Trinity 2021 - University of Oxford

Cardiovascular Disease under the Lens of Automated 3D Analysis of the Heart



Jorge Corral Acero

Oxford Institute of Biomedical Engineering

St. Hugh's College

Supervisors:

Prof. Vicente Grau, Prof. Alfonso Bueno-Orovio & Prof. Pablo Lamata

Thesis submitted for the degree of

Doctor of Philosophy

***

A mi padre y a mi madre, que me dieron la vida, y la alegría para vivirla.

***

Acknowledgements

First of all, I would like to thank my supervisors, Vicente, Alfonso, and Pablo, for all their support and guidance over the last three years. Thanks for your endless patience, your generous availability, the freedom and trust you gave me from the very first day, and your always positive mindset. It has been a pleasure and a privilege to work side-by-side and learn from you. I really appreciate your huge involvement. You have been much more than supervisors for me. Likewise, I would like to acknowledge Ernesto for being such a great mentor and friend during my PhD. He knows literally everything about medical imaging and always has a practical solution ready for any sort of problem. I miss a lot our lunch breaks and the late evening discussions.

I would also like to thank the rest of colleagues from my research group for the enriching Cake Club discussions and for making the lab such a wonderful environment to work in. The same applies to my PIC fellows. I would never have expected to find such a brilliant and at the same time crazy people in a Marie Skłodowska-Curie program. You have made each of our international meetings an unforgettable adventure. Besides, this work would have never been possible without my external collaborators at IBM and at the German Centre for Cardiovascular Research. I am very grateful to Andreas and Slava for their active engagement despite their busy agendas and all the Covid-19 disruptions.

This Thesis is the culmen of an intense academic journey. Thus, I would like to thank all the lecturers that ever shared their knowledge and supported me along this way. Raul, thanks a lot for encouraging me to start this PhD. Brothers Mato, Juan, and all the 'IQ' crew, thanks for such wonderful years and the sprit that you conveyed to me. Also, special credits to Sonja, my friend and the best English teacher in the world. Finally, I would like to acknowledge all the institutions and funding entities that have made this possible. Thanks to my home university, the University of Valladolid, for always supporting me.

The PhD has been a roller coaster ride, with peaks of extreme joy and valleys of infinite frustration. I would like to deeply thank my family and friends for helping me stay sane and fit during this journey.

Arturo, Kike, Jeni, Sonia, the ‘Pulpito Paul’ crew and all of you, thanks for always being there and up for plans whenever I could take a break. Thanks to the people from my lovely village, Villacón, for all their kindness and to the ‘Costanas’ crew & company for being so easily motivated to make me a visit anywhere. Miguel, take what you can, give nothing back. Luca, 小老弟, 谢谢你一直相信.

My dear family, thanks, thanks, and thanks. ¡Tengo las mejores tías, primos y sobrinitos del mundo! Maya and Leo, my two lovely goddaughters, I feel extremely privileged to be your godfather. Mamá, gracias por todo tu apoyo, tu dedicación, tu cariño y por recordarme las cosas que son importantes de verdad. Nunca te rindas. Papa, allá donde estés, me guías, iluminas mi camino y cuidas de mí. Gracias. I would like to conclude by thanking the person that has supported me the most during these 3 last years. Nieves, I could not be more grateful for your patience, love, and care. Thanks also to your family for their help during the Covid-19 crisis. We started and finished this Thesis together. More adventures are yet to come.

Abstract

Improving healthcare systems mandates a shift towards personalised and preventive medicine, especially in the management of cardiovascular diseases, the world leading cause of death. The anatomy and function of the heart can be accurately and non-invasively assessed using cardiac magnetic resonance. Nevertheless, it is not yet fully understood how their interplay modulates disease outcomes. In this regard, this Thesis (1) reviews the crucial role of computational modelling towards precision cardiology, by building the 'digital twin' of a patient; (2) develops a fully automated pipeline to construct 3D computational models of human left ventricle from 2D cardiac magnetic images, using state-of-art deep learning approaches; and (3) proposes an enhanced structural and functional analysis to disentangle how adverse outcomes are modulated in cardiovascular disease. This is showcased in acute myocardial infarction, the deadliest cardiovascular disease not only on first manifestation, but on reoccurrence and secondary adverse outputs. In this scenario, the approach introduced in this Thesis is shown to provide additional prognostic information towards a better post-infarction management. Moreover, it identifies 3D shape and contraction patterns that are related to adverse outcomes, building a more profound understanding of remodelling after infarction. Furthermore, the automation of the analysis removes the burden of manual annotations and intra- and inter-observer variability, as confounders of prognosis. All this argues for the clinical impact of this Thesis. Ultimately, the Thesis evidences the potential of 3D computational models for cardiovascular disease management, towards personalised and preventive care.

List of Contents

List of Figures **8**

List of Abbreviations **9**

1. Introduction **12**

1.1. Motivation 12

1.2. Research Aims and Contributions 13

1.3. Thesis Overview 16

1.4. Research Publications, Presentations and Awards 18

1.4.1. List of Publications 18

1.4.2. Presentations in Conferences and Workshops 19

1.4.3. Public Engagement and Media Interviews 19

1.4.4. Awards 20

2. Background and Literature Review **23**

2.1. Cardiac Computational Modelling 23

2.1.1. Statistical Modelling 23

2.1.2. Mechanistic Modelling 27

2.2. Clinical Background 30

2.2.1. Basic Cardiac Anatomy and Function 30

2.2.2. Cardiovascular Diseases 34

2.2.3. Hypertrophic Cardiomyopathy 34

2.2.4. Acute Myocardial Infarction 36

2.3. Cardiac Magnetic Resonance 38

2.3.1. CMR in Imaging Diagnosis 38

2.3.2. CMR Principles 40

2.3.3. CMR Analysis in Clinical Practice 42

2.3.4. CMR and ML in Acute Myocardial Infarction 44

2.4. Automated 3D CMR analysis 46

2.4.1. Traditional Methods for Automated CMR Segmentation 46

2.4.2. ML Methods for Automated CMR Segmentation 47

2.4.3. 3D Shape Reconstruction 53

2.4.4. 3D Shape Analysis 57

3. Mechanistic and Statistical Modelling Synergy towards Precision Cardiology **61**

3.1. Chapter Publications 61

3.2. Chapter Outline 61

Paper I 63

4. Accurate and Fully Automated Estimation of the 3D Heart Shape from 2D CMR 78
4.1. Chapter Publications 78
4.2. Chapter Outline 78
Paper II 82
Paper III 93
Paper IV 106
5. Analysis of the Heart Anatomy from 3D Meshes 122
5.1. Chapter Publications 122
5.2. Chapter Outline 122
Paper V 125
6. Enhanced Cardiac Function Characterization Using CMR 144
6.1. Chapter publications 144
6.2. Chapter Outline 144
Paper VI 147
7. Discussion and Conclusions 168
7.1. Key Contributions and Main Findings 168
7.2. Future Work 170
7.3. Conclusion 172
REFERENCES 175
APPENDICES 202
Appendix A: Supplemental Material Paper V 203
Appendix B: Supplemental Material Paper VI 229

List of Figures

1.1 Thesis goal and research objectives 13

2.1 Statistical modelling techniques 26

2.2 Mechanistic modelling 28

2.3 Basic cardiac anatomy, function and electrical activation 31

2.4 Left Ventricular volume transient and cardiac phases 32

2.5 ECG diagram of a healthy heart in normal sinus rhythm 33

2.6 Hypertrophic cardiomyopathy and acute myocardial infarction 35

2.7 Normal vs ST-elevated ECG 37

2.8 Cardiac magnetic resonance standard acquisition planes 42

2.9 U-Net architecture 49

List of Abbreviations

2D, 3D	Two-, Three-dimensional
AE	Autoencoders
AMI	Acute myocardial infarction
AUC	Area under the ROC curve
CMR	Cardiac magnetic resonance
CMR-FT	Cardiac magnetic resonance feature tracking
CNN	Convolutional neural network
CVD	Cardiovascular diseases
DL	Deep Learning
ECG	Electrocardiography
ED	End-diastole
ES	End-systole
ESV	End-systolic volume
GAN	Generative adversarial networks
HCM	Hypertrophic cardiomyopathy
LV	Left ventricle
LVEF	Left ventricular ejection fraction
LVH	Left ventricle hypertrophy
MACE	Major adverse cardiac events
ML	Machine Learning
MRI	Magnetic resonance imaging
NN	Neural Network
NSTEMI	Non-ST-segment–elevation myocardial infarction
PCA	Principal Component Analysis
ROI	Region of interest
RV	Right ventricle
SAx	Short-axis
SMOD	Statistical models of deformation
STEMI	ST-segment–elevation myocardial infarction
SV	Stroke volume
VAE	Variational autoencoders

CHAPTER 1

Introduction

1. Introduction

1.1. Motivation

In the current socio-political framework, nations are facing tightening financial constraints and an aging population. Given this scenario, preventive and personalised medicine is crucial to improve both the efficiency and efficacy of healthcare systems [1], [2]. This is particularly important in cardiovascular disease (CVD), the world leading cause of death, responsible for 42% of the European mortality and costing 169€ billion/year only in Europe [3]–[6]. CVD is usually associated with changes in the structure and function of the heart. Nevertheless, their interplay in modulating disease outcomes is not fully understood, resulting in suboptimal treatments. My personal research interest is therefore to gain a better understanding of cardiac structural and functional remodelling, to contribute towards early-stage disease identification and personalised medicine. This remodelling has been proven to impact survival after acute myocardial infarction (AMI), a disease that accounts for millions of deaths worldwide per year [7]–[12]. The burden of AMI and the limited understanding of the way remodelling patterns affect survival is a direct motivation for this research.

Large population studies have been facilitated by the recent expansion of big data in cardiovascular medicine [13], [14]. Likewise, increasing computational power has boosted cardiac modelling, enabling the concept of the 'digital twin' of a patient towards precision cardiology [2]. Facilitated by this large dataset and computational power availability, recent developments in deep learning (DL) are revolutionising the medical imaging field [15], [16]. They remove the burden of manual analysis and expand the possibilities of accurate and automated analysis of cardiac images, including cardiac magnetic resonance (CMR) datasets, the gold standard for quantitative assessment of the heart [15]–[19]. Altogether, these advances facilitate the integration of anatomy with function in 3D computational models to further disentangle the anatomical features and contraction patterns that modulate CVD outcomes. The final motivation of this Thesis is the translation to clinical practice, by addressing methodological challenges that allow building knowledge and ultimately improving

patient management. A scenario of breakthroughs emerges, and I am excited to have become part of it.

1.2. Research Aims and Contributions

The goal of this Thesis is to gain a better understanding of how cardiac anatomy and function modulate adverse outcomes in human CVD. This is structured into three research objectives, as depicted in Fig. 1.1:

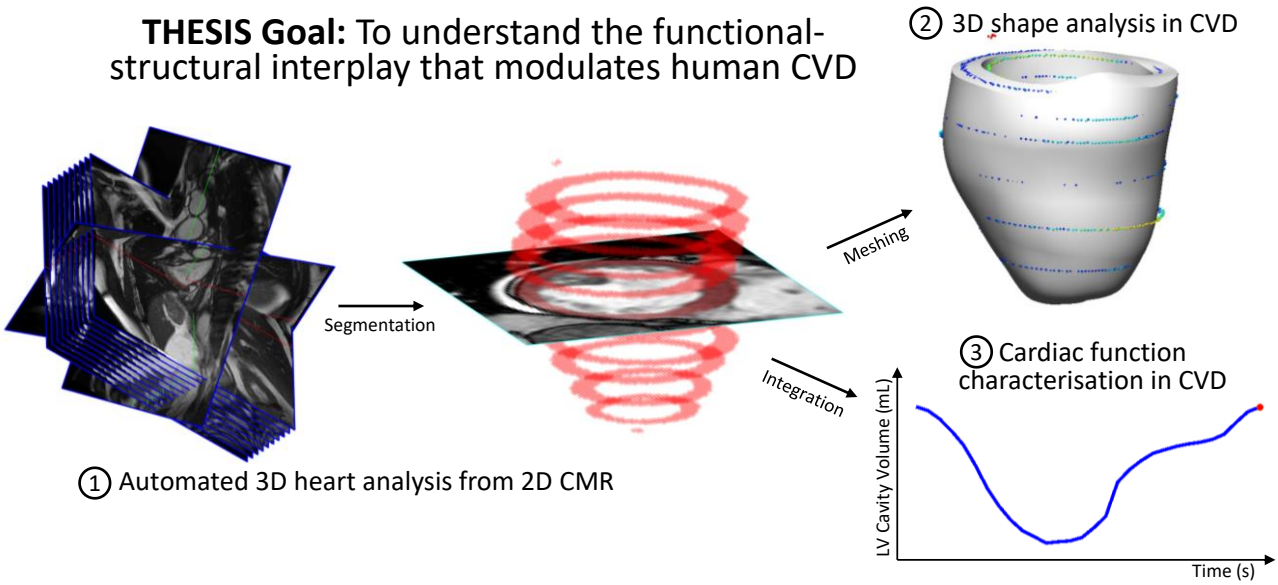


Fig. 1.1: Outline of the Thesis goal and its 3 core research objectives. **LEFT**: CMR scan sample, including all the standard 2D planes. **MIDDLE**: Segmentation of the 2D images aligned with the short axis plane of the left ventricle (LV), arranged in 3D. **RIGHT**: Resulting 3D LV mesh in an instance of the cardiac cycle (top), and LV volume transient (bottom).

The first objective is the development of a fully automated pipeline to accurately estimate the 3D shape of the heart from 2D CMR data. It focuses on methodological aspects to overcome the challenges of the 2D CMR segmentation and its translation into 3D meshes. Most of the contributions within this first objective are on the field of segmentation, where state-of-the-art DL techniques are extended to account for: (1) the limited training data availability in clinical environments; (2) the variability in structure and appearance of the anatomies to be segmented; and (3) the segmentation labels imbalance. Contributions within this objective include:

- Development of a data augmentation technique based on statistical models of deformation (SMOD), to enrich the space of structural variability of the heart by introducing anatomically meaningful deformations.
- Validation of SMOD showing a significant improvement in segmentation accuracy, especially in scenarios where the training data available is very limited or the structures to segment are complex and highly variable (hypertrophic cardiomyopathy case study).
- Development of a fully automated pipeline for left ventricular (LV) quantification from 2D CRM scans. The proposed pipeline is based on a novel 2-step DL architecture, able to tackle segmentation labels imbalance. The pipeline is trained with SMOD and aligns the LV orientation to a canonical reference position for metrics quantification. The development includes a complete assessment of the overall impact of the SMOD hyperparameters and the neural network (NN) architecture on the pipeline performance.
- Coupling of the developed 2-step DL segmentation pipeline with domain adaptation and domain unlearning techniques. This overcomes challenges associated with inter-domain variability in CMR appearance, enabling the use of the pipeline for CMR multicentre and multi-vendor applications.
- Development of a 3D fully automated pipeline to estimate the 3D shape and contraction of the human LV. This is done by coupling the 2D CMR segmentation NNs with a mesh personalisation algorithm based on cubic Hermite variational mesh warping.

The second objective builds on and expands the contributions from the first one. It consists of the 3D analysis of heart anatomy in CVD, specifically in AMI patients. Shape analysis and dimensionality reduction techniques are coupled with multivariate predictive models towards the following contributions:

- Comprehensive analysis of 3D LV patterns of shape and contraction on AMI patients, within days after the infarct event, in a large-sized multicentre CMR study for risk stratification.

- Building of the first 3D reference atlas of the post-infarcted human LV, which captures the average LV anatomy from 1000+ AMI patients along with its main variations (PCA modes).
- Identification of the 3D patterns of the LV related to AMI prognosis, and studied their interplay with other clinical parameters, such as infarct size and micro-vascular obstruction.
- Assessment of the performance of alternative 3D descriptions of end-systolic volume (ESV) and LV ejection fraction (LVEF), as main clinical biomarkers for AMI risk stratification.
- Investigation of the gains in performance of the proposed 3D metrics for the prognosis of adverse effects after AMI, when embedded into a multivariable setting including CMR markers, cardiovascular risk factors, and basic patient characteristics.

The third research objective expands my investigations on cardiac function. LV contraction after AMI is characterised in further detail via the analysis of LV volume transients and myocardial strains. Strain analysis is enabled by exploiting CMR feature tracking (CMR-FT) as an additional imaging modality. Altogether, the shape analysis from the second research objective incorporates 3D detail, while LV volume transients and CMR-FT strains provide temporal dynamics and bulk tissue level information, respectively. These 3 directions to enhance the description of contraction are combined to further improve risk prediction in the AMI case study, resulting in the following contributions:

- Extension of the project pipeline to automatically derive the personalised LV volume transient and calculate its systolic and diastolic relevant clinical metrics from 2D CMR scans.
- Characterisation of LV contraction after AMI via the 3 proposed directions of enhancement in a large-sized multicentre CMR study.
- Identification of LV contraction patterns related to AMI adverse events via machine learning (ML) techniques. As an additional result of the ML analysis, I have built a LV volume transient atlas for AMI patients, which describes the post-infarct average contraction along with its main modes of variation.

- Evaluation of the prognostic value of each of the three proposed directions to enhance the characterisation of cardiac function.
- Combining the proposed directions of enhancement, on top of established CMR indexes and cardiovascular risk factors, towards a final prediction model of major adverse cardiac effects after AMI.

Taken together, this Thesis demonstrates the potential of computational models to improve CVD management. To further accelerate this vision and facilitate future studies, the final translational contributions of the Thesis include:

- Software development to make publicly available to the community the pipeline, atlases, and risk stratification models derived during my Thesis.
- Writing of a position paper, published in European Heart Journal, highlighting the role of the synergistic combination of statistical and mechanistic models towards precision and preventive cardiovascular medicine [2].

1.3. Thesis Overview

This dissertation is written as an integrated Thesis, where original contribution chapters are substituted by scientific papers and complemented by a position paper. For those chapters, a brief introduction is provided explaining the way in which the papers contribute to the global objectives of the Thesis. A brief outline of the chapters is given below:

Chapter 1 introduces the Thesis and its motivation. This chapter also summarises the research objectives and contributions, presents the structure of the Thesis and lists the accomplishments achieved towards its completion.

Chapter 2 aims at introducing the basic principles of cardiovascular disease, cardiac modelling, CMR, and deep learning. It illustrates the state-of-the-art trends and place the pieces of the Thesis in

context, to help the reader understand the implications and contributions of the work. The literature review is further extended in the introduction of each of the scientific papers. Therefore, a certain degree of overlap is expected given the integrated thesis format.

Chapter 3 elaborates on the role of mechanistic and statistical models towards personalised and preventive medicine, and emphasises their combined use to accelerate clinical translation. Such a vision guides this Thesis to maximise its clinical implications, and it is adopted when developing and selecting the methods. This chapter includes 1 scientific publication (Paper I).

Chapter 4 corresponds to the first objective of the Thesis, i.e., the development of a fully automated pipeline to obtain the 3D shape of the heart from 2D CMR scans. It discusses the main challenges and how these have been overcome towards the completion of the pipeline. This chapter resulted in 3 publications (Papers II-IV).

Chapter 5 concentrates on the 3D analysis of heart anatomy in CVD, the second objective of the Thesis. It identifies 3D LV patterns related to risk of adverse effects after AMI, and proves they provide additional prognostic information for clinical management. This chapter contributed to 1 scientific publication (Paper V).

Chapter 6 further focuses on the quantitative analysis of cardiac function. LV contraction is comprehensively characterised using CMR-FT strains, volume transients, and shape analysis. This expands the AMI case study presented in the previous chapter, aiming at gaining a better understanding of the structural and functional interplay that modulates adverse post-infarction events. The chapter led to 1 additional publication (Paper VI).

