—0.096)	0.002	-0.267 (-0.429—0.105)	0.001	-0.279 (-0.441—0.116)	0.001	-0.33 (-0.522—0.138)	0.001
T-wave amplitude (V3), mV	0.0001 (-0.001–0.001)	0.859	0.0003 (-0.001–0.001)	0.590	0.0002 (-0.001–0.001)	0.642	-0.001 (-0.002–0.000)	0.190
T-wave amplitude (aVL), mV	-0.001 (-0.002–0.000)	0.108	-0.001 (-0.002–0.000)	0.098	-0.001 (-0.002–0.000)	0.087	-0.002 (-0.003—0.001)	0.034
J-point amplitude (V3), mV	0.000 (-0.0001–0.0003)	0.547	0.0001 (-0.0001–0.0004)	0.384	0.000 (-0.0001–0.0004)	0.466	0.000 (-0.0003–0.0004)	0.816
J-point amplitude (aVL), mV	0.000 (-0.0002–0.0002)	0.869	0.000 (-0.0002–0.0002)	0.817	0.000 (-0.0002–0.0002)	0.925	0.000 (-0.0002–0.0002)	0.784
Sokolow-Lyon index, mm	0.010 (-0.023–0.042)	0.555	0.010 (-0.022–0.042)	0.558	0.008 (-0.024–0.039)	0.638	0.006 (-0.031–0.044)	0.736
CMR								
LV ED volume, ml	-0.363 (-0.837–0.111)	0.133	-0.347 (-0.814–0.121)	0.146	-0.323 (-0.797–0.151)	0.182	-0.256 (-0.852–0.340)	0.400
LV stroke volume, ml	-0.135 (-0.256—0.014)	0.029	-0.132 (-0.247—0.018)	0.023	-0.122 (-0.237—0.007)	0.038	-0.104 (-0.243–0.036)	0.145
LV global average wall thickness, mm	-0.002 (-0.006–0.002)	0.378	-0.001 (-0.005–0.002)	0.408	-0.0003 (-0.003–0.003)	0.860	0.001 (-0.003–0.004)	0.607