Chapter 7 discusses the results and the transversal implications of the three research objectives chapters. It analyses the Thesis as a whole and summarises its major scientific contributions and findings, along with the potential future directions for further research. Finally, it closes the Thesis with the concluding remarks and the take-home message.

1.4. Research Publications, Presentations and Awards

1.4.1. List of Publications

- **J. Corral Acero**, F. Margara, M. Marciniak, et al. "The 'Digital Twin' to enable the vision of precision cardiology". *Eur. Heart J.*, Mar. 2020, 41:4556–64. This publication corresponds to Paper I, included in Chapter 3.

- **J. Corral Acero**, E. Zacur, H. Xu, et al. "SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI". *FIMH 2019*. Lecture Notes in Computer Science, vol 11504. Springer, Cham; 2019. p361-69. This publication corresponds to Paper II, included in Chapter 4.

- **J. Corral Acero**, H. Xu, E. Zacur, et al. "Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation". *STACOM 2019*. Lecture Notes in Computer Science, vol 12009. Springer, Cham. 2019. p384–94. This publication corresponds to Paper III, included in Chapter 4.

- **J. Corral Acero**, V. Sundaresan, N. Dinsdale, et al. "A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation". *STACOM 2020*. Lecture Notes in Computer Science. Springer, Cham; 2021. p. 196–207. This publication corresponds to Paper IV, included in Chapter 4.

- **J. Corral Acero**, A. Schuster, E. Zacur, et al. "Understanding and Improving Risk Assessment after Myocardial Infarction using Automated Left Ventricular Shape Analysis". Under 3rd revision in *JACC Cardiovasc Imaging*. This publication corresponds to Paper V, included in Chapter 5.

- **J. Corral Acero**, P. Lamata, I. Eitel, et al. "Comprehensive Characterization of Cardiac Function for Improved Post-Infarction Risk Assessment". Under 1st revision in *Circulation.* This publication corresponds to Paper VI, included in Chapter 6.

- V.M. Campello, P. Gkontra, C. Izquierdo, C. Martín-Isla, A. Sojoudi, P.M. Full, K. Maier-Hein, Y. Zhang, Z. He, J. Ma, M. Parreno, A. Albiol, F. Kong, S.C. Shadden, **J. Corral Acero**, et al. "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge". *IEEE Trans Med Imaging*. 2021 Jun 17; PP.

- J. Parikh, T. Rumbell, X. Butova, T. Myachina, **J. Corral Acero**, et al. "Generative Adversarial Networks for construction of virtual populations of mechanistic models: Simulations to study Omecamtiv Mecarbil action". *J Pharmacokinet Pharmacodyn*. 2021 October 29; PP.

- M. Beetz, **J. Corral Acero**, and V. Grau, "A Multi-View Cross-Over Attention U-net Cascade With Fourier Domain Adaptation For Multi-Domain Cardiac MRI Segmentation". Accepted for publication in *M&Ms2 - STACOM 2021*. Lecture Notes in Computer Science. Springer, Cham.

1.4.2. Presentations in Conferences and Workshops

- Personalised in-silico Cardiology Winter School, Oxford (November 2020)
- MICCAI 2020 conference, STACOM workshop, Lima (October 2020)
- Efficient use of animal and virtual models in evaluation, Maastricht (November 2019).
- MICCAI 2019 conference, STACOM workshop, Shenzhen (October 2019).
- FIMH 2019 Conference, Bordeaux (June 2019).
- VPH Summer School, Barcelona (June 2019).
- Department of Engineering Science Researchers Conference, Oxford (March 2019).
- Provision of diagnostic and prognostic value by personalisation, Canfranc (March 2019).
- Statistical & mechanistic modelling: mutual synergies in cardiology, Bordeaux (November 2018).
- Medical Imaging Summer School 2018, Sicily (July 2018).
- SIMULA Summer School in Computational Physiology, Oslo (June 2018).
- Personalised in-silico Cardiology European Network Kick-off event 2018, London (March 2018).

1.4.3. Public Engagement and Media Interviews

- TEDx Youth@EEB3 (https://youtu.be/mHu7xdq_5oo).
- Academia and Industry debate moderator at MICCAI 2020.
- Diario de Valladolid (http://www.diariodevalladolid.es/amp/noticias/innovadores/ingeniero-llega-corazon_133248.html).

- University of Oxford (https://www.eng.ox.ac.uk/news/best-paper-award-is-huge-boost-for-cardiac-imaging-researchers/).

1.4.4. Awards

- Best paper in Medical Imaging, Functional Imaging and Modelling of the Heart 2019.
- First prize, Left Ventricle Full Quantification Challenge MICCAI 2019.

CHAPTER 2

Background and Literature Review

2. Background and Literature Review

2.1. Cardiac Computational Modelling

The current tightening financial constraints and aging population that nations are facing urge to improve healthcare systems, mandating a shift towards a personalised and preventive management of disease [1], [2]. This effectively means to consider phenotypic variability in disease to a greater extent, and transition from one-size-fits-all treatments to therapies tailored to each patient [20]. The recent explosion of computer modelling, facilitated by increasing computational power, has proven crucial in this direction [21], [22]. Computer models are able to integrate physiological knowledge and clinical data into the 'digital twin' of a patient, i.e. their personalised *in silico* representation towards optimal therapy and preventive medicine [2]. This integration can occur at any level, from genetic and molecular data, up to the cell, tissue, and organ pathophysiology, lifestyle, and environment of the patient [23]. Computer models are built by induction, learning from data, and/or by deduction, through mechanistic modelling and simulation, integrating multiscale knowledge and data [14], [24]. The former are called statistical models, and the latter, mechanistic models. While both types of models are introduced next, more information can be found on [13], [14], [25] and [23], [24], respectively. Paper I, which constitutes the core of Chapter 3, further elaborates on this topic [2].

2.1.1. Statistical Modelling

Statistical models are built on historical data and capture the evidence of past observations, encapsulating the knowledge and relations induced from the data [2], [25]. They often lack the possibility of explaining causality, and are thus sometimes referred to as a 'black-box' tool. However, they can be extremely useful in extracting key dataset features that would otherwise remain unrevealed [2]. This type of modelling is central to the increasingly popular ML techniques [25]. This is especially relevant in medical imaging, a vanguard for application of ML in clinical practice, to

guide and optimise the acquisition [26]–[28], automatically analyse the image [29]–[34], or even perform diagnosis [25], [35]–[38].

Statistical modelling can be either supervised, where the models are enforced to look for patterns that fit an outcome of interest, or unsupervised, where pattern recognition freely develops within the supplied data without any prior specification of outcome [14], [25], [39]. Examples of supervised training range from image segmentation [40]–[44] or the direct estimation of measurements [31], [45], to classification [46], [47], automated diagnosis and patient management [2], [25], [35], [48]–[52], among others. Unsupervised models are applied to combine different sources of data towards disease phenotyping and to gain understanding of the similarities between patients [53], [54], or to pre-process data for other algorithms [25], for instance. Indeed, both supervised and unsupervised approaches can be used in combination, i.e., unsupervised learning can be applied to patients with known outcomes to identify novel features, which could then be fed into supervised models [25], [55], [56]. Likewise, several models trained for the same task can be used in parallel (i.e., bagging) or in sequence (i.e., boosting) to create stronger meta-models called ensembles, which reduce variance and bias in the predictions [57], [58].

Regression analysis, one of the simplest supervised techniques, has been widely used in medicine due to its relatively straightforward interpretably [39]. It consists of an algorithm that describes a relationship (linear, logistic, etc.) between one or more variables and the outcome of interest. In this type of modelling, variables may be automatically added or removed based on the strength of their association with the outcome, until a significant model is learned (stepwise analysis) [25]. Support vector machines and random forests can capture highly nonlinear models and more complex variable-outcome relationships [59], [60]. While the former defines planes of separation in a high dimensional space, the latter identifies the threshold decision values to separate out features and predict a certain outcome [25].

Neural network models have been lately established as the state-of-the-art technique for complex supervised tasks [14], [48], [61]. “Deep” networks, i.e., networks with many layers, have proven particularly effective, giving rise to the field known as Deep Learning. Neural networks consist of a collection of connected neurons. Inspired by the activation in the human brain, these neurons usually have multiple inputs and one output, with the output generally representing the active or inactive states of biological neurons [61], [62]. This is mathematically formulated by: (1) multiplying (dot product) the input vector of features times a weight vector; (2) adding a bias; and (3) finally applying a (generally nonlinear) activation function, to output the state [62], [63]. Values for the network’s weights (including bias terms) are learnt in order to minimise a loss function, selected to represent the performance on the objective task. The loss function gradient determines the direction in which weights are updated, usually in a process known as backpropagation. In backpropagation this gradient flows back through the network to update all its parameters applying a certain learning rate [64], [65]. While in fully connected NNs all neurons of subsequent network layers are connected, convolutional NNs (CNN) rely on locally connected kernels as feature detectors that are convolved over the respective layer inputs [66], [67]. The local and global features are often integrated in CNNs via pooling operations that downsample or upsample the layer resolution [40]. Fig. 2.1 illustrates some examples of the mentioned techniques in the context of classification. As indicated above, these techniques can be used in a wide range of applications, including the regression of continuous outputs (outputs are not necessarily binary or discrete, as in classification).

Clustering analysis is extensively applied in unsupervised learning to identify consistent patterns in the data [14], [68]. This comprises a vast variety of algorithms, from hierarchical and partitioning algorithms to model-, grid- or density-based clustering. These algorithms are able to structure the features based on distance, similarity, or distribution, among others [25], [35]. Unsupervised learning algorithms are many times combined with dimensionality reduction techniques, which simplify data samples by representing them with a reduced number of features. In this Thesis I use Principal

Component Analysis (PCA), which selects the features that account for the most variation in a dataset [25], [68]. These PCA features correspond to the directions of orthogonal projection of the data onto a lower dimensional linear space that maximise the variance. PCA is particularly popular in geometrical modelling, to condense complex meshes, and their node features, into a reduced number of parameters [52], [55], [68]–[71]. Appendices A and B illustrate this in detail in the context of LV shape and volume transient analyses.

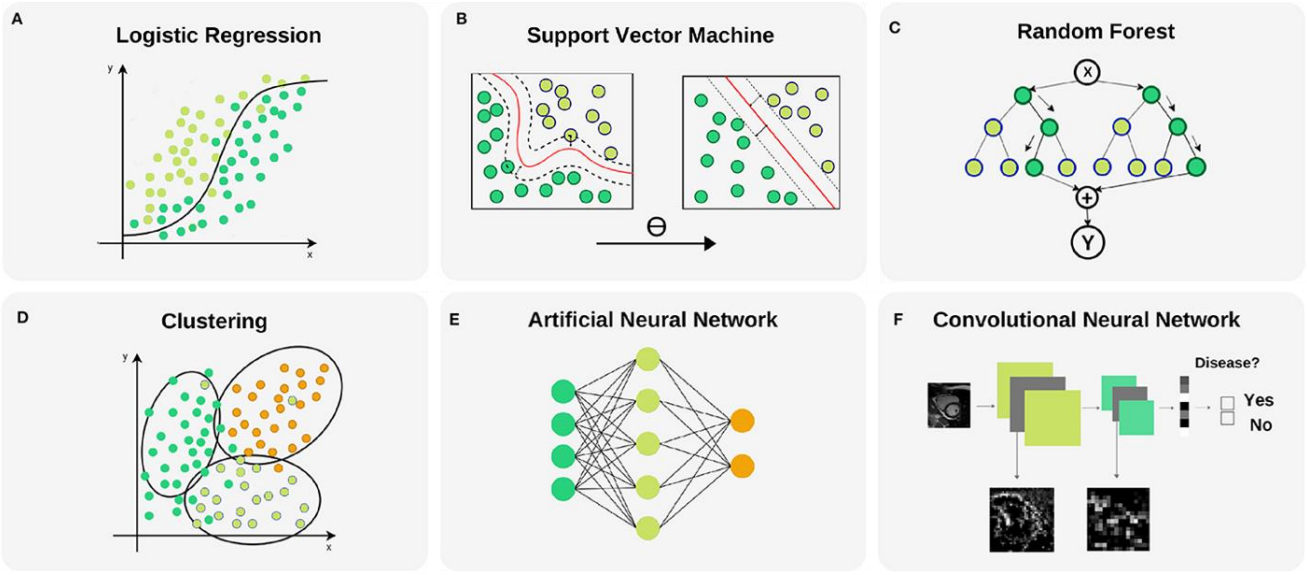


Fig. 2.1: Examples of statistical models applied to classification (discrete output). **A**: Logistic Regression models the probability of a binary outcome, y axis, based on the continuous input variable, x axis. **B**: Support Vector Machine model transforms via kernel operations a non-linear boundary to a linear one maximising the distances between classes. **C**: Random Forest combines several Decision Trees to reduce prediction uncertainty. **C**: Clustering finds subgroups along data in an unsupervised manner. **E**: Artificial NNs model complex non-linear relations between input variables and outcomes by connecting layers of neurons. **F**: CNNs use kernels as feature detectors towards improved neurons connectivity and are recently widely used in medical imaging. Figure adapted under Open Access license from [35].

Regardless of the statistical modelling approach deployed, the quality, accuracy, and richness of the data ultimately determines the usefulness of the model [13], [25], [48]. In order to build successful applications, its mandatory that: (1) the data is relevant and detailed enough to answer the question; (2) the modelling technique is appropriate to the type, complexity, and amount of data available; and (3) the model is rigorously validated [14], [25]. NN models require particularly large amounts of data to generalise properly. Only recently have they been facilitated by the expansion of big data in cardiovascular medicine, through the creation of biobanks and bioresources [13], [14], [35], [72].

Nevertheless, confidentiality issues, technical, ethical, and financial constraints still limit data availability, and, in consequence, the use of these approaches in clinical practice [13], [14]. To overcome such limitation and avoid overfitting, data augmentation and transfer learning techniques are routinely applied [61], [73], [74]. Data augmentation enlarges the dataset by introducing modifications or generating synthetic instances from the original ones. In transfer learning, a model developed for a task is reused as the starting point for training on a second task.

Model validation relies on a full separation of the data into training and testing datasets. This can be ensured by keeping an independent set out for testing. Alternatively, the entire dataset can be divided into ‘k’ roughly equal parts (folds) and iteratively train on ‘k-1’ folds and test on the remaining unseen fold [25]. This is called k-fold cross-validation, and has proven to have smaller bias than a traditional training-testing split approach [25], [75]. However, even when a rigorous cross-validation is applied, a further multicentre external validation is recommended before clinical deployment. For this type of applications, homogenisation of clinical data recordings and imaging protocols is essential [25]. Finally, approaches to improve the interpretability, credibility and understanding of ML models, and especially the deep learning ones, have been explored [76]–[80]. This mainly includes modifications in the NN architecture to interpret the features along the whole NN, and to illustrate the NN output reasoning based on the input data. These modifications usually come at the cost of performance in the supervised task.

2.1.2. Mechanistic Modelling

Mechanistic models rely on the bio-physical knowledge of the system of interest, which is expressed through sets of mathematical equations describing the casual mechanisms behind the observed phenomena [81], [82]. They provide a framework to integrate and augment clinical and experimental data, towards the identification of mechanisms and the prediction of outcomes [2]. Since the first model of a ventricular cardiomyocyte was published over five decades ago, cardiac modelling has

pushed the boundaries of our understanding of cardiac physiology and pathophysiology. It does so by providing useful insight from cell up to whole organ [83]–[85]. Phenomena such as electrical activation, heart contraction, cavity pressure, and blood flow can be accurately simulated [81]. Besides, *in-silico* models can be used not only as knowledge building tools, but also to assist evidence-based decision making both in clinical and regulatory settings (see Fig 2.2).

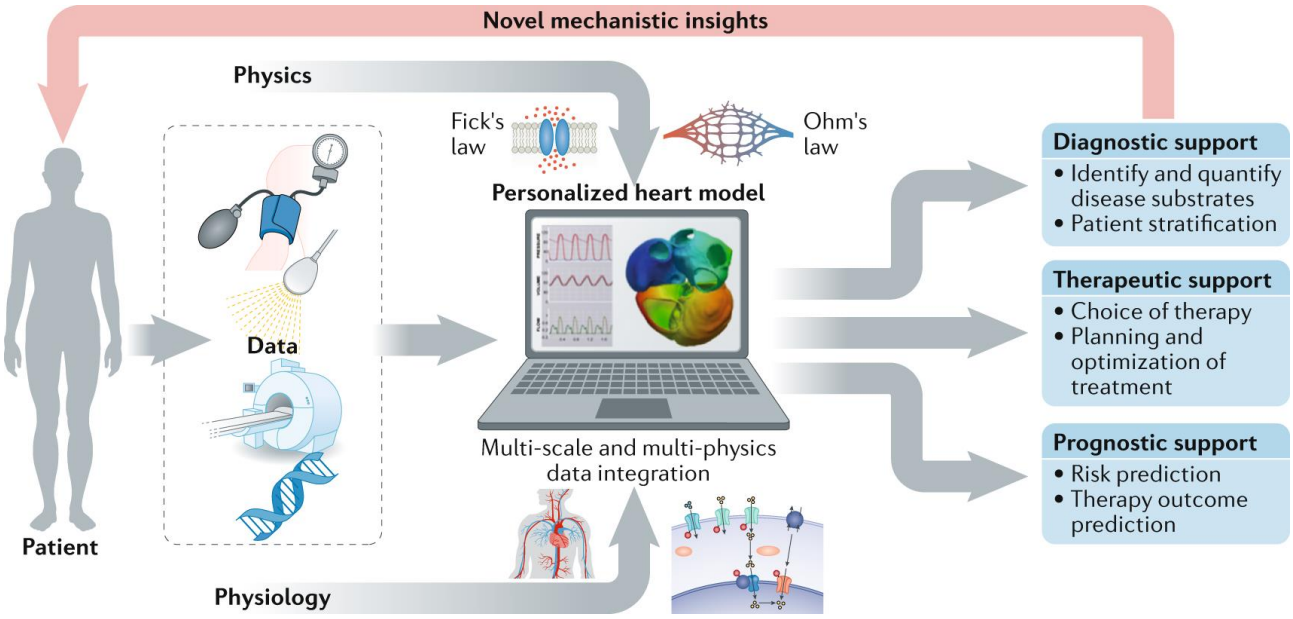


Fig. 2.2: Computational mechanistic models are able to integrate diagnostic data from different and heterogenous modalities and combine it with the physiological knowledge and physics principles towards patient-specific personalised models. This integration can occur at multi-scale level. The models allow for diagnostic and therapy-decision improvement and, in the end, an enhanced cardiology care. Figure reproduced under permission from [24].

The combination of cardiac modelling with recent advances in medical imaging has facilitated the evolution from generic to personalised 3D heart models [20], [86]. These models faithfully represent the anatomy and function of a given subject [22], [87], [88]. Hence, simulation-based personalised medicine may predict endpoints, or indicate whether a patient is an appropriate candidate for a medical device or pharmaceutical treatment. This is achieved, for instance, by simulating invasive clinical procedures or dosage effects to predict unexpected outcomes before selecting a therapy [89], [90]. In combination with imaging, mechanistic modelling may also produce surrogate measurements for biomarkers that cannot be directly measured non-invasively [91]–[94]. Personalised models have proven useful in a variety of clinical scenarios from myocardial infarction, heart failure, cardiac

resynchronisation and ablation, to hypertrophic cardiomyopathy [53], [81], [82], [95]–[101]. In addition, the recent clearance from the U.S.A. Food and Drug Administration to biophysical model-based products, such as HeartFlow® and CardioInsight®, further supports the use of personalised modelling in clinical practice [81], [102], [103]. *In-silico* modelling is well-established in drug design [82]. It has the potential to simulate safety of therapy prior to first-in-human use, or under scenarios not ethically possible in human [104]. Thus, clinical trials may be augmented accounting for virtual patients that represent different phenotypes; animal use, reduced; and gaps in the transition from *in-vitro* to *in-vivo*, filled [104]–[107].