C
Appendix for Chapter 6

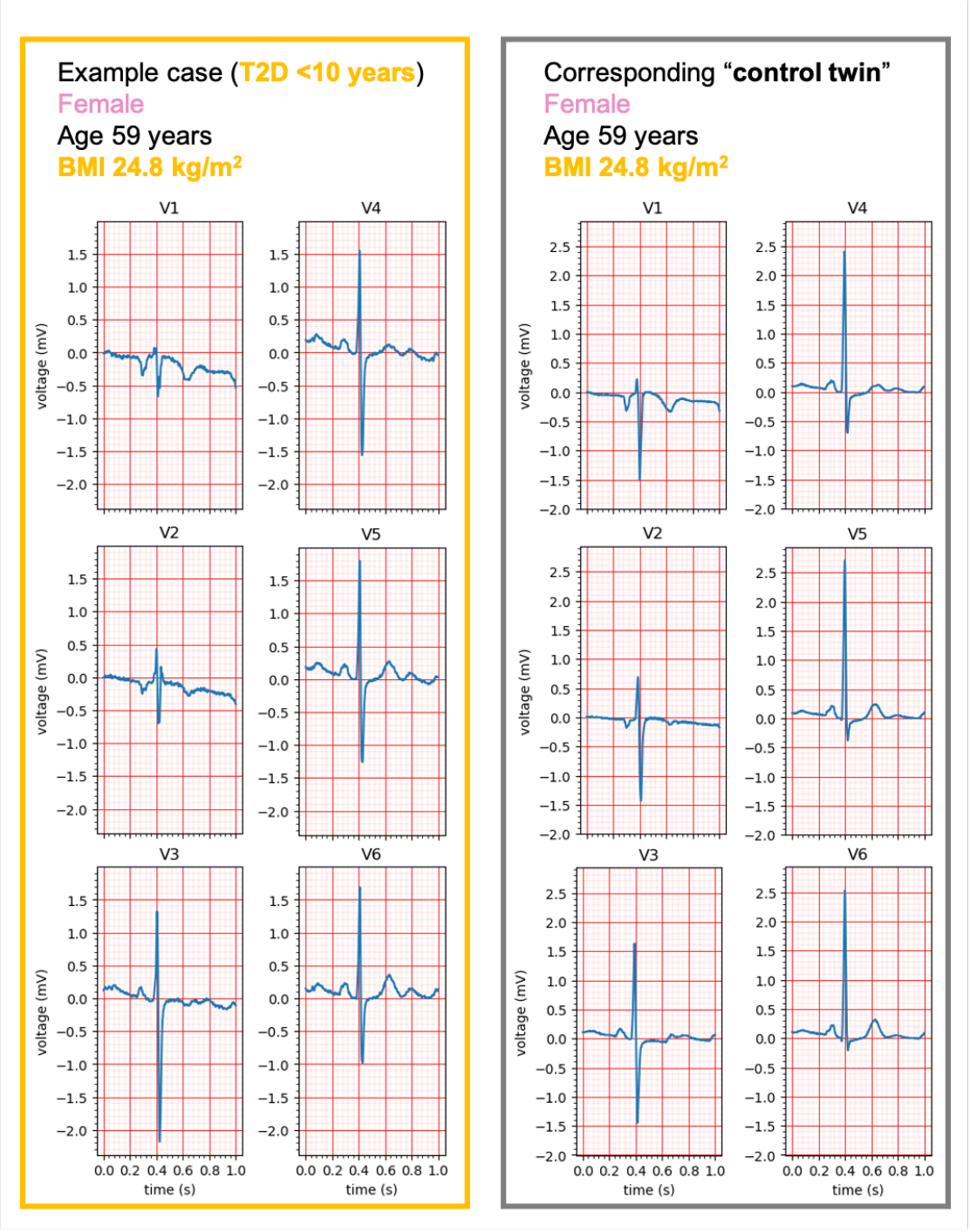

Figure C.1: Illustrative comparison of a pair ECG chest leads from a selected UK Biobank type 2 diabetes subject and their corresponding "control twin".

Table C.1: List of selected drugs with a known QT-prolonging effect considered as exclusion criteria in the case selection process. Class Ia anti-arrhythmics moderately block fast sodium and potassium channels, Class Ic anti-arrhythmics are potent sodium channel blockers, Class III are potassium channel blockers. ATC: Anatomical Therapeutic Chemical classification system, SSRI: Selective Serotonin Reuptake Inhibitor.

ATC code	**Name**	**Class**
C01BA01	Quinidine	Class Ia anti-arrhythmic
C01BA02	Procainamide	Class Ia anti-arrhythmic
C01BC04	Flecanaide	Class Ic anti-arrhythmic
C01BD01	Amiodarone	Class III anti-arrhythmic
C01BD04	Dofetilide	Class III anti-arrhythmic
C07AA07	Sotalol	Non-selective beta blocker
J01FA01	Erythromycin	Macrolide antibiotic
J02AC01	Fluconazole	Antifungal
N06AB04	Citalopram	SSRI antidepressant
N05AD01	Haloperidol	Antipsychotic
A04AA01	Ondansetron	Antiemetic (5HT3)

C.1 Additional methods

C.1.1 Virtual torso and cardiac anatomy reconstruction

Step 1: Generating sparse points clouds of the heart

For each selected UK Biobank participant, I converted the set of original 2D CMR image files into an initial 3D representation of the given individual's heart as points scattered in 3D virtual space representing the tissue, known as a point cloud [290, 406]. The files were read in their raw Digital Imaging and Communications in Medicine (DICOM) format, and relevant 2D image slices automatically selected. This took about 30s per case. Then, I obtained contours of the myocardium from these slices by deploying a pre-trained convolutional neural network, enabling rapid and automatic segmentation of the ventricles. This network was pre-trained on over 90,000 short-axis and long-axis images, annotated by clinical experts, from 4,875 subjects in the UK Biobank [289]. The test set was made up of 600 cases, of which 39 with prior identified CVD which showed very similar performance metrics compared to the complete test set, suggesting that the network is able to segment the myocardium accurately despite potential structural abnormalities. More details on the protocol, validation, and results of this work can be found in [289]. The deployed segmentation network took between 1-2min per case, with a further final 30s post-processing step to generate the sparse point clouds from the 2D contours. For each case, I generated three separate point clouds, namely one for the left ventricular endocardium, one for the left ventricular epicardium, and one for the right ventricle, which is too thin to separate transmurally.

Step 2: Converting point clouds from a sparse to a dense representation

I converted the sparse point clouds generated in Step 2 to dense point clouds containing a higher density of points in the areas that represent the myocardium. First, contours from the input point clouds were extracted and smoothed. Then, the coordinate system was adjusted to align the point clouds to a common virtual space and to correct short-axis slice misalignment in-plane and out-of-plane. Finally, I applied a point cloud completion network to generate point clouds from the corrected 3D contours of the sparse point clouds. This network was developed and pre-trained on selected image slices from 1000 UK Biobank participants in a previous study [406].

Step 3: Surface mesh generation

I converted the dense point clouds generated in Step 3 to triangular surface meshes, connecting independent points in the dense point clouds to their nearest neighbours using a ball pivoting algorithm, which was successfully applied for the same purpose in previous works [406]. This algorithm forms triangles by "pivoting" a virtual ball of fixed radius over the point cloud, using local surface normals to identify triplets of points that the ball simultaneously touches. Starting from an initial triangle, the ball pivots around the mesh edges, iteratively adding new triangles, effectively capturing the surface's topology. This short but crucial step of the pipeline was achieved with the technical support of Dr Ruben Doste (Department of Computer Science, University of Oxford).

Step 4: Volumetric mesh generation

In this final step, I applied a novel pipeline [408] that converts surface meshes into coarse volumetric meshes made up of tetrahedral elements, finer volumetric meshes made up of hexahedral elements, and a number of fields containing descriptive anatomical characteristics including cell transmurality, myofiber orientation, tissue type (e.g. healthy,

fibrotic, etc.), and ionic gradients, determined using the using the Cobiveco universal biventricular coordinates [440]. Together, these files represent a comprehensively detailed virtual cardiac geometry ready to use for whole-organ simulations. I set the resolution of the fine hexahedral meshes to 400µm.

Step 5: Torso mesh generation and virtual ECG lead locations

For each case, I also generated patient-specific virtual torso anatomies and ECG lead locations by adapting an existing pipeline [407]. Similarly to Step 1, this pipeline reads in raw 2D UK Biobank DICOM files and automatically segments the patient's torso, returning a smooth surface mesh as well as coordinates of 10 virtual ECG leads, placed at relevant limb and chest locations on the body's surface.

Software requirements and reproducibility

Step 1 required a Windows operating system and MATLAB, version 2014a, for the code to run correctly. Additionally required a Microsoft Visual C++ redistributable, 2013 version. Microsoft Visual C++ 2022 and beyond are not compatible with MATLAB R2014a; the older version is required to run MATLAB executable (MEX) files generated. The segmentation part of Step 1 is run using Python 3.7 externally to the main MATLAB environment.

Step 2 requires the installation of the following Microsoft Visual C++ packages: C++ Development Tools, Desktop development with C++ (under "Workloads"), C++ CMake tools for Windows and MSVC v143 VS 2022 C++ x64/x86 build tools (under "Individual Components").

Step 3 and **Step 4** requires MATLAB version 2021b or later, and Windows 10 or a Linux operating system. An NVIDIA GPU is recommended for better performance.

Step 5 requires MATLAB 2014a on Windows, along with the pipeline's own requirements detailed in [408]. The machine I used to generate the meshes had a Windows 11 operating system, an NVIDIA Quadro P400 graphics card, and Intel(R) UHD Graphics P630.

C.1.2 Automatic ECG delineation tool

To analyse the simulated ECGs generated in Chapter 6, I developed an automatic delineation algorithm to detect the QRS complex and T wave based on signal deviations from its baseline.

The algorithm defines the onset of the QRS complex as the first point where the signal exceeds 5% of its maximum absolute amplitude. The end of the QRS is then determined by looking for a return to baseline, using a sliding window of 20 ms (20 samples, given a sampling rate of 1 ms per sample) where no more than two samples cross the threshold. The QT interval is measured from the start of the QRS complex to the end of the T wave, which I define as the last clear deflection in the signal followed by a return to baseline, again using the same sliding window approach.

To handle cases where the signal doesn't return to baseline cleanly, I included fallback rules. If the end of the QRS complex isn't found within the expected window, the algorithm defaults to a fixed duration of 150 ms from the QRS onset; a conservative value, given that healthy QRS durations typically fall between 80-120 ms. Similarly, if the T wave offset isn't clearly detected, for example due to a broad or elevated T wave, the algorithm uses the end of the signal as the QT endpoint. These fallback detections are flagged during later analysis so they can be reviewed or excluded if needed. Overall, this approach allows for consistent interval measurements while still being flexible enough to handle a range of signal shapes.

Finally, the algorithm extracts the T wave amplitude by identifying the maximum and minimum deflections within the segment between the end of the QRS complex and the end of the QT interval. The largest of these deflections, in absolute value, is taken as

the T wave peak, and its amplitude is measured relative to the signal baseline, defined as the median of the last 50 samples of the ECG trace. The index of this peak within the full signal is also recorded to locate the T wave position in time.