Unlike statistical models, cardiac mechanistic modelling has the advantage of explaining causality. Moreover, *in-silico* models can be exploited in scenarios of highly limited data, where statistical approaches suffer the most [13]. Should the prior knowledge be adequate, they are predictive under unseen scenarios without the need of being re-trained for specific purposes [82]. Additionally, in the context of generating knowledge and hypothesis testing, they are extremely important in pointing towards the next most valuable piece of data to be collected [108]. Furthermore, *in-silico* models have also proven useful to test scenarios that cannot be evaluated in real patients as well as to reliably infer data that cannot be directly measured [109], [110]. They can also be used to generate phantom data to train statistical models [111], [112].

However, as our understanding of physiology expands, the complexity and non-linearity of the models increase accordingly [82], [113]. Thus, larger parameter sets are needed for model parametrisation, and more uncertainty is introduced [114], [115]. Sensitivity analysis to assess the effect of this uncertainty on model outputs is a common practice, involving large number of computationally heavy runs [113]–[116]. In addition, either optimisation or the solution of inverse problems for model personalization may also require a high number of simulations. Despite the recent increase in computer power, this remains a challenge [113], [114], [116].

In these scenarios, where high complexity or a large number of runs are needed, the use of statistical frameworks as emulators or meta-models can be exploited to alleviate the computational cost of mechanistic models. This enables the feasibility of wider parameter explorations and longer runs (e.g., number of beats) [113]–[115]. Effectively, emulators treat the mechanistic *in-silico* model as a black box [116]. Given a certain input, they regress the output based on historical data (training data) generated using the mechanistic model [115]. Several emulation approaches, such as polynomial chaos expansion or partial least squares linear regression, have been proposed [116], [117]. Gaussian processes, which regress outputs along with predicting uncertainty, are most used since first introduced in the field of cardiac modelling in 1989 [113], [114], [118], [119]. Gaussian processes benefit from handling input uncertainty explicitly. However, they assume normal distribution in both inputs and outputs, as well as the outputs to be smooth and continuous functions of the inputs, which is not always the case [113]. Emulators based on recent DL techniques have proven successful to solve high non-linearities, and may be alternatively considered [120]–[123]. Such DL frameworks can be trained for multiple outputs simultaneously and their excellent performance in weather forecast or climate modelling can serve as reference [121], [124]–[127]. Finally, to account for heart geometry and combine this information with inputs of other natures, approaches such as that in [128] may be followed. In that work, the authors condensed the geometrical information into a small number of continuous variables using PCA, which were subsequently used for model optimization to estimate load-free geometries.

2.2. Clinical Background

2.2.1. Basic Cardiac Anatomy and Function

The heart is a muscular organ responsible for uninterruptedly pumping blood between the lungs and the rest of the body. It is at the centre of the cardiovascular system, which ensures a continuous flow of oxygen, nutrients, and hormones to the organism. The top of the heart, where the valves are located,

is called the base, and the bottom, the apex. The heart consists of 4 chambers: the left and right atria, which behave as reservoirs; and the left and right ventricles (LV and RV), which act as intermittent pumps [129]. Fig 2.3.A depicts the anatomy and blood flow in a healthy heart.

The right atrium receives de-oxygenated blood from the body through the superior and inferior vena cava, and the RV pumps it through the pulmonary artery to the lungs, where it is oxygenated. The fresh blood then flows via the pulmonary vein to the left atrium, and is propelled by the LV back to the rest of the body through the aorta. This completes a cardiac cycle of pressure and volume changes [129]. Fig. 2.4 illustrates the evolution of the LV volume along the cycle, and summarises its different phases. The phase where the ventricles fill up, firstly by difference of pressures (passive filling) and secondly by active contraction of the atria (active filling), is called diastole. The phase where the ventricles contract ejecting blood is known as systole.

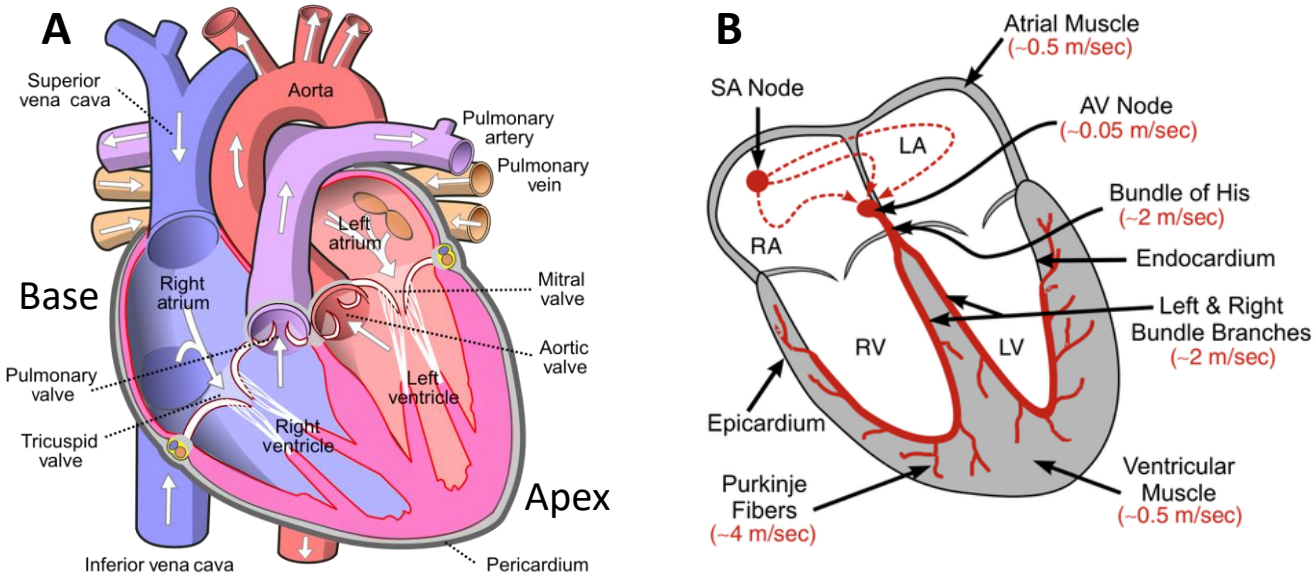


Fig. 2.3: **A**: Basic anatomy and cardiovascular circulation system of a healthy heart. Adapted under Open Access license from Wikipedia.org [130]. **B:** Cardiac electrical activation of a healthy heart. The electrical signal originates in the sinoatrial (SA) node, propagates to the atrioventricular (AV) node, and then, to the bundle of His and the Purkinje system. The propagation speed in indicated in parenthesis. Reproduced from [131], under Open Access license for non-commercial use.

The septum is the muscular wall that separates the right and left sides of the heart, preventing the blood from mixing. The direction of blood flow into the RV and LV, and out into the lungs and body,

is controlled by the tricuspid, mitral, pulmonary, and aortic valves, respectively. The heart itself receives blood and nutrients from the coronary arteries. These are a tree-like vessel structure located on top of the RV and LV that branches from the aorta. The RV shape is pyramidal, and the LV resembles a prolate ellipsoid, with thicker walls to face the higher systemic cardiovascular resistance of supplying the whole body. The ventricular and septum walls are composed of three layers. The epicardium is the outer and protective thin layer, composed of elastic and adipose tissue. The myocardium is the middle and thicker layer made up of cardiomyocytes, the cardiac muscle cells that allows the heart to contract and thus to pump blood. The endocardium is the smooth and thin layer composed by endothelial cells in contact with the blood pool. Inside the ventricles the papillary muscles and trabeculations prevent blood regurgitation and can also assist the contraction.

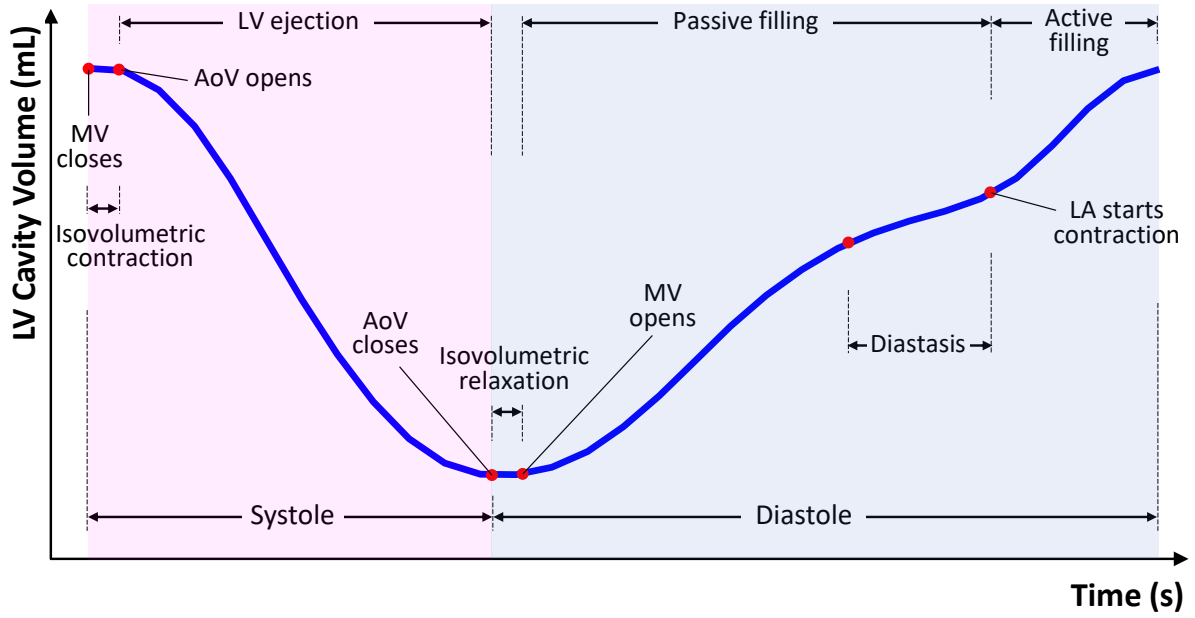


Fig. 2.4: Representative LV volume transient of a healthy heart. AoV indicates aortic valve; MV, mitral valve; and LA, left atrium. During isovolumetric relaxation, LV pressure decreases until it is below the left atrium pressure causing the mitral valve to open, initiating the filling of the ventricle. During isovolumetric contraction, LV pressure increases until it is higher than the aortic pressure, triggering the opening of aortic valve. Diastasis is the stage of diastole where the initial passive filling of the ventricles has slowed, but before the atria contract to complete the active filling [132].

Physiologically, the mechanical contraction is generated by an electrical impulse that propagates through the conducting myocytes. This impulse triggers the shortening of the sarcomeres, the basic contractile units of the contracting myocytes, through an ionic exchange. The electrical signal

originates in the sinoatrial node, a group of self-excitable cells that acts as a natural pacemaker controlling the heart rate. It then stimulates both atria, inducing their contraction. After that, it spreads to the atrioventricular node in around 50-100 ms, which ensures that the ventricles are filled. Finally, it rapidly propagates to the ventricles through the bundle of His and the Purkinje fibres, and from there to the rest of the LV and RV, causing the ventricular contraction [131].

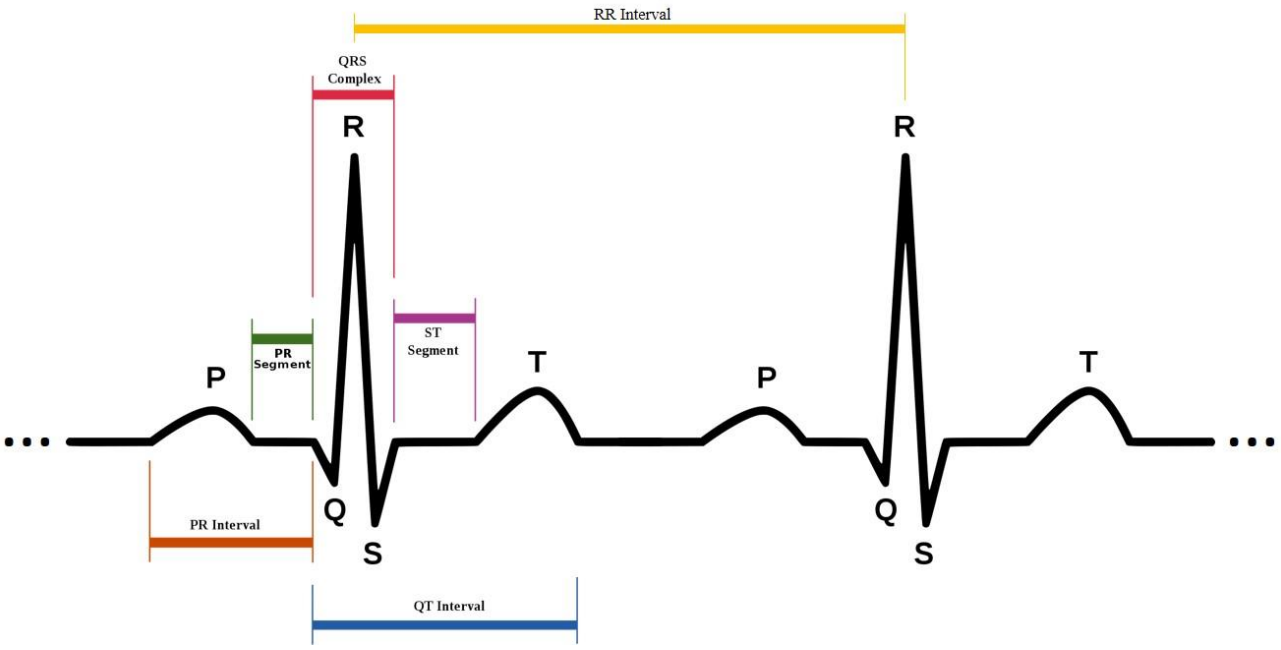


Fig. 2.5: ECG diagram of a healthy heart in normal sinus rhythm. Adapted under Open Access license from Wikipedia.org [133].

The described electrical propagation, illustrated in Fig. 2.3.B, can be assessed non-invasively using electrocardiography (ECG). This technique places electrodes on the skin to record the small electrical signals resulting from the cardiac muscle depolarisation and the subsequent repolarisation of the different regions of the heart [134]. The conventional 12-lead ECG uses ten electrodes, six placed on the chest and four on the limbs at specific positions, to measure the voltage of the heart over time from twelve different angles. Thus, the overall magnitude and the direction of the electrical activity of the heart is captured. In a healthy heart, three main components of an ECG can be identified throughout the cardiac cycle: (1) the P wave, representing the depolarisation of the atria; (2) the QRS complex, corresponding to the depolarisation of the ventricles; and (3) the T wave, representing the

final repolarisation of the ventricles [131]. The PR segment connects the P wave and the QRS complex, and the ST segment joins the QRS complex to the T wave. The QRS complex is itself composed of 3 waves: the Q wave (septum depolarisation, sometimes not visible), the R wave (ventricular early depolarisation) and the S wave (ventricular late depolarisation). Since the R wave is the most identifiable signal on the ECG, the heart rate is calculated using the R-to-R wave interval, also known as the R-R interval. Fig. 2.5 depicts the described components, waves, and segments.

2.2.2. Cardiovascular Diseases

Cardiovascular diseases, a group of heart and blood vessels disorders including hypertrophic cardiomyopathy (HCM) and AMI, are the leading cause of death both in industrialised and developing countries. They are responsible for 17.9 million annual deaths (31% of all deaths) worldwide, with the death toll further climbing each decade [3], [6]. This translates into 1.9 million deaths in the European Union (42% European mortality), with an intrinsic total cost of €169 billion annually. Just in the United Kingdom, 160,000 deaths are registered per year [6]. Cardiovascular research attempts to achieve a deeper understanding of cardiac structure and its function, to improve diagnosis and treatment, and thus to reduce this burden [3]–[5]. While important advances have been made in improving detection and therapies in CVD, the topic remains one of the most important challenges in medical research.

2.2.3. Hypertrophic Cardiomyopathy

HCM is a monogenic autosomal disease caused primarily by mutations in genes encoding sarcomeric proteins, responsible for myofilament contraction. This disease can lead to left ventricle hypertrophy (LVH), myocyte disarray, fibrosis, electrophysiological abnormalities, and ionic remodelling [135]–[137]. While most patients have normal longevity and remain asymptomatic, major complications including progressive heart failure, atrial fibrillation, embolic stroke, and sudden cardiac death may be developed [136], [138]. HCM is the most common genetic cardiomyopathy, affecting 1 in 500

individuals (~1 million in the European Union), and the principal cause of sudden cardiac death in the young. The identification of patients at risk remains a challenge [53], [136], [139]. Some of the barriers are the high phenotype heterogeneity, a variable disease onset and the low incidence (annual mortality rate < 1% per year), which complicates endpoint studies [136]. In addition, although genetic testing can confirm some positive genotypes, HCM is usually not diagnosed until late disease, when the first symptoms appear. Thus, the relation between clinical outcome and genetic mutation has not yet been established.

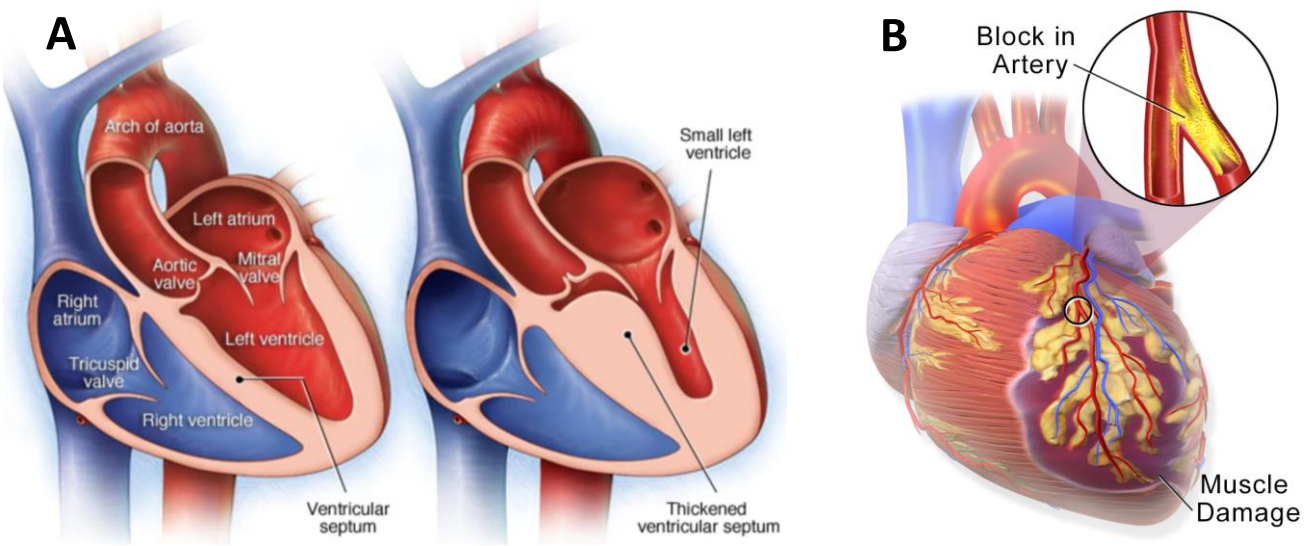


Fig. 2.6: **A**: Normal anatomy (left) vs HCM characteristic anatomy positive in LVH (right). Reproduced under Open Access license for non-commercial use from Mayo Foundation for Medical Education and Research [140]. **B:** Illustration of a coronary artery blockage, which prevents blood from perfusing to the myocardium. This leads to ischemia and, ultimately, to myocytes necrosis and fibrosis, that is, myocardial infarction. Reproduced under Open Access license from Wipedia.org [141].

According to phenotype, patients in HCM can be grouped in: a) phenotype negative; b) phenotype positive in the form of abnormalities such as myocardial fibrosis, but LVH negative; and c) phenotype and LVH positive, diagnosed as maximal LV wall thickness > 15 mm in adults in the absence of other cardiac disease [138], [142] (see Fig. 2.6.A). In the first group, typically consisting of younger patients, sudden cardiac death incidence is higher. It usually occurs with no warning, triggered by heavy physical effort [136]. However, patients in group three usually score at higher risk, since they are likely to suffer from frequent episodes of ventricular tachyarrhythmia and syncope. Treatment

depends on the prognosis of the patients. It usually involves implantable cardioverter-defibrillators to prevent ventricular arrhythmias, the precipitating mechanism in sudden cardiac death. Surgical ablation to remove outflow obstructions, anti-arrhythmic drugs (i.e., beta blockers), LV assist devices, or even heart transplants may also be considered [138]. The invasiveness, side-effects, and costs of the treatment prevent low implementation thresholds (i.e., high sensitivity at the cost of low specificity) and highlight the importance of risk stratification. Current risk stratification algorithms are based on factors such as age, family history, maximal LVH, unexplained collapse, non-sustainable ventricular tachycardia, left atrium size, and blood pressure drop on exercise [138], [143], [144]. Despite the new prediction model included in the latest European Society of Cardiology guidelines, accurate and personalised prediction of outcomes is limited, has low sensitivity, and is still far from being solved [145], [146].

2.2.4. Acute Myocardial Infarction

AMI is defined as myocardial cell necrosis due to prolonged ischemia. It occurs when the blood flow, and therefore the oxygen supply, to a part of the heart is partially or totally blocked [147]. Coronary heart diseases in the form of coronary artery occlusion, usually due to the rupture of an atherosclerotic plaque, represent the main cause of AMI (70% of the events) [7], [147]. This is illustrated in Fig 2.6.B. Other less common obstructing mechanisms include plaque-free coronary artery spasms triggered by emotional stress, extreme cold, and cocaine consumption, among others [7], [147]. Chest pain, discomfort, and shortness of breath are the most common symptoms. Nonetheless, about 30% of people have atypical symptoms and even “silent” AMI may occur, especially in the elderly. Complications during and after AMI include arrhythmias, cardiogenic shock, or even heart rupture. Treatment depends on the severity of the AMI, the overall condition of the patient, and the elapsed time since AMI started. It primarily consists in the use of antiplatelet agents (i.e. aspirin, heparin, etc.), thrombolytic drugs to dissolve the blood clot, and/or surgery to restore the blood flow (percutaneous coronary intervention, coronary artery bypass grafting, etc.) [7]. ECG is used to

classify AMI depending on severity in ST segment elevation (STEMI, popularly known as "heart attack" – See Fig 2.7) and non-ST elevation myocardial infarction (NSTEMI). While the blockage of a coronary artery is usually total in STEMI, the blood flow may be only partially blocked in NSTEMI, resulting in a smaller damaged section of the heart [7]. The actual heart damage is usually assessed with a blood test measuring troponin. This is a protein integral to cardiac muscle, which is released into the bloodstream upon myocyte necrosis. Coronary angiography, an invasive technique in which an X-ray contrast is delivered through a catheter, may be used to locate the exact location of the blockage in the arteries [7].

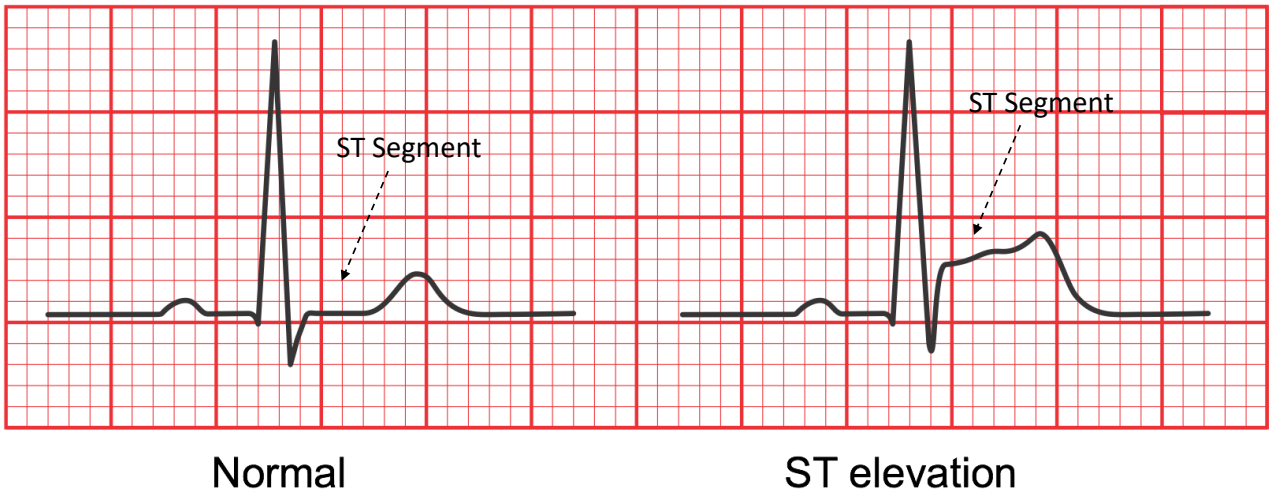


Fig. 2.7: Diagram of a normal ECG (left) vs a ST-elevated ECG (right). Adapted from [131], under Open Access license for non-commercial use.

In 2017, about 10.6 million AMI episodes were registered worldwide, 3 million of those being severe (STEMI) [12]. Despite a huge improvement in the last decade, mortality within 6 months after infarct sits at 12% [148].

2.2.4.1. Left Ventricular Remodelling After Myocardial Infarction

The acute myocardial cell necrosis, and the associated increase in LV load, induces a cascade of adaptative changes in both the infarcted and remote myocardium, which typically include the formation of scar tissue, hypertrophy, and ventricular dilation [149], [150]. Within hours of myocyte injury, during an early remodelling phase, an increase in wall thinning and ventricular dilation are

usually observed [149]. This results in a higher wall stress and triggers an increase in heart rate and myocyte shortening as temporary compensating mechanisms. Late remodelling, in turn, involves changes at tissue level and the formation of collagen scar to stabilize the distending forces [149].

This late LV remodelling is a crucial factor for recovery [8], [151]. Patients with pumping function deterioration are likely to suffer from severe cardiac events and exhibit an increase in end-systolic volume. In contrast, patients with pumping function improvement are characterised by an increase in LV wall thickening, both in the infarct as well as in the remote zones [10]. Post-infarction remodelling involves complex multi-scale processes regulated by mechanical, neurohormonal and genetic factors [149], [150]. For this reason, it is not completely understood how acute structural-functional changes modulate long-term remodelling and outcomes.

This thesis focuses on the understanding of the mechanical (rather than electrophysiological) adaptative responses. In this context, CMR has been proven successful to account for remodelling and predict prognosis after episodes of AMI (see Sections 2.3.1 and 2.3.4) [9], [152].

2.3. Cardiac Magnetic Resonance

2.3.1. CMR in Imaging Diagnosis

Early and accurate diagnosis is essential in the clinical management of CVD. Medical imaging has contributed to assess cardiac function since sonomicrometry was first used in 1956 [153], [154]. Sonomicrometry was based on tracking piezoelectric crystals that were previously embedded in the ventricular wall, and falls into the category of invasive techniques. Non-invasive techniques such as echocardiography, CMR, and X-ray computed tomography have since been developed [155]. Echocardiography is the leading cardiac imaging technique for a first exploration because of its availability and affordability [18]. CMR is the gold standard for quantitative assessment of cardiac structure and function whenever disease is suspected [18], [19], [156], [157]. Moreover, CMR does not use ionising radiation in contrast to computed tomography.

Non-invasiveness, accuracy, reproducibility, and spatial resolution justify the routine use of CMR in clinical practice [19]. CMR provides high quality information about cardiovascular anatomy. This allows accurate quantification of parameters such as ventricular volume, myocardial mass, and wall thickness [19], [156]. Furthermore, the time dimension of the technique provides functional information by capturing several instances of the cardiac cycle. Important indicators of AMI such as ejection fraction, the percentage of blood that the LV pumps in each cycle, can therefore be calculated [156]. The well-established late gadolinium enhancement CMR, based on the diffusion of a magnetic contrast agent for the identification of fibrotic regions, further expands the potential of CMR to in-vivo tissue characterisation [138], [158]. More recent techniques include T1 mapping and diffusion tensor imaging. While T1 mapping is based on magnetisation relaxation times (see Section 2.3.2), in diffusion tensor imaging the natural mobility of water molecules in tissue is measured. These two techniques could also play an important role in clinical tissue characterisation including assessment of fibrosis or detection of stenosis [138], [159]–[161]. However, they have not yet been adopted in clinical routine.

In addition to tissue characterisation, cardiac motion can be accurately tracked via CMR tagging and, more recently, strain-encoded imaging [162], [163]. Myocardial tissue planes are marked at the beginning of the cycle with radio frequency pulses, orthogonal to the image plane in CMR tagging and parallel in strain-encoded imaging. This allows to calculate velocity and strain in different areas of the heart [156], [163]. The standard 17-segment model of the heart may be used as template for comparison and dysfunction identification [164]. An alternative to these two methods for strain quantification is CMR-FT, a post-processing technique based on the tracking of image features. While less robust, CMR-FT does not require acquisition of additional imaging sequences. For this reason, CMR-FT is the preferred technique for myocardial deformation assessment in clinical practice [165], [166]. Among the techniques mentioned above, this Thesis will primarily focus on standard CMR,

but it also uses late gadolinium enhancement and CMR-FT in Chapters 5 and 6 to build advanced risk models for post-infarction risk prediction (see Section 2.3.4).

There are, however, some challenges for the use of CMR in clinical practice [167]. CMR scans are usually more costly and require larger acquisition times than computed tomography or X-ray imaging. Moreover, CMR is susceptible to imaging artifacts as consequence of cardiac, respiratory, and patient motion (see 2.3.2. CMR principles) [168]. Finally, its use is limited for patients suffering from claustrophobia, patients unable to hold the breath, or in patients wearing metal implants, given the strong magnetisation process involved in the acquisition [169]. The challenges specific to the analysis of CMR scans, such as the imaging heterogeneity introduced by different scanners and protocols, are discussed in Section 2.4.

Given the above discussed potential of CMR to non-invasively provide structural and functional information, the use of this modality in HCM for detailed imaging and biomarker phenotyping studies is strongly advocated [136], [142], [170]. CMR is likewise the gold standard in AMI to assess infarct segments, both location and severity, and LV acute remodelling according to clinical guidelines [9], [11], [18], [19], [152], [158]. Hence, the use of CMR in this Thesis to investigate the anatomical and functional interplay that modulates adverse outcomes is motivated.

2.3.2. CMR Principles

CMR is based on the spatial variations of the electromagnetic energy that the anatomical tissues emit or absorb when excited [155], [171]. The spin of hydrogen nuclei (protons), naturally abundant in biological tissues, aligns to a strong static magnetic field in the scanner. In response to a radio frequency pulse, protons, excited, tilt away from alignment into the transverse plane [171]. When the pulse is removed, magnetisation returns gradually to its equilibrium state releasing energy in the form of radio frequency signal, captured by the scanner. The magnetic response, which can be characterised by T1 (longitudinal magnetisation recovery) and T2 (transverse magnetisation decay) relaxation,

depends on tissue type [171], [172]. Three magnetic gradients are used to achieve spatial localisation by populating the k-space, which is reconstructed as an image using Fourier transforms [171], [173].

Cine CMR is the most common type of study [18]. In this CMR modality, the heart is imaged at multiple time points throughout the cardiac cycle and therefore rich information of cardiac function is acquired. The magnetic sampling is limited to dozens of milliseconds to minimise cardiac motion artefacts during each image acquisition. Thus, the relaxation time (image quality) and the number of phase encodings (spatial resolution) are highly restricted. Steady-State Free Precession pulse sequences are preferred since they achieve high signal-to-noise and contrast-to-noise ratios, while maintaining low sensitivity to motion [172]. ECG gating, usually based on the R-wave, is used to synchronise acquisition with heart motion. This ensures that the heart is always imaged at the same phase points, and therefore reduces artefacts because of cardiac motion [171].

The long acquisition times along with respiratory and cardiac motion critically compromise resolution in 3D acquisition [174]. Hence, acquisition of 2D slices for different image planes is preferred. Whenever the patient is able to do so, each of these 2D slices is acquired within a breath-hold and includes an instance of the different cardiac phases, each of them imaged within a different heartbeat [171]. For this reason, CMR is particularly susceptible to motion artifacts, as mentioned before. This leads to slice misalignment and, ultimately, to 3D inconsistencies. Slice-to-slice [175] and slice-to-volume-registration [176]–[178] algorithms, based on the interplay between cardiac and respiratory motion, have been proposed to correct for motion artefacts [179]–[181]. In case of poor breath-holding or arrhythmia, information for diagnosis can still be obtained with single-shot acquisition or arrhythmia rejection algorithms [171].

The image planes are standardised as follows: two perpendicular planes transecting the axes that connect the apex and the centre of the mitral valve, the horizontal and vertical long-axes; and a stack of planes perpendicular to those from the apex to the base, short-axis (SAx) [164]. Fig. 2.8 depicts

these standard CMR planes. A left ventricular outflow tract and a right ventricular long axis plane (parallel to the LV vertical long axis plane) may also be acquired.

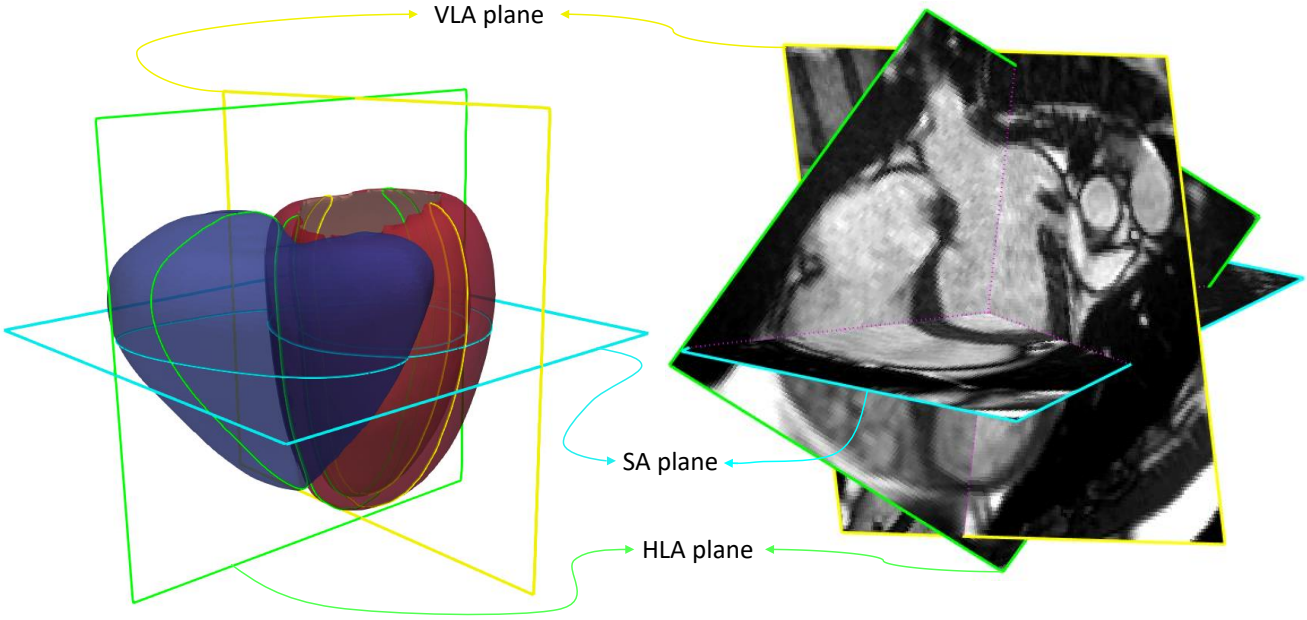


Fig. 2.8: CMR standard planes: model (left) and real acquisitions (right). HLA: horizontal long axis; VLA: vertical long axis; SA: short axis.

2.3.3. CMR Analysis in Clinical Practice

In clinical practice, the analysis of CMR images is manually conducted by a radiologist [182]–[185]. The end-systolic (ES) phase is usually determined by visual assessment, selecting the phase in which the LV blood pool is smallest at mid-ventricular level. The end-diastolic (ED) phase is assumed to be the first frame in the temporal sequence, because of the ECG gating [182]. The LV endocardium, LV epicardium, left atrium, and RV endocardium are then contoured, and metrics are derived from them to evaluate cardiac function [156], [182], [186]. LV mass, ejection fraction, stroke volume (SV) and cardiac output are the principal global metrics along with ESV and EDV [156], [182], [186], [187]. EDV represents the total amount of blood filling the ventricles at ED. It is indicative of the size of the ventricle, and sits around 120 mL for normal patients. LV mass is calculated as the volume between the LV epi and endocardium surfaces times the myocardium density (ρ_m = 1.05 g/cm^3) [186]. Consistency in LV mass in ES and ED might be used as a contouring quality indicator [182].

SV and cardiac output account for the blood pumped per heartbeat and per minute, respectively [186]. The former is estimated as EDV minus ESV, and the latter is equal to SV times the heart rate. They can be calculated in both ventricles, obtaining similar values in normal hearts [182], [186]. SV and cardiac output are frequently normalised to total body surface area or weight for inter-patient comparison. Ejection fraction usually refers to LVEF, and represents the ratio between SV and ED volumes, in percentage [186]. While SV is roughly estimated to be around 75 mL, normal values of LVEF are in the range of 65-70% [188]. The Simpson's rule, based on an interpolated summation of the SAx cross-sectional areas, or the bi-plane area-length method, are most commonly used for volume calculation [182]–[184]. However, the assumption of orthogonality of SAx planes with respect to the ventricle may lead to error [186]. Local metrics such as wall thickening and strain may also be considered in clinical studies [156], [186].

Manual contouring is laborious, tedious, and highly time-consuming. For these reasons, only the ES and ED phases are usually segmented, even though metrics from the whole cycle could be systematically retrieved. Moreover, manual assessment is subject to intra- and inter-observer variability and prone to errors [182], [189]. The aforementioned expansion of big data in cardiovascular medicine has led to the consequent enlargement of datasets to thousands of patients [13], [14], [16], [48]. The UK Biobank project, which had recruited more than 500,000 subjects by 2019, including 25,000 CMR scans, serves as an example of this [72]. Manual analysis of such number of images is no longer feasible and further highlights the need of automated segmentation and analysis [156], [157], [190].

The above-mentioned metrics have proven successful in both diagnosis and prognosis of CVD [18], [19], [156]. For instance, the LVEF threshold of 35% is widely established to stratify patients in low (LVEF>35%) and high (LVEF<35%) risk, and to provide treatment to them accordingly. Nevertheless, there might be additional information other than ventricular volumes encoded in the heart shape that could potentially lead to more accurate and personalised predictions. This would

ultimately provide a better understanding of CVD. Heart shape analysis has therefore become an active field of cardiac research [186], [191]–[193].