The algorithm was tested and validated using my own simulated ECGs only; further validation would be required to use the algorithm on ECGs of different format. Outputs of this algorithm are illustrated in one ECG lead for each of my six selected cases (Figure C.2).

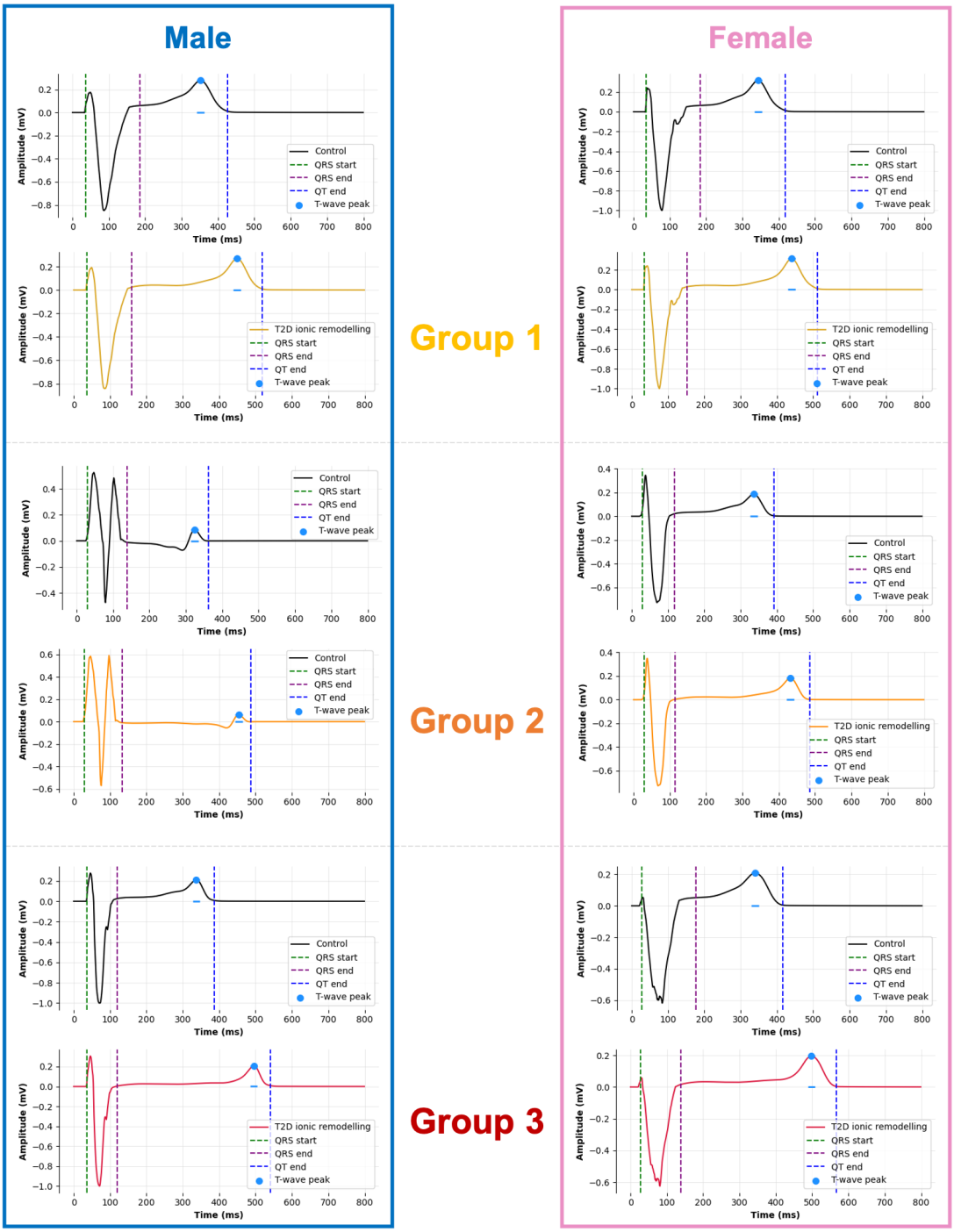

Figure C.2: QRS complex, QT interval and T wave amplitude automatically delineated in the simulated ECGs. Traces in lead V6 are illustrated for each case, under baseline control conditions (black signal) and type 2 diabetes ionic remodelling (yellow, orange, red signals). Groups 1, 2, and 3 refer to the disease-specific groups presented in Section 6.2.1.