2.3.4. CMR and ML in Acute Myocardial Infarction

Remodelling after an infarct episode is central to early prognosis, as explained in Section 2.2.4 [7], [10]. Accordingly, LVEF and ESV, which quantify the LV macro-function, have been established as the preferred imaging markers for post-infarction risk assessment and therapy decision-making [11], [194]–[198]. CMR has proven uniquely suitable to regress these markers. Moreover, CMR can further assess and describe anatomical and functional myocardial alterations after an infarct, which provide incremental prognostic information over LVEF and ESV [9], [199], [200]. As an example, late gadolinium enhancement CMR has become the gold standard to quantify infarct size and microvascular obstruction. These two markers have emerged as robust outcome measures to assess micro-myocardial injury in AMI patients towards risk stratification [159, 164–166]. Given its spatial resolution, CMR has been likewise applied to quantify remodelling changes at different regions of the ventricle [201], [202]. CMR-FT has been proven successful to quantify myocardial local deformation indexes and mechanical uniformity contraction alterations. These contraction-related markers are associated with hemodynamic alterations, adverse LV remodelling, and, ultimately, clinical outcomes [200], [203]–[206].

Multiple studies have shown the strong prognostic value of CMR-based risk models in AMI management [194], [205]. For instance, Siermaier et al. proposed a simple risk score system, which attributes ‘risk’ points based on LVEF, infarct size, and microvascular obstruction being superior to certain thresholds [207]. This model was validated in an independent cohort and proven significantly superior to the LVEF risk model proposed in clinical guidelines [11], [207]. More examples can be found in [200], where Reindl et al. provide an overview of the current impact of CMR imaging on optimised risk assessment in STEMI. Nevertheless, these models have not yet relegated the sole

LVEF-based risk assessment in clinical practice, often done via echocardiography [194], [205]. The more common availability of echocardiography, the length of the CMR acquisitions, the requirement for manual interaction, and the complexity of CMR multi-parameter models have been among the barriers for adoption [11], [205]. However, CMR post-infarction protocols have been shortened and CMR availability has significantly increased in recent years. Moreover, metrics such as ESV or LVEF can be fully-automatically determined using commercially available software [157], [190], [208]. In addition, complex CMR prognostic predictors can be synthesised into simple risk score models facilitating their use in clinical practice, as in the Siermaier et al. example [205], [207]. Despite the proven prognostic value of CMR multi-parameter models and the limitations being addressed, further randomised trials are needed to accelerate the use of CMR as risk stratification tool in daily clinical routine [200], [206]. This topic is further discussed in Paper V and Paper VI.

Methodologically, ML approaches are increasingly used to build risk models for post-infarction management [209], [210]. Indeed, the vast majority of the state-of-the-art ML studies focused on AMI are aimed to predict secondary adverse outcomes [210]. This applies not only to CMR-based models, but to AMI risk models in general. Alternative biomarkers range from blood tests, ECG recordings, to demographic data [210]–[213]. Approaches based on NN, random forests, support vector machines, and Bayesian techniques have been proposed to predict adverse events after infarction [210]–[214]. For instance, Kwon et al. trained a NN on demographic information and laboratory data to predict 12-month mortality [212]. This study is particularly interesting given the extraordinarily large size of the testing set (10,000+ patients) that was used to validate their proposed method. Likewise, Li et al. explored up to six different ML approaches towards the same purpose [213]. The authors of these studies reported a significant improvement in performance of the ML approaches compared to the risk models proposed in the clinical guidelines [212], [213]. Nevertheless, there is controversy in the literature with respect to the actual benefits of ML classifiers for AMI risk stratification in comparison to conventional statistical models (i.e., linear, logistic and

Cox regression) [209]–[211], [214]. Only a few months ago, Min-Cho et al. systematically reviewed the studies that were published in the last decades comparing ML and conventional models for prognosis after infarct [209]. They concluded that although the ML approaches may lead to a prediction improvement, the absolute differences in prediction are small and the proposed ML models usually lack a rigorous validation [209].

Other applications of ML in AMI can be found in diagnosis and detection of infarction [210], [215]–[217]. For example, Younghoon et al. coupled a variational autoencoder (see Section 2.4) to a fully connected NN to detect infarcts using only 6-lead ECG, as acquired in some wearable devices [217]. Than et al. successfully applied DL to personalise (based on the age, sex, and time between samples) the troponin thresholds that are used for the diagnosis of infarction [215]. As a final example, Wu et al. also proposed a fully connected NN to classify patients with chest pain into NSTEMI or angina based on cardiovascular risk factors and laboratory data [216]. While the NN deployed was very basic, they reported a classification AUC close to one. Indeed, the troponin test used in clinical practice for AMI diagnosis has already a high specificity and sensitivity [11]. This could explain why most of the ML applications in AMI focus on post-infarction management, where the room for improvement is much larger.

2.4. Automated 3D CMR analysis

2.4.1. Traditional Methods for Automated CMR Segmentation

Image segmentation consists in the pixel-wise labelling of an image into predefined region or classes. Cardiac segmentation is usually a pre-requisite for the calculation of image biomarkers, as explained in the sections above, and for shape analysis. Despite the satisfactory soft tissue contrast resolution of CMR, the automated segmentation of cardiac images is not an easy task. Accurate, consistent, and fast automated segmentation of CMR scans has not yet been fully implemented in clinical routine

[156]. The anatomical variability, imaging heterogeneity, and the intrinsic complexity of cardiac dynamics and geometry represent a challenge [156].

Traditionally, semi- or fully-automated segmentation methods are grouped into image-driven and model-driven approaches. The former are based on weak prior or non-prior knowledge, while the latter use strong priors. However, it is difficult to classify methods in practice, since they frequently combine both types of approaches. Thresholding, region-growing, and active contour-based techniques fall into the first category, and are often used in combination [156], [218], [219]. Thresholding localises the region of interest (ROI) based on pixel intensity histograms. Region-growing algorithms initialise seeds placed in the ROI, which expand searching for similar pixels nearby until none of the surrounding pixels qualify any longer [218], [219]. Active contour methods do not search for pixels, but for contours instead. They do so by minimising the energy of a parameterised curve, usually roughly initialised by hand [220], [221]. The energy functions include several forces accounting for smoothness, target fitting, and convergence, among others [156]. On the other hand, atlas-based segmentation and statistical shape modelling are examples of model-driven methods. In the former, pixel classification is performed by registering the image to one or several atlases, which represent the reference values and variations within the population [25]. Statistical shape modelling approaches segment the images by finding the set of PCA mode weights of the model that best match the image [88], [191], [222]–[225].

2.4.2. ML Methods for Automated CMR Segmentation

ML techniques have been used in CMR segmentation for a long time [156], [226]. Supervised methods have been explored for pixel and voxel classification, such as random forests, k-nearest neighbours, support vector machines, and NN [227]–[229]. Unsupervised methods have been likewise applied to segmentation, including k-means clustering and Gaussian mixture models [230].

Bayesian formulations and regression forests have also been used for direct estimation of cardiac metrics, that is, without segmenting the images first [231], [232].

2.4.2.1. The U-Net Architecture

In 2015, the U-Net architecture introduced by Ronneberger et al. [40] marked a watershed in segmentation. This accelerated a shift in the segmentation field towards DL approaches, which have proven to achieve similar performance to expert manual segmentation [15], [226], [233]. In brief, U-Net is a CNN that consists of a contraction and an expansion path, as shown in Fig. 2.9. The contraction path extracts the features of the image. This is done by concatenating sequences of convolutional layers, followed by a rectified linear unit activation and a max-pooling operation. The expansion path recovers the resolution of the image by combining spatial information with "skip connections", following the order of resolutions. These skip connections act as a shortcut between feature maps in the contraction and expansion paths that are at the same resolution [40]. Then, blocks of up-convolutions and concatenations, to incorporate the mentioned information from skip connections, integrate the expansion path.

Since then, a number of variations of CNNs have been explored reporting excellent performance [41], [233]–[237]. For instance, Newell et al. proposed the 'stacked hourglass' network, an architecture based on successive pooling and up-sampling steps, conceptually similar to several U-Nets in sequence [238]. The underlaying hypothesis is that the re-evaluation and assessment of initial predictions will improve the performance. The authors demonstrated the potential of this approach for pose estimation. Vigneault et al. used this 'stacked hourglass' in combination with a classical U-Net that pre-segments and transforms the images into a canonical orientation [235]. The pre-processing step removes variance and improves label balance, which they hypothesised facilitates the fine segmentation task. Zhoe et al. introduced an architecture where the U-Nets are nested, called U-Net++ [239]. They argued that thanks to their re-design the feature maps from the decoder and encoder paths are semantically similar, and that this eases the optimisation of the NN in training.

Valanarasu et al. deployed an architecture where an additional ‘inverted’ U-Net is run in parallel, which claims to better preserve detailed image information [240]. This architectural modification specifically targeted the detection of small anatomical structures with blurred and noisy boundaries, where they showed an improvement in comparison to a baseline U-Net.

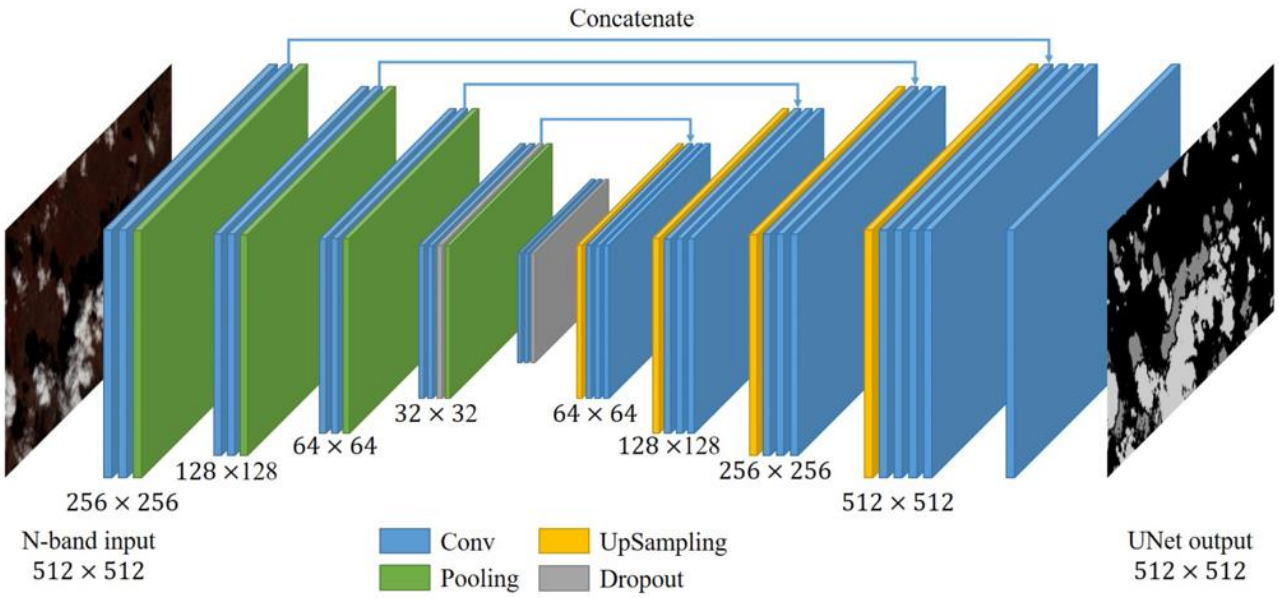


Fig. 2.9: Illustration of U-Net architecture, reproduced under Open Access license from [241].

Baumgartner et al. introduced a Bayesian framework in the up-sampling path to output not only the segmentation mask, but also a prediction of uncertainty [242]. They reasoned that this accounts for the inherently ambiguous nature of the segmentation of anatomical structures. Oktay et al. developed an Attention U-Net for pancreas segmentation, which automatically learns to focus on target structures of varying shapes and sizes [243]. This architecture was later successfully applied to CMR, achieving impressive results in segmentation performance [244]. Bai et al. introduced a variation of VGG-16 to segment 2D cine CMR [41]. In VGG-16, the feature maps learnt at different levels are up-sampled and concatenated to the input resolution in a single step, contrary to U-Net [245]. In their work, they used a dataset of 5,008 subjects, the largest to date in CMR segmentation [41]. Therefore, their achieved performance (average Dice metric of 0.94 for the LV cavity, 0.88 for the LV myocardium and 0.90 for the RV cavity) may be used as a benchmarking framework. It could

be argued that it is not fully clear whether the contribution to performance comes from the proposed architecture or from the size of the dataset. A large dataset allows an increase of CNN depth without over-fitting, and so a potentially more complex non-linear mapping that could yield a better performance.

2.4.2.2. Incorporation of Anatomical Knowledge

Several methods have been explored to incorporate anatomical knowledge more explicitly into the final segmentation. Many include this anatomical information either via pre-trained autoencoders (AEs) or by modifying the segmentation NN to enforce the learning of anatomical features [236], [246], [247]. The architecture of the AE is similar to that of the U-Net, but without skip connections. AE consists of an encoder network to convert the input into a dense and compact representation (latent space) and a decoding branch that reconstructs the images [61], [248]. Both Larrazabal et al. and Painchaud et al. proposed the use of AEs to postprocess the segmentation of a separately trained U-Net [246], [249]. These AEs, trained using ground truth segmentations, learn the space of plausible anatomies to transform the output of the U-Net into an anatomically meaningful segmentation. Nevertheless, the anatomical improvement may come at the cost of a slight drop in prediction performance [249].

Yue et al. also used a postprocessing AE unit trained on ground truth segmentation, but its output was included in the loss function to train the segmentation U-Net [250]. While they claimed superior performance to other state-of-the-art regularization schemes, the dice scores obtained were relatively modest. Alternatively, Chen et al. used pre-trained AEs to encode information from other CMR planes and add it to the latent space of the segmentation U-Net [251]. In this way, the 2D U-Net incorporates spatial context from multiple and complementary views, which span a 3D space. This showed to outperform both baseline 2D and 3D U-Net architectures, while keeping a higher data efficiency. Nevertheless, there is no guarantee that the final segmentation output is anatomically consistent,

unlike in the previously described works. Future work could attempt to combine the Chen et al. approach with postprocessing AEs to overcome this limitation.

Regarding other NN modifications, Otkay et al. introduced the use of a regularisation block on top of the CNN [236]. This block enforces the model predictions to follow the distribution of the learnt low dimensional representations or priors. The authors showed that this approach could be potentially applied to both segmentation and super resolution tasks. Designing the NN to explicitly factorise the images into spatial anatomical factors and imaging modality factors has also been proposed [252]. Chartsias et al. argued that this explicit disentanglement of the features is particularly useful in multimodality and multitasking settings [252]. Other examples include the combination of DL with atlas propagation [247], [253]. Duam et al. used this type of approaches to correct for 2D SAx misalignment towards the generation of anatomically meaningful 3D cardiac shapes [253]. Modification in the loss function to ensure continuity and consistency have also been proposed [254]–[260]. For instance, Byrne et al. designed topological loss functions for the evaluation of image segments against an explicit prior to improve segmentation quality. They claimed to resolve segmentation inconsistencies without sacrificing overlap performance [260]. More recently, the use of graph NN to introduce shape reasoning has been explored [261], [262]. This latter type of approaches has shown great potential, especially for the segmentation of highly variable structures. Nonetheless, the prediction improvement with respect to CNN for general segmentation applications is arguable.

2.4.2.3. Temporal Dynamic Modelling and Direct Methods

Temporal dynamic modelling aims to incorporate cardiac motion in the segmentation process, ensuring consistency throughout the cardiac cycle [30]. The underlaying hypothesis is that the temporal information would compensate the errors introduced by frame-by-frame quantification. Recurrent NNs, proven useful in video analysis [263]; CNNs, extensively used in language processing [264]; and, more recently, attention-based NNs, have been explored for this purpose [30], [265]–

[269]. Nevertheless, they have not yet replaced the use of static segmentation. This is probably because, in most datasets, the ground truth segmentations remain contoured frame-by-frame and are not necessarily time continuous, limiting the performance of these approaches. Direct quantification methods, both temporal and static, based on CNNs have also been proposed [30], [31], [45], [270]. However, they do not outperform the indirect regression of metrics from automated segmentations, as shown in externally validated segmentation challenges [271], [272]. Likewise, direct contouring of the structures, rather than pixel-wise segmentation, has been explored via graph NNs [273]–[275]. While these NNs efficiently handle the coordinates of the contours as nodes of a graph, an improvement in performance has not been consistently proven.

2.4.2.4. Augmentation and Transfer Learning Techniques

U-Nets therefore remain the state-of-the-art method for automated segmentation and quantification in CMR. As a learning-based approach (see Section 2.1.1), data availability critically affects performance and may be a limitation [61], [73], [74], [190], [276]. In this context, data augmentation is standard practice to enlarge datasets. Image cropping, translation, rotation, scaling, flipping, and even deformation are routinely applied to enlarge datasets [61], [233], [276], [277]. Paper II further elaborates on the state-of-the-art augmentation techniques for segmentation [278].

Another major challenge towards automated segmentation is the imaging heterogeneity introduced by different scanners, centres, and acquisition protocols [279], [280]. While techniques such as transfer learning have proven useful for changes in imaging heterogeneity or modality [281]–[283], retraining for every unseen domain is still required [284]. In order to address this limitation, solutions based on appearance augmentation techniques and domain adaptation techniques have been developed. Histogram matching is an example of the former, which focuses on exploiting the training data available [285]. While very simple, it has proven powerful for multidomain applications in recent segmentation challenges [272]. In contrast, domain adaptation methods introduce modifications in the NN to enforce a single feature representation for all domains that is invariant to domain but

discriminative for the task of interest [244], [272], [286]–[290]. This is explained in more detail in Paper IV, where two domain adaption techniques are implemented [291].

In addition, pre-processing steps and NN architectural modifications have been explored to address the class imbalance of the segmentation labels [235], [292]. The previously described configuration proposed by Vigneault et al., or my work presented in Paper III, may serve as example of this [235], [293]. Finally, the use of ensembles is strongly recommended for improved precision and accuracy in the predictions [26], [272], [294], [295].

2.4.3. 3D Shape Reconstruction

The spatial consistency of 3D analysis leads to more accurate estimations of global metrics compared to 2D estimates. Moroseos et al. evidenced this for RV measurements, by comparing the Simpson's method estimations based on 2D segmentations against estimations from a 3D reconstruction of the ventricle [296]. Higher dimensional anatomical information may also enhance the analysis of regional function, as shown by Medrano-Gracia et al. in the context of regional wall motion [297], [298]. In addition, personalised 3D shape models are increasingly used for surgical planning, as well as for population or electromechanical modelling studies (see Section 2.1.2) [88], [299]–[305].

Obtaining smooth and dense 3D meshes, however, represents a challenge due to the nature of CMR studies. As mentioned before in Section 2.3.2, 3D acquisition does not currently provide sufficient spatial resolution. In turn, 2D scans are likely to be affected by motion artefacts because of multiple breath holdings and possible patient movement during acquisitions. This can cause spatial inconsistency [299]. Moreover, the inter-slice resolution, usually around 10 mm, is rather limited. The discrepancy in segmentation that may result from a 2D analysis introduces noise (i.e., papillary muscles included in horizontal long-axis segmentation but excluded in SAx). Furthermore, a 3D ground truth is usually non-existent [299]. The overall result is a sparse, non-uniform, noisy and misaligned set of data points [306]. Pre-processing of these data usually includes the correction for

misalignment mentioned in Section 2.3.2. Alignment methods are either based on the registration of the individual slices to a reference 3D volume or among them, minimising dissimilarity, which can be formulated on either feature or voxel terms [175]–[178], [307].

2.4.3.1. Types of surface reconstruction methods

Surface reconstruction can be classified into point-based and model-based approaches [299], [306]. Point-based ones are further divided into region growing, computational geometry, and functional methods. Whereas in region growing methods an initial seed propagates to all the points connecting them, computational geometry algorithms explore the surrounding neighbours of all the surface points deciding the best connectivity [299]. Delaunay triangulation is an example of the latter, widely used in CMR [308], [309]. Given a set of points, functional methods aim to find the function that best approximates the surface, such as Poisson surface reconstruction [310]. In contrast, model-based approaches fit a template mesh to the cloud of points (e.g. [311]). Despite their extensive use, the resulting mesh is inevitably biased to the template. Hence, since the mesh template is typically learnt from normal anatomies, these model-based methods may be inaccurate in the presence of anatomical abnormalities and pathological cases [223].

There is always a trade-off between data-fitting and smoothness of the resulting mesh [299]. Some of the explored solutions to account for both include iterative processes that smooth the mesh while increasing its density (i.e., piecewise smooth subdivision surfaces) [296], [312], [313]. Alternatively, Villard et al. used compositions of smooth deformations on a smooth surface, where the smoothing constrain is progressively relaxed as the number of fitting interactions increases [299]. Clinically, the most accepted 3D reconstruction method is the guide-point modelling approach, where a model shape is fitted to the guided points, interactively introduced (or removed) by the user in the 2D slices [314], [315]. The guide-point modelling approach is in use in clinical practice, distributed by Siemens, the largest MRI manufacturer. Moreover, it has been validated against autopsy tissue and independent

estimates of cardiac output. All this justifies the use of guide-point modelling approach as reference mesh when no ground truth is available [314].

2.4.3.2. Deep learning in 3D heart reconstruction.

Deep learning approaches have become overwhelmingly popular for 3D segmentation and reconstruction [16], [17], [316]. While comparable accuracy is achieved, these DL models reduce the computation time to the order of seconds versus the several hours that the previously mentioned algorithms may take for accurate 3D reconstructions [316]. Moreover, they are able to account for misalignment correction and shape reconstruction all at once and in a fully automated manner. Examples of DL configurations towards these aims include the 3D U-Net proposed by Xu et al. [317]. Given a set of 2D CMR sparse and misaligned segmentations, this U-Net outputs a 3D consistent and dense voxel grid, from which a bi-ventricular mesh is extracted applying the marching cubes algorithm [318]. Another DL alternative can be found in the pipeline developed in [319], which combines a CNN with a registration and a deformation step to reconstruct the LV cavity in 3D.

Generative models have also been used for 3D segmentation and reconstruction, facilitated by the representation of shapes by voxel grids [17], [320]. Among them, generative adversarial networks (GANs) and variational autoencoders (VAEs) are used the most. GANs are based on the idea of two NNs trained simultaneously that compete against each other [321]. While one generates synthetic instances (generator), the other discriminates between real and synthetic samples (discriminator). As the discriminator improves, the generator learns to produce even more realistic instances to confuse the discriminator [321]. While 3D U-Nets combine low and high-level features, they cannot guarantee spatial consistency in the final segmentation map. However, they are good generator candidates in a GAN architecture, where the discriminator acts as a shape regulariser affecting the internal features of the segmenter to ensure 3D consistency [17], [322]–[324]. With respect to VAEs, their main advantage in comparison to the classical AEs is that the latent space is biased to ensure continuity and has a Gaussian distribution [325], [326]. Thus, new samples can be easily drawn from this latent

space configuration to effectively generate synthetic data [325], [326]. Conditional VAEs allow for additional information for regularisation, such as image intensity features or labels [327], [328]. The same applies to conditional GANs [329], [330].

Dynamic modelling could also be considered for shape reconstruction, for instance including in the latent space long-short-term-memory modules, a type of recurrent NNs [331], [332]. Furthermore, DL methods can also implement template priors for shape reconstruction. Wickramasinghe et al. presented a pipeline from which the final mesh is obtained through an iterative process of deformations, ruled by successive feature maps obtained from the original 3D image [333]. This approach avoids the challenging post-processing step usually needed in most DL methods to obtain a mesh from the resulting labelled volume (voxel grid). In turn, Tóthová et al. implemented a DL method to estimate the weights (and the uncertainty) of a PCA shape model to fit a sparse volumetric image [334].

None of these deep learning approaches can handle irregular and unstructured data [335]. Thus, the segmentation point clouds are transformed into 3D regular grids of voxels, rendering the resulting data unnecessarily voluminous and introducing quantification artefacts [335]. Moreover, GPU requirements grow cubically with volume size in 3D convolution-based methods, limiting the input to subvolumes of the actual 3D instances [17], [336], [337]. To overcome these limitations, Qi et al. proposed the PointNet network, which works directly with point clouds of different size [335]. In theory, the network is invariant to point permutations and rigid motions, as well as robust to input perturbation and corruption [335]. PointNet has proven useful for imaging classification and segmentation [335], [338], [339]. It was formally applied to 3D cardiac shape reconstruction for the first time by our research group only a few months ago, at ISBI 2021 [340]. Other alternative solutions include hierarchical or sparse representations, or the use of structured meshes. Examples of these can be found in [341] and [342], respectively.

2.4.4. 3D Shape Analysis

Shape analysis of the resulting reconstructed 3D meshes can be based on simple shape descriptors such as moments (volume, centroid, etc.) or on the distributions of all the points composing the mesh [343]. Examples of the former are LV volume, wall thickness, endocardial surface curvature, and sphericity index [344], [345]. These are intuitive and do not require point correspondence among meshes, which effectively implies registering all the meshes (usually to a pre-defined template) to find homologous points [346]. However, the high dimensionality of point-distribution approaches may result in better prediction accuracy by finding hidden trends in data [192], [297], [347]. In these approaches, PCA is usually applied to reduce the dimensionality once the point correspondence has been established. The resulting modes may serve as new potential biomarkers and their prediction power can be accordingly assessed [52], [71], [301], [347]–[353].

Contrary to linear discriminant analysis, PCA does not consider class membership for dimensionality reduction, and does not maximise classification power [354]. Nevertheless, linear discriminant analysis cannot be applied if the number of variables is greater than the minimum number of class elements. In shape analysis this is the usual scenario, with thousands of points (variables) but only hundreds of samples [355]. PCA can be applied as a dimension reduction pre-step, to later apply linear discriminant analysis on the resulting PCA modes to maximise the prediction power of the modes in combination [355]. Information-maximising component analysis may also be used for the optimisation of classification tasks. Contrary to linear discriminant analysis, it does not assume Gaussianity in class distribution [354]. In addition, other more complex dimensionality reduction techniques, alternative to PCA, can be considered to bias the orthogonal decomposition to further improve interpretability. For instance, Zhang et al. applied partial least square regression to associate the shape modes of variation with well-established LV indexes of cardiac remodelling [356].

Other analysis approaches consider the deformation fields resulting from the registration of the template to the meshes or from inter-frame registration, rather than coordinates of corresponding points [192], [357]. Aderkani et al. used the former to discriminate among different aetiologies of LV hypertrophy (HCM vs hypertensive heart disease) [192]. Yang et al. applied the latter to analyse 3D desynchrony, and to ultimately improve prediction of response to cardiac resynchronization therapy [358]. Furthermore, DL autoencoders could be also explored for dimensionality reduction [359]. Finally, DL methods for shape analysis and outcome prediction have been developed to directly handle meshes as inputs, rather than 3D images [360], [361]. However, the use of such approaches in medical imaging is still scarce and most of the DL attempts for shape analysis are based on heavy 3D arrays, such in [362].

CHAPTER 3

Mechanistic and Statistical Modelling Synergy towards Precision Cardiology

3. Mechanistic and Statistical Modelling Synergy towards Precision Cardiology

3.1. Chapter Publications

- **Paper I**: J. Corral Acero et al. "The 'Digital Twin' to enable the vision of precision cardiology". *Eur. Heart J.*, Mar. 2020, 41:4556–64

3.2. Chapter Outline

Computational modelling of human heart function has been recently boosted by the big data explosion in cardiology and the increasing power of computers and algorithms [13], [14], [24]. This translates into an intersection between the increased capacity to capture extensive data about individual patients and the ability to digest it. Section 2.1 has highlighted the major role that computer models play in present and future precision cardiology, and it has outlined two different modelling approaches [2]. The 'deductive' mechanistic models are constrained by their premises, that is, their assumptions and principles [13], [14], [25]. Hence, they represent a good choice when the underlying mechanistic understanding of the system is profound. In contrast, the 'inductive' statistical models are a good fit when the mechanisms of the system are rather complex or remain poorly understood. However, they are constrained by the observations available, that is, the amount and diversity of data [23], [24]. Therefore, there exist clinical needs that cannot be solved by deploying a single modelling approach.

The position paper that drives this chapter argues for the combination of the mechanistic and statistical approaches to address the above-mentioned limitations. The article introduces the current role of computational cardiology in personalised medicine and the basic modelling concepts. It then elaborates on the synergistic use of the two modelling approaches towards the 'digital twin' of a patient. The first section of the article focuses on this model synergy towards clinical diagnosis, therapy decision-making, and prognosis. The next section explores their combined use to accelerate clinical translation. It includes examples towards model validation and as critical tools for regulatory

decision-making. A final section discusses the impact, timeline for adoption, and organisational, societal, and ethical challenges ahead. The manuscript concludes with recommendations to facilitate their intake in clinical contexts. The supplemental material of this article, which can be found online in [2], reviews the model synergies for exploiting, maximising, and integrating clinical data.

The input of the 14 Personalised in-silico Cardiology fellows of my European International Training Network, co-authors of this position paper, is acknowledged. Likewise, the contribution of Prof. Pablo Lamata, without whom this paper would have never been possible, shall be deeply thanked. My contribution to the manuscript is detailed in the Statement of Authorship, located at the end of this chapter. In essence, I was one of the main contributors to the literature review and the preparation of the manuscript.

The vision herein presented drives this Thesis, and, therefore, the next chapters.

Paper I

The 'Digital Twin' to enable the vision of precision cardiology

Jorge Corral-Acero, Francesca Margara, Maciej Marciniak, Cristobal Rodero, Filip Loncaric, Yingjing Feng, Andrew Gilbert, Joao F Fernandes, Hassaan A Bukhari, Ali Wajdan, Manuel Villegas Martinez, Mariana Sousa Santos, Mehrdad Shamohammdi, Hongxing Luo, Philip Westphal, Paul Leeson, Paolo DiAchille, Viatcheslav Gurev, Manuel Mayr, Liesbet Geris, Pras Pathmanathan, Tina Morrison, Richard Cornelussen, Frits Prinzen, Tammo Delhaas, Ada Doltra, Marta Sitges, Edward J Vigmond, Ernesto Zacur, Vicente Grau, Blanca Rodriguez, Espen W Remme, Steven Niederer, Peter Mortier, Kristin McLeod, Mark Potse, Esther Pueyo, Alfonso Bueno-Orovio, Pablo Lamata

European Heart Journal, Volume 41, Issue 48, 21 December 2020, Pages 4556–4564
(https://doi.org/10.1093/eurheartj/ehaa159)

ESC European Society of Cardiology

European Heart Journal (2020) **41**, 4556–4564
doi:10.1093/eurheartj/ehaa159

STATE OF THE ART REVIEW
Disease management

The 'Digital Twin' to enable the vision of precision cardiology

Jorge Corral-Acero[1], Francesca Margara [2], Maciej Marciniak [3], Cristobal Rodero [3], Filip Loncaric [4], Yingjing Feng [5,6], Andrew Gilbert [7], Joao F. Fernandes [3], Hassaan A. Bukhari[6,8], Ali Wajdan[9], Manuel Villegas Martinez[9], Mariana Sousa Santos[10], Mehrdad Shamohammdi[11], Hongxing Luo [11], Philip Westphal[12], Paul Leeson [13], Paolo DiAchille [14], Viatcheslav Gurev [14], Manuel Mayr [15], Liesbet Geris [16], Pras Pathmanathan[17], Tina Morrison[17], Richard Cornelussen[12], Frits Prinzen[11], Tammo Delhaas [11], Ada Doltra [4], Marta Sitges [4,18], Edward J. Vigmond [5,6], Ernesto Zacur [1], Vicente Grau [1], Blanca Rodriguez [2], Espen W. Remme[9], Steven Niederer [3], Peter Mortier[10], Kristin McLeod [7], Mark Potse [5,6,19], Esther Pueyo [8,20], Alfonso Bueno-Orovio [2], and Pablo Lamata [3]*

[1]Department of Engineering Science, University of Oxford, Oxford, UK; [2]Department of Computer Science, British Heart Foundation Centre of Research Excellence, University of Oxford, Oxford, UK; [3]Department of Biomedical Engineering, Division of Imaging Sciences and Biomedical Engineering, King's College London, London, UK; [4]Institut Clínic Cardiovascular, Hospital Clínic, Universitat de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain; [5]IHU Liryc, Electrophysiology and Heart Modeling Institute, fondation Bordeaux Université, Pessac-Bordeaux F-33600, France; [6]IMB, UMR 5251, University of Bordeaux, Talence F-33400, France; [7]GE Vingmed Ultrasound AS, Horton, Norway; [8]Aragón Institute of Engineering Research, Universidad de Zaragoza, IIS Aragón, Zaragoza, Spain; [9]The Intervention Centre, Oslo University Hospital, Rikshospitalet, Oslo, Norway; [10]FEops NV, Ghent, Belgium; [11]CARIM School for Cardiovascular Diseases, Maastricht University, Maastricht, The Netherlands; [12]Medtronic PLC, Bakken Research Center, Maastricht, the Netherlands; [13]Radcliffe Department of Medicine, Division of Cardiovascular Medicine, Oxford Cardiovascular Clinical Research Facility, John Radcliffe Hospital, University of Oxford, Oxford, UK; [14]Healthcare and Life Sciences Research, IBM T.J. Watson Research Center, Yorktown Heights, NY, USA; [15]King's British Heart Foundation Centre, King's College London, London, UK; [16]Virtual Physiological Human Institute, Leuven, Belgium; [17]Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD, USA; [18]CIBERCV, Instituto de Salud Carlos III, (CB16/11/00354), CERCA Programme/Generalitat de, Catalunya, Spain; [19]Inria Bordeaux Sud-Ouest, CARMEN team, Talence F-33400, France; and [20]CIBER in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain

Received 6 September 2019; revised 29 November 2019; editorial decision 16 February 2020; accepted 24 February 2020; online publish-ahead-of-print 4 March 2020

Providing therapies tailored to each patient is the vision of precision medicine, enabled by the increasing ability to capture extensive data about individual patients. In this position paper, we argue that the second enabling pillar towards this vision is the increasing power of computers and algorithms to learn, reason, and build the 'digital twin' of a patient. Computational models are boosting the capacity to draw diagnosis and prognosis, and future treatments will be tailored not only to current health status and data, but also to an accurate projection of the pathways to restore health by model predictions. The early steps of the digital twin in the area of cardiovascular medicine are reviewed in this article, together with a discussion of the challenges and opportunities ahead. We emphasize the synergies between mechanistic and statistical models in accelerating cardiovascular research and enabling the vision of precision medicine.

Keywords Precision medicine • Digital twin • Computational modelling • Artificial intelligence

* Corresponding author. Tel: (+44) 20 784 89563, Email: pablo.lamata@kcl.ac.uk
© The Author(s) 2020. Published by Oxford University Press on behalf of the European Society of Cardiology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

Introduction

Providing therapies that are tailored to each patient, and that maximize the efficacy and efficiency of our healthcare system, is the broad goal of precision medicine. The main shift from current clinical practice is to take inter-individual variability into greater account. This exciting vision has been championed by the -omics revolution, i.e., the increasing ability to capture extensive data about the pathophysiology of the patient.[1,2] This -omics approach has already delivered great achievements, especially in the management of specific cancer conditions.[3] Nevertheless, the initial conception of precision medicine has already been criticized for being too centred in genomics and failing to address challenges of clinical management.[4] The concept is thus gradually widening, shifting from the original gene-centric perspective to the wide spectrum of lifestyle, environment, and biology data.[5,6]

In this context, we argue that the definition of optimal therapy options requires a mechanistic understanding that links all levels from genetic and molecular traces to the pathophysiology, lifestyle and environment of the patient, and back. Precision medicine requires, not only better and more detailed data, but also the increasing ability of computers to analyse, integrate, and exploit these data, and to construct the 'digital twin' of a patient. In health care, the 'digital twin' denotes the vision of a comprehensive, virtual tool that integrates coherently and dynamically the clinical data acquired over time for an individual using mechanistic and statistical models.[7] This borrows but expands the concept of 'digital twin' used in engineering industries, where *in silico* representations of a physical system, such as an engine or a wind farm, are used to optimize design or control processes, with a real-time connection between the physical system and the model.[8]

This position paper claims that precision cardiology will be delivered in a synergetic fashion that combines induction, by using statistical models learnt from data, and deduction, through mechanistic modelling and simulation integrating multiscale knowledge and data.[9] These are the two pillars of the digital twin (*Figure 1*). We review the state of the art of the interplay between such models that supports this vision, considering that there are already excellent independent review papers in the fields of statistical[14–16] and mechanistic[17,18] models for cardiovascular medicine.

Mechanistic models encapsulate our knowledge of physiology and the fundamental laws of physics and chemistry. They provide a framework to integrate and augment experimental and clinical data, enabling the identification of mechanisms and/or the prediction of outcomes, even under unseen scenarios without the need for retraining.[19] Examples of such mechanistic models are the bidomain equations for cardiac electrophysiology[20] or the Navier–Stokes equations for coronary blood flow.[21] In a complementary manner, statistical models encapsulate the knowledge and relations induced from the data. They allow the extraction and optimal combination of individualized biomarkers with mathematical rules. Examples of statistical models applied to computational cardiology are random forests for assessment of heart failure severity[22] or Gaussian processes to capture heart rate variability.[23]

There are clinical needs that can be solved with a single modelling approach. But both mechanistic and statistical models have limitations that can be addressed by combining them. Mechanistic models are constrained by their premises (assumptions and principles), while statistical models are constrained by the observations available (the amount and diversity of data). A mechanistic model may be a good choice when a good understanding of the system is available. A statistical model, on the other hand, can serve to find predictive relations even when the underlying mechanisms are poorly understood or are too complex to be modelled mechanistically. The rest of the article describes the synergies between mechanistic and statistic models (see *Figure 2* for an overview), motivated by actual clinical problems and needs, with specific representative components of the digital twin. Supplementary material online reviews the model synergies for exploiting and integrating clinical data.

Mechanistic and statistical model synergy for improving clinical decisions

Technical, ethical, and financial constraints limit the data acquisition needed to assist clinical decision-making.[14,15] Synergy between mechanistic and statistical models has shown value in aiding diagnosis, treatment, and prognosis evaluation. A fully developed digital twin will combine population and individual representations to optimally inform clinical decisions (*Figure 3*).

Model synergy in aiding diagnosis

Models can pinpoint the most valuable piece of diagnostic data. An example is the simulation study that revealed that fibrosis and other pulmonary vein properties may better characterize susceptibility to atrial fibrillation.[24] Models can also reliably infer biomarkers that cannot be directly measured or that require invasive procedures. For instance, the combination of cardiovascular imaging and computational fluid dynamics enables non-invasive characterizations of flow fields and the calculation of diagnostic metrics in the domains of coronary artery disease, aortic aneurysm, aortic dissection, valve prostheses, and stent design.[25–29]

The key to guide diagnosis is the personalization of a mechanistic model to the actual health status of the patient as captured in available clinical data. In this personalization process, statistical models enable robust and reproducible analysis of clinical data and infer missing parameters. An example of this synergy is the assessment of left ventricular myocardial stiffness and decaying diastolic active tension by fitting mechanical models to pressure data and images during diastole.[30,31] Another example is the non-invasive computation of pressure drops in flow obstructions,[32,33] such as aortic stenosis or aortic coarctation, which has been proven more accurate than methods recommended in clinical guidelines.[34] Models have also been used to derive fractional flow reserve from computed tomography (CT) to non-invasively identify ischaemia in patients with suspected coronary artery disease, avoiding invasive catheterized procedures.[29,35–37]

Some diagnostic medical devices based on personalized mechanistic models have already reached their industrial translation and clinical adoption. HeartFlow FFR_{CT} Analysis (HeartFlow, USA) and CardioInsight (Medtronic, USA) use patient-specific mechanistic

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

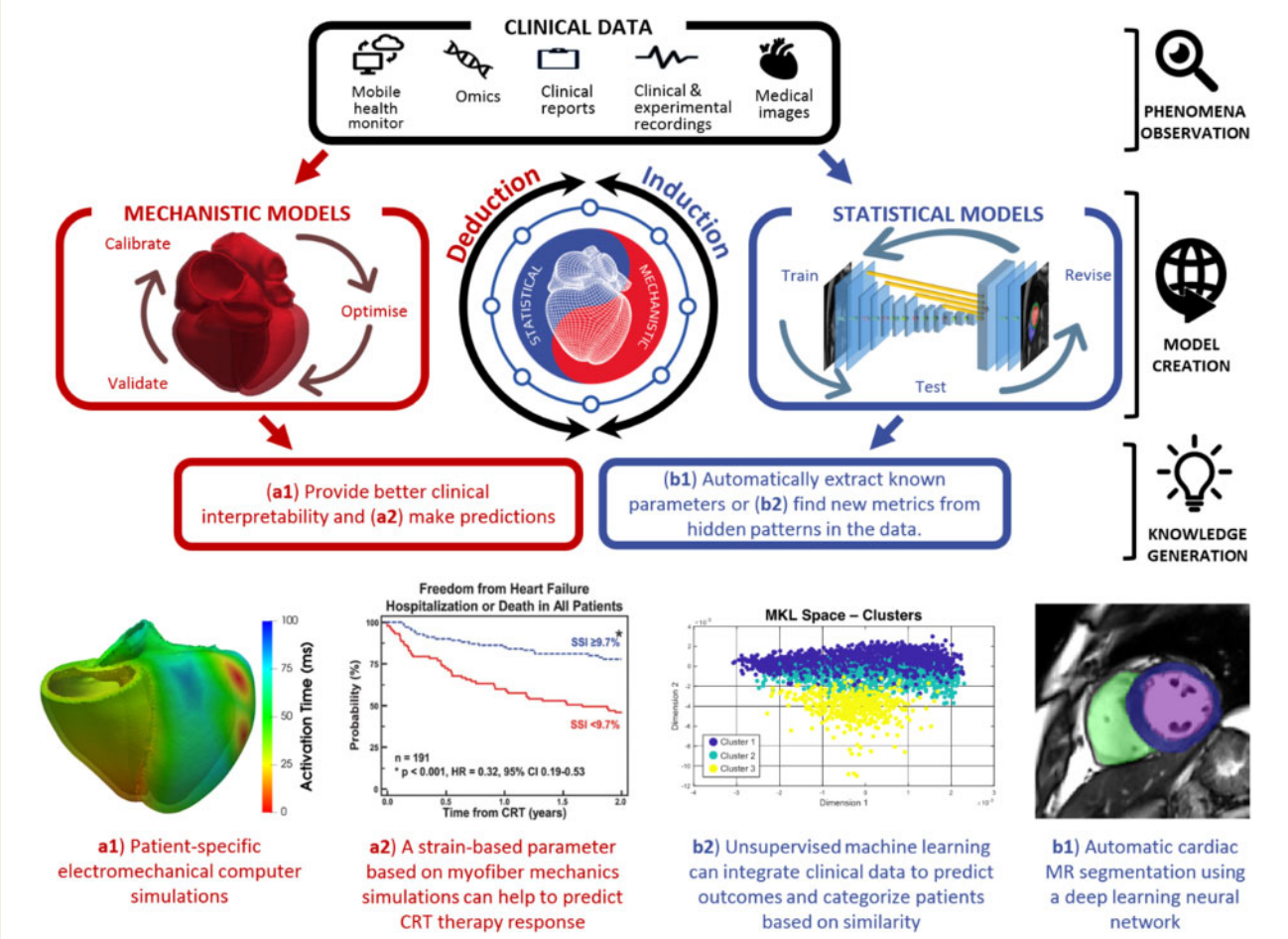


Figure 1 The two pillars of the digital twin, mechanistic and statistical models, illustrating its construction and four examples of use: **a1**,[10] **a2**,[11] **b1**,[12] **b2**.[13]

models to non-invasively calculate clinically relevant diagnostic indexes and have received clearance from the USA Food and Drug Administration (FDA).[38] HeartFlow predicts fractional flow reserve by means of a personalized 3D model of blood flow in the coronary arteries.[36] In the CardioInsight mapping system, the electrical activity on the heart surface is recovered from body surface potentials using a personalized model of the patient's heart and torso.[39]

Model synergy in guiding treatments

A digital twin may indicate whether a medical device or pharmaceutical treatment is appropriate for a patient by simulating device response or dosage effects before a specific therapy is selected.

The benefits of cardiac resynchronization therapy (CRT) have been demonstrated in patients with prolonged QRS duration. However, uncertainty remains in patients with more intermediate electrocardiogram (ECG) criteria.[40] To guide decision-making in this 'grey zone', approaches using mechanistic modelling have investigated the role of different aetiologies of mechanical discoordination in CRT response.[10] For example, a novel radial strain-based metric was defined based on simulations of the human heart and circulation to differentiate patterns of mechanical discoordination, suggesting that the response to CRT could be predicted from the presence of non-electrical substrates.[11] Statistical methods were used to verify these findings in a clinical cohort, and the novel index remained useful in predicting response in the clinical 'grey zone', creating the opportunity to improve patient selection in the group with intermediate ECG.

Another example is the improvements in ablation guidance of infarct-related ventricular tachycardia, where the accurate identification of patient-specific optimal targets is provided before the clinical procedure.[41] Mechanistic models can propose novel electro-anatomical mapping indices to locate critical sites of re-entry formation in scar-related arrhythmias, aid acquisition and quantitative interpretation of electrophysiological data, and optimize future clinical use.[42]

The industrial translation and clinical adoption of models for guiding treatment are exemplified by the optimal planning of valve prosthesis with the HEARTguide™ platform (FEops nv, Belgium), or by the platform to guide ventricular tachycardia ablations (inHeart, France).

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

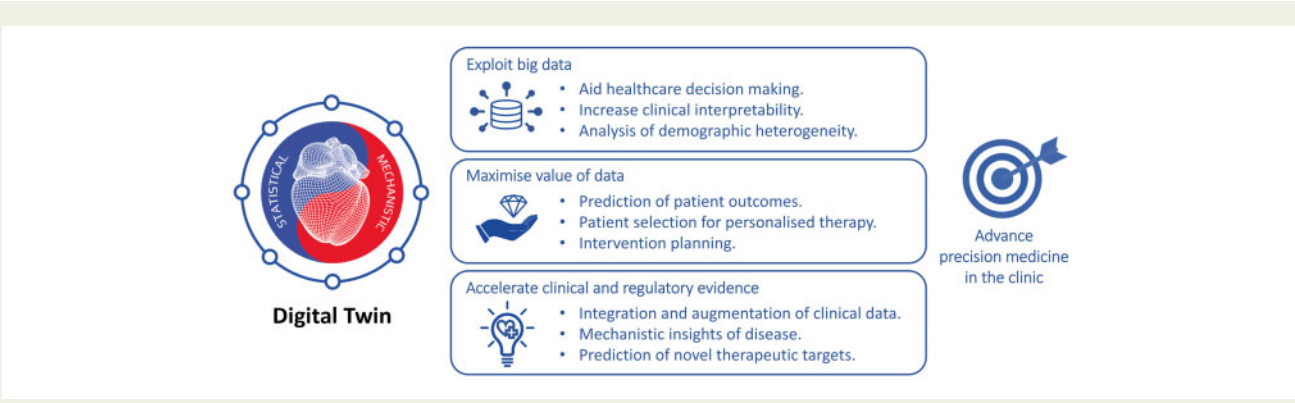


Figure 2 Conceptual summary of the main benefits of digital twin technologies.

Model synergy in evaluating prognosis

While statistical modelling allows for categorizing patients based on the probability of various outcomes, mechanistic modelling provides more insights to support or reject the categorization.

For example, model synergies represent an exciting approach to interpret structure–function relationships and improve risk prediction in inherited disease conditions, such as hypertrophic cardiomyopathy (HCM). Relationships among specific ECG changes, ventricle morphologies, and sudden cardiac death have been inferred from observations.[43,44] However, the complex process of translating underlying heterogeneous substrates in HCM to ECG findings is still poorly understood, and there exists a 'grey zone' of clinical decision-making in the low-risk patient subgroups, specifically when deciding on restriction of involvement in professional sports.[45] In this context, by using methods of statistical inference and mathematical modelling (see *Figure 4*), HCM patients were categorized into phenogroups based on ECG biomarkers extracted from 24-h ECG recordings,[48] and the aetiology of each ECG phenogroup linked with different underlying substrates, suggesting ion-channel and conduction system abnormalities.[46,47] The results directly highlighted the potential of personalized anti-arrhythmic approaches in the treatment of HCM patients, and addressed the low-risk patients, showing that a normal ECG might indeed be the discriminatory factor signalling minimal ionic remodelling, fibrosis, disarray, and ischaemia in these 'grey zone' patients.

Models have also been used in the prediction of arrhythmic events in post-myocardial infarction, outperforming existing clinical metrics including ejection fraction.[49] When the amount of data is not sufficient to inform state-of-the-art machine learning methods, statistical methods can still prove useful. An example is the use of principal component analysis to account for right ventricular motion in predicting survival in pulmonary hypertension,[50] or to identify signatures of anatomical remodelling that predict a patient's prognosis following CRT implantation.[51]

While statistical models allow predictions, mechanistic models provide the underlying explanations. Understanding the actual meaning of the selected features improves the plausibility of findings and increases their credibility. For both approaches, quantifying uncertainty of prediction can help identify cases that may require further review, while building trust in cases where models are shown to be robust.[52,53]

Mechanistic and statistical model synergy to accelerate evidence generation

While digital twin technologies in cardiology show promising research results, only a small number of models have reached clinical translation. The difficulties encountered include the need to increase validation, lack of clinical interpretability, and potentially obscure model failures.[54] Therefore, solid evidence for the generalization of preliminary findings and efficient testing strategies are needed. Even when these barriers are overcome, rigid assessment of algorithmic performance and quality control from regulatory bodies can slow down the adoption. In this context, model synergy can be used to accelerate the integration of novel technologies into clinical practice by increasing clinical interpretability, validating generality of findings, and accelerating regulatory decision-making.

Model validation towards generality of findings

The goal after validating an initial concept is to extend it to a more general patient cohort, with less controlled characteristics. The problem of sampling bias, based on both intrinsic (physiological) and extrinsic (environmental) demographic heterogeneity of the population, becomes relevant when implementing solutions for broader patient cohorts.[55,56] Consequently, models (as clinical guidelines) may need recalibrations when used on populations from different countries or ethnicities, or even from different centres in the same country. In recent years, only 6% of artificial intelligence algorithms had external evaluation performed (note this is beyond the minimum requirement of using the learning, validation, and testing partitions of the data), and none adopted the three design criteria of a robust validation: diagnostic cohort design, the inclusion of multiple institutions, and prospective data collection.[57] The quality of datasets also needs to be thoroughly validated to avoid possible biases before the models developed from them can be integrated in clinical decision-making.[58]

To address this issue, an increasing number of institutions are creating initiatives for data-sharing platforms, aiming at reusing existing

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

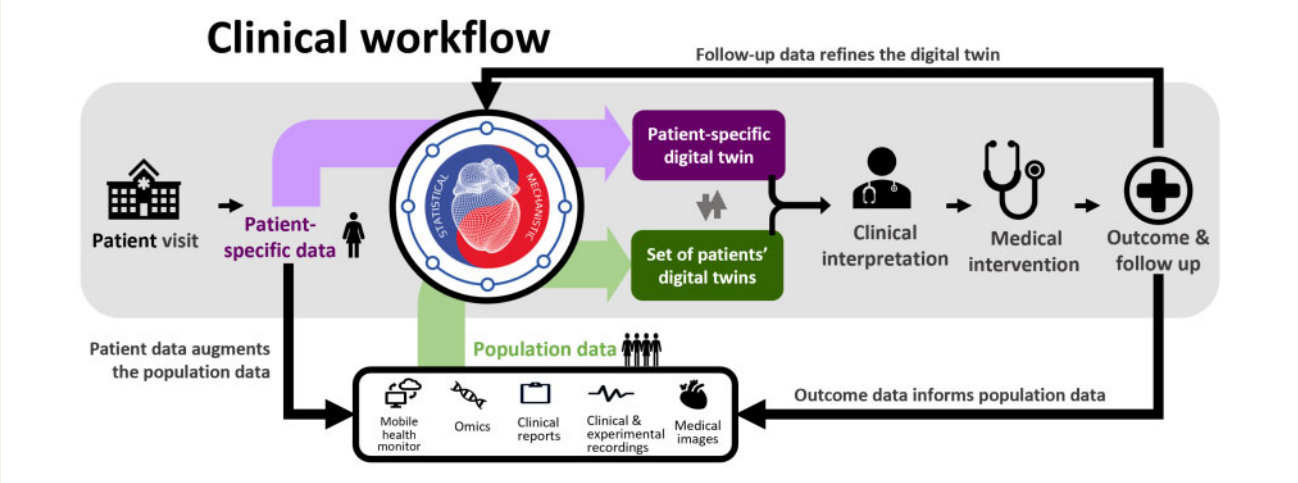


Figure 3 Envisioned clinical workflow using the fully developed digital twin concept. Population data, collected from preceding patients and study cohorts, are used to create and validate statistical and mechanistic models, as well as to create a population-based digital twin (green). Novel patient data are analysed with the help of the existing models and integrated to form the patient's digital twin (purple). The comparison and interaction between digital twins give valuable insight (phenotyping, risk assessment, prediction of disease development. . .) that is clinically interpreted and combined with traditional data to aid in the process of clinical decision-making. The digital twin develops in line with the patient's condition—adjusting and improving in accordance with the follow-up data. Resulting outcomes are supplemented to shape population data and refine the follow-up data.

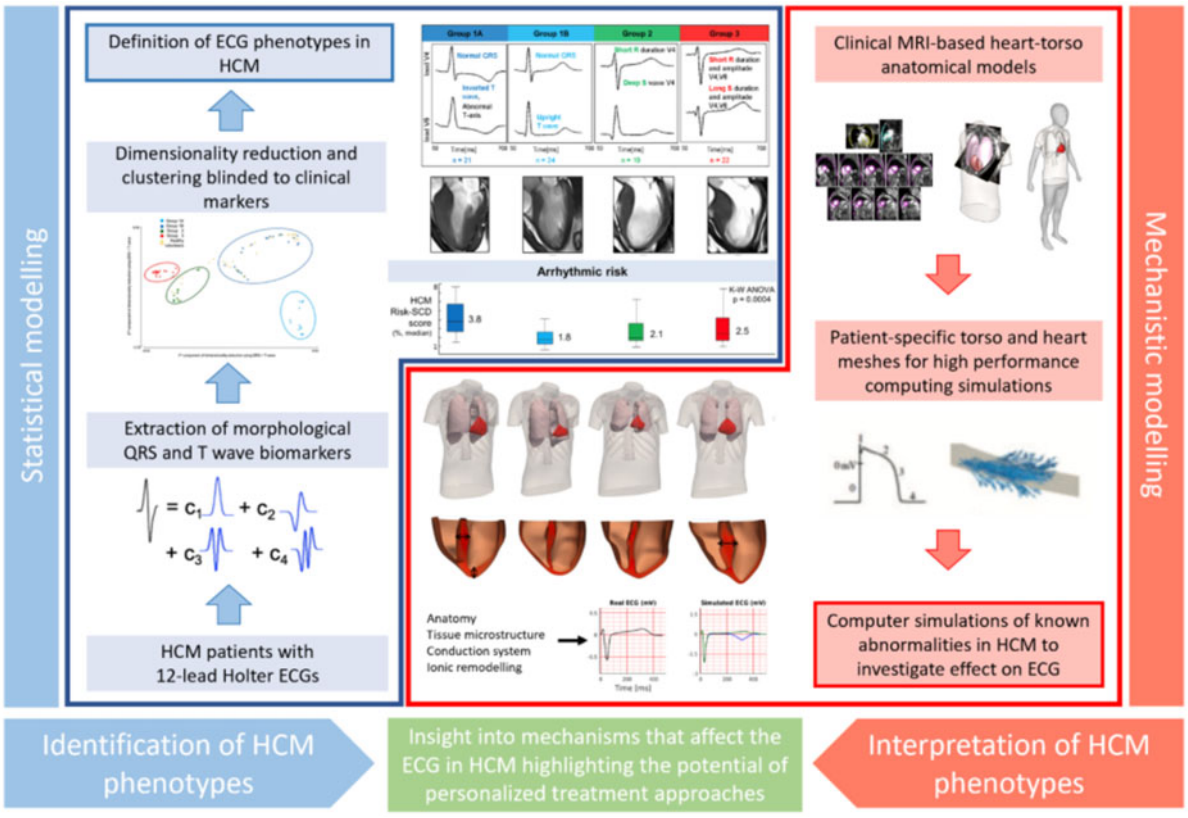


Figure 4 Synergy between mechanistic and statistical models in the definition of electrocardiogram (ECG) biomarkers for the management of hypertrophic cardiomyopathy.[47,48]

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

datasets and verifying published research works.[59] Governments, regulatory agencies, and philanthropic funders are promoting the open science culture, enforcing publishing patient-level data by means of compliance to product launching, funding application, and journal publishing.[60]

Another approach to improve the generality of data is the generation of synthetic cases of a representative wider population. The core idea is to expand the average mechanistic model to obtain populations of models, all of them parameterized within the range of physiological variability obtained by experimental protocols.[61,62] Such an approach, which allows investigating many more scenarios than possible experimental acquisitions, is not only able to evaluate the impact of physiological variability but to explain the mechanisms underpinning inter-individual variability in therapy response (e.g. adverse drug reactions), and to identify sub-populations at higher risk.[63,64] Statistical shape modelling techniques can represent inter-patient anatomical variability for a cohort, and be used in combination of mechanistic models for clinical decision support systems.[65]

As in traditional scientific research, mechanistic and statistical models are complementary tools to verify the findings derived from one another. Finding a mechanistic explanation of an inductive inference from statistical models increases its plausibility, such as the redistribution of work in the left bundle branch block to explain the remodelling pattern that predicts response to CRT.[51] Equivalently, data computed from mechanistic models need to be scrutinized quantitatively as it was done in the comparison of clinical and simulation groups to validate a model for acute normovolaemic haemodilution.[66]

An important final remark is that randomized control trials will always be needed to establish evidence that can never be obtained from large observational databases.[67]

Models as critical tools for accelerating regulatory decision-making

Clinical decisions are built on evidence from bench to bedside. Regulatory decisions, on the contrary, are often based on heterogeneous, limited, or completely absent human data, as in the case of approval for first-in-human clinical trials. In this regard, the results of computational models can now be accepted for some regulatory submissions.[68,69] Digital evidence obtained using computer simulations can be used for safety of therapy prior to first-in-human use, or under scenarios not ethically possible in human.[38] Computational models have an increasingly important role in the overall product life cycle management, proving useful in the processes of design optimization for development and testing, supplemental non-clinical testing, and post-market design changes and failure assessment.[27]

The development process for medical devices involves manufacturing and testing samples under a wide range of scenarios, which is often time-consuming and financially overwhelming. Moreover, pre-clinical testing conditions are often very simplified with respect to the actual patient environment. Statistical and mechanistic models synergistically offer to streamline this process, where statistical models can be used to collect a representative virtual patient cohort, and mechanistic models can then be used to simulate the device behaviour under defined scenarios. In this way, new devices can be tested in a representative virtual patient population, thereby decreasing the risk before moving to an actual clinical trial. An example is HEARTguide™ (FEops nv, Belgium), where device–patient interactions after transcatheter aortic valve implantation can be predicted.[25]

The augmentation of clinical trial design with virtual patients is also an evolving idea.[70–72] This would overcome limitations of current empirical trials, where patients burdened with comorbidities or complex treatment regimens are often excluded from the trials, and enrolled individuals are handled under reductionist approaches, assuming they share a common phenotype. Such approaches often fail to capture differences in response to treatment.[70] Alternatively, computational evidence can inform collection of novel evidence from clinical trials,[13,38] where models can improve patient selection by derived biomarkers and predictions. This offers an opportunity to answer questions traditionally restricted by financial or ethical considerations, and to investigate therapy efficacy in more clinically relevant cases. Computational modelling can also facilitate safe methods to explore treatment effects in sub-populations clinically more complex to address, such as patients with rare diseases or paediatric cohorts, and therefore may allow for insights not possible in the current clinical trial practice.

One of the first examples in which digital evidence (i.e. an *in silico* trial) replaced any additional clinical evidence was in the approval of the Advisa MRI SureScan pacemaker (Medtronic, Inc.).[73] Another powerful example is a computer simulator of type 1 diabetes mellitus,[74] which was accepted by the FDA as a substitute to animal trials for the pre-clinical testing of control strategies in artificial pancreas studies. Later, an investigational device exemption (i.e. the approval needed to initiate a clinical study), issued solely on the basis of modelling testing, was granted by the FDA for a closed-loop control clinical trial of the safety and effectiveness of the proposed artificial pancreas algorithm.

In the context of drug safety and efficacy assessment, an unmet need is filling the gaps between animal translation or *in vitro* preparations and prediction of the human response. Mechanistic models may assist in scaling observations into humans.[75] This is, for example, the goal of the CIPA initiative,[69] sponsored by the FDA among others, aiming at facilitating the adoption of a new paradigm for assessment of potential risk of clinical Torsades de Pointes, where mechanistic models of human electrophysiology will play a crucial role. This is reinforced by a recent study in which human *in silico* trials outperformed animal models in predicting clinical pro-arrhythmic cardiotoxicity, so they might be soon integrated into existing drug safety assessment pipelines.[63]

Finally, after a product is launched, mechanistic models can be still used for post-market re-evaluation and failure assessment in order to identify any potential underlying problems. This creates a valuable opportunity for simulations to evaluate any design changes planned for next-generation productions, ultimately closing the product life cycle loop, and demonstrating the ubiquitous presence and utility of statistical and mechanistic models in the future of medical product regulation.

Discussion

The digital twin, i.e., the dynamic integration and augmentation of patient data using mechanistic and statistical models, is the actual

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

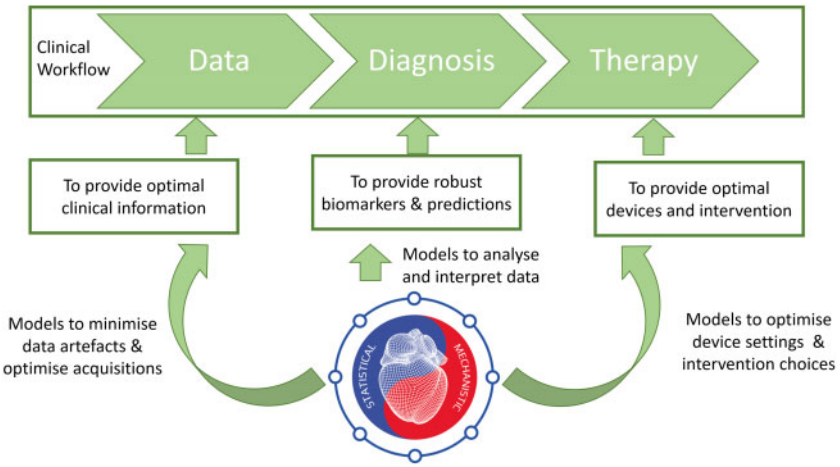


Figure 5 The vision of a personalized *in silico* cardiology, where the digital twin informs all the stages through the clinical workflow. Models are used (i) to optimize data acquisition and the information extracted from it, (ii) to evaluate current health status and inform diagnosis and risk stratification, and (iii) to optimize clinical devices and drug selection to deliver a personalized therapy.

pathway towards the vision of precision medicine. Simple and fragmented components of the digital twin are already used in clinical practice: a decision tree in a clinical guideline encapsulates the best-documented evidence that is based in statistical and mechanistic insights. The digital twin will gradually include tailored computer-enabled decision points, and create the transition from healthcare systems founded on describing disease to healthcare systems focused on predicting response, and thus shifting treatment selection from being based on the state of the patient today to optimizing the state of the patient tomorrow.

Envisioned impact and timeline

The digital twin provides a pathway to map current patient observations into a predictive framework, combining inductive and deductive reasoning. Early components of the digital twin are already making a clinical impact. In a generic clinical workflow divided in the stages of data acquisition, diagnosis, and therapy planning, computational models can provide value in the three stages, see *Figure 5*. To improve data acquisition techniques, there are already statistical models to automate the image analysis tasks.[16] To provide better diagnosis, a virtual fractional flow reserve can replace an invasive catheter,[29,37] or the body surface recordings can be mapped to the surface of the heart.[39] With regards to therapy planning, a virtual deployment of the valve replacement[25,76] or a roadmap to guide ablation procedures[77,78] represents existing techniques (statistical and mechanistic) that have been implemented into the clinical workflow. These solutions have thus met regulatory approval, where they are referred to as 'software as a medical device', and where guidelines from the International Medical Device Regulators Forum are accepted by the EU and the USA.

A digital twin will follow the life journey of each person and harness both data collected by wearable sensors and lifestyle information that patients may register, shifting the clinical approach towards preventive healthcare. A notable challenge is the integration of these data with healthcare organizations, where security and confidentiality of the sensitive information remain paramount.

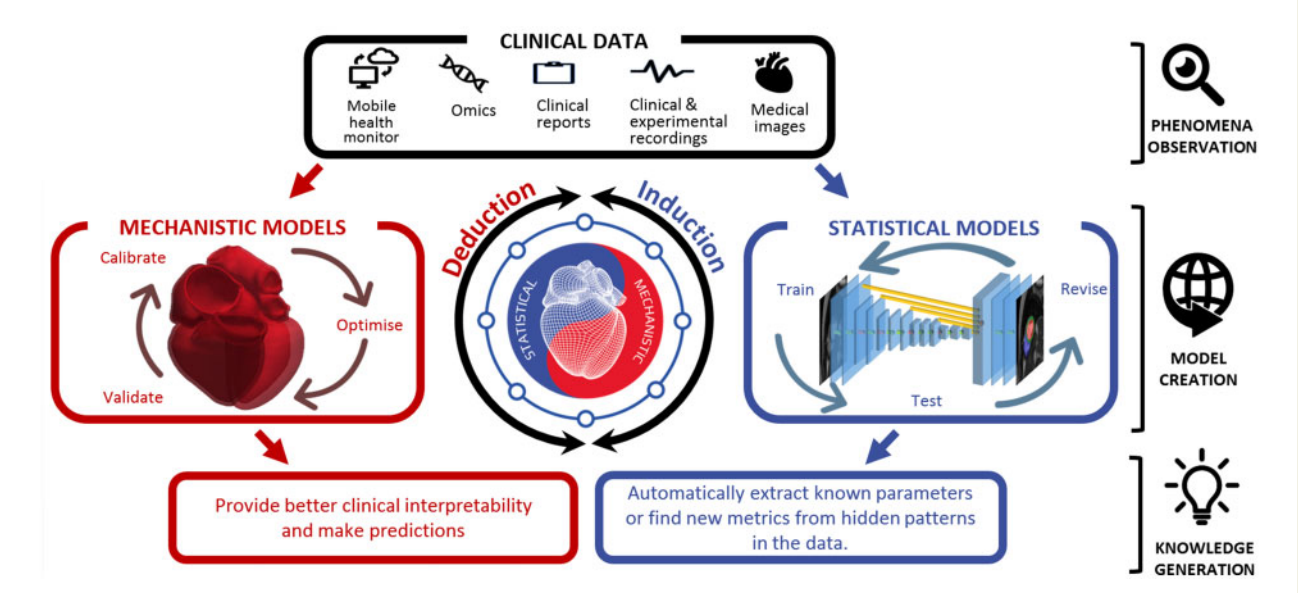
The currently still fragmented and incipient concept of the digital twin will be gradually crystallized and adopted during the next 5–10 years. The holistic integration of a Digital Twin is the aspiration that will be reached through two complementary and synergetic pathways: the first is the refinement of key decision points in the management of cardiac disease, driven by personalized mechanistic models that are informed by key pieces of patient's data; and the second is the disease-centred optimization of the patient's lifetime journey through the healthcare system, driven by statistical models being informed by the electronic health record of a large population.

On the actual implementation of the digital twin, we envision that the evolution will be towards a gradually better inter-operability of current health information systems, leading to a distributed location of the information. Digital twin users will mainly be citizens and physicians, with different interfaces that retrieve the relevant data and trigger the analysis capabilities hosted in the local device or remote cloud resources. The analysis may also require specialized skills that may be delivered by industry, or even by *computational cardiologists* inside healthcare organizations.

Organizational and societal challenges ahead

Access to data is the main challenge in both the development and the clinical translation of the digital twin, caused by infrastructural, regulatory, and societal reasons. Information systems and electronic health records are fragmented, highly heterogeneous and difficult to interoperate. Information is often contained in unstructured format, and its extraction requires either manual work or further research efforts of automation through natural language processing technologies.[79] Simulations may also require specialized skills and supercomputers.

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021



Take home figure The cardiovascular digital twin that will deliver the vision of precision medicine by the synergetic combination of computer-enhanced induction (using statistical models learnt from data) and deduction (mechanistic modelling and simulation integrating multi-scale knowledge).

In this context, provision of digital twin technologies may be enabled by cloud infrastructures (e.g. HeartFlow FFR_{CT} Analysis).

Consent and confidentiality are key ingredients to address the societal concerns when handling the personal data needed to develop and validate digital twin technologies. The EU General Data Protection Regulation (GDPR) has imposed new legal requirements, such as the right to withdraw consent and the right to be forgotten, causing controversy about the cost and feasibility of its enforcement.[80] Any digital twin solution that holds enough information to identify a patient needs to carefully watch these requirements, that also apply to retrospective data and safety backups.

Potential professional, cultural, and ethical issues

As more clinical tasks are performed by models, the fear of replacement of physicians by machines may arise. In some scenarios, machines may match or even outperform physicians.[81] In other scenarios, human experts, by not practising on the easy problems solved by the machine, may lose the skills that may still be needed when dealing with difficult cases.

The second professional barrier is the mistrust that originates from a 'black box', where predictions derived by algorithms are not matched with a plausible explanation. Generation of evidence is one clear way to generate trust. Another solution is to use methods to illustrate the logic inside the box, including clustering and association techniques,[82] which may help to identify the causes and mechanisms.

From the patient's perspective, personalization creates the opportunity of more involvement in healthcare decisions. Patients will be empowered to better manage their disease using the digital twin to gain information about their current and predicted state, and potentially to adopt optimized lifestyle suggestions. A well-informed patient shall have more efficient discussions with physicians, and consent and decide faster on diagnostic or treatment procedures.

Finally, on the ethical side, there is a risk of models to create or exacerbate existing racial or societal biases in healthcare systems: if a group is misrepresented in the data used to train models, that group may receive a sub-optimal treatment.[83]

Recommendations

The pathway to accelerate the clinical impact with digital twin technologies is to generate trust among researchers, clinicians, and society.

Research communities shall avoid inflating expectations. Claims about generality and potential impact should be based on rigorous methodology, with external cohorts to demonstrate the validity of inferences, and with the quantification of the uncertainty of predictions.[84] Any model is a simplified representation of the reality, with a limited scope and dependence on assumptions made. The opportunity is an adequate handling of these limitations, with models able to identify data inconsistencies, and with data used to constrain and verify the model assumptions.[85]

As an emerging field, the digital twin needs guidelines, gold-standards, and benchmark tests.[86,87] Scientific organizations and regulatory bodies have released guidelines that can be used to establish the level of rigour needed for computational modelling.[27] Such guidelines and standards are useful tools as they allow regulators to judge computational evidence and industry to understand regulatory requirements for computational models, leveraging a substantial part of the risk and uncertainty associated to the development of these new technologies. They can even increase and facilitate their translational impact, as the quality and robustness of the models and their reporting will

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

increase by adhering to such guidance during model development. Further effort is needed to widen the scope of these first multi-stakeholder consensuses involving industry, academia, and regulators. Current initiatives that develop visions, technologies, or infrastructure relevant to the 'digital twin' community are Elixir (https://elixir-europe.org/), FAIRDOM (https://fair-dom.org/), and EOSC (https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud).

The education of citizens, care providers, physicians, and researchers in the uses and possibilities of digital twin technologies is key for its adoption and acceptance. University education systems should also allow for the exchange of knowledge at the earliest stages of the career: medical students should have some computational training, just as engineers in biomedical industry should be trained in cardiology during their studies.[88] And postgraduate training programmes should bridge remaining cultural and language gaps between disciplines, such as our Personalised In-silico Cardiology EU funded Innovative Training Network (https://picnet.eu).

Conclusion

Precision cardiology will be delivered, not only by data, but also by the inductive and deductive reasoning built in the digital twin of each patient. Treatment and prevention of cardiovascular disease will be based on accurate predictions of both the underlying causes of disease and the pathways to sustain or restore health. These predictions will be provided and validated by the synergistic interplay between mechanistic and statistical models. The early steps towards this vision have been taken, and the next ones depend on the coordinated drive from scientific, clinical, industrial, and regulatory stakeholders in order to build the evidence and tackle the organizational and societal challenges ahead.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Conflict of interest: none declared.

Funding

This work was supported by the EU's Horizon 2020 Marie Sklodowska-Curie ITN Projects (g.a. 764738 and 766082), the EU's Horizon 2020 research and innovation programme (g.a. 675451 and 823712), the Wellcome/EPSRC Centre for Medical Engineering (WT 203148/Z/16/Z), the National Research Agency (ANR) (g.a. ANR-10-IAHU-04), the NC3RS (NC/P001076/1) and the British Heart Foundation (RE/13/2/30182, RE/13/1/30181, TG/17/3/33406, PG/16/75/32383, FS/17/22/32644, CH/16/3/21406, RG/16/14/32397). E.Pueyo holds an ERC Starting Grant (g.a. 638284). B. Rodriguez and P.Lamata hold Wellcome Trust Senior Research Fellowships (214290/Z/18/Z, 209450/Z/17/Z).

References

1. Antman EM, Loscalzo J. Precision medicine in cardiology. *Nat Rev Cardiol* 2016;**13**:591–602.
2. Trayanova N. From genetics to smart watches: developments in precision cardiology. *Nat Rev Cardiol* 2019;**16**:72–73.
3. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;**372**:793–795.
4. Joyner MJ, Paneth N. Promises, promises, and precision medicine. *J Clin Invest* 2019;**129**:946–948.
5. Khoury MJ. Precision medicine vs preventive medicine. *JAMA* 2019;**321**:406.
6. Noble D. Evolution beyond neo-Darwinism: a new conceptual framework. *J Exp Biol* 2015;**218**:7–13.
7. Alber M, Buganza Tepole A, Cannon WR, De S, Dura-Bernal S, Garikipati K, Karniadakis G, Lytton WW, Perdikaris P, Petzold L, Kuhl E. Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *NPJ Digit Med* 2019;**2**:115.
8. Tao F, Cheng J, Qi Q, Zhang M, Zhang H, Sui F. Digital twin-driven product design, manufacturing and service with big data. *Int J Adv Manuf Technol* 2018;**94**:3563–3576.
9. Lamata P. Teaching cardiovascular medicine to machines. *Cardiovasc Res* 2018;**114**:e62–e64.
10. Niederer SA, Plank G, Chinchapatnam P, Ginks M, Lamata P, Rhode KS, Rinaldi CA, Razavi R, Smith NP. Length-dependent tension in the failing heart and the efficacy of cardiac resynchronization therapy. *Cardiovasc Res* 2011;**89**:336–343.
11. Lumens J, Tayal B, Walmsley J, Delgado-Montero A, Huntjens PR, Schwartzman D, Althouse AD, Delhaas T, Prinzen FW, Gorcsan J. Differentiating electromechanical from non-electrical substrates of mechanical discoordination to identify responders to cardiac resynchronization therapy. *Circ Cardiovasc Imaging* 2015;**8**:e003744.
12. Corral Acero J, Zacur E, Xu H, Ariga R, Bueno-Orovio A, Lamata P, Grau V. *SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI*. FIMH 2019: Functional Imaging and Modeling of the Heart - pp. 361–369. doi: 10.1007/978-3-030-21949-9_39.
13. Cikes M, Sanchez Martinez S, Claggett B, Solomon SD, Bijnens B. Machine-learning integration of complex echocardiographic patterns and clinical parameters from cohorts and trials. *Eur Heart J* 2019;**40**: doi: 10.1093/eurheartj/ehz745.0147.
14. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? *Heart* 2018;**104**:1156–1164.
15. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. *Nat Rev Cardiol* 2016;**13**:350–359.
16. Dey D, Slomka PJ, Leeson P, Comaniciu D, Shrestha S, Sengupta PP, Marwick TH. Artificial intelligence in cardiovascular imaging: JACC state-of-the-art review. *J Am Coll Cardiol* 2019;**73**:1317–1335.
17. Niederer SA, Lumens J, Trayanova NA. Computational models in cardiology. *Nat Rev Cardiol* 2019;**16**:100–111.
18. Johnson KW, Shameer K, Glicksberg BS, Readhead B, Sengupta PP, Björkegren JLM, Kovacic JC, Dudley JT. Enabling precision cardiology through multiscale biology and systems medicine. *JACC Basic Transl Sci* 2017;**2**:311–327.
19. Davies MR, Wang K, Mirams GR, Caruso A, Noble D, Walz A, Lavé T, Schuler F, Singer T, Polonchuk L. Recent developments in using mechanistic cardiac modelling for drug safety evaluation. *Drug Discov Today* 2016;**21**:924–938.
20. Tung L. A bi-domain model for describing ischemic myocardial D-C potentials. 1978. https://dspace.mit.edu/handle/1721.1/16177 (29 February 2020).
21. Sherwin SJ, Formaggia L, Peiró J, Franke V. Computational modelling of 1D blood flow with variable mechanical properties and its application to the simulation of wave propagation in the human arterial system. *Int J Numer Methods Fluids* 2003;**43**:673–700.
22. Guidi G, Pettenati MC, Miniati R, Iadanza E. Random forest for automatic assessment of heart failure severity in a telemonitoring scenario. In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 3230–3233.
23. Stegle O, Fallert SV, MacKay DJ, Brage S. Gaussian process robust regression for noisy heart rate data. *IEEE Trans Biomed Eng* 2008;**55**:2143–2151.
24. Roney CH, Bayer JD, Cochet H, Meo M, Dubois R, Jaïs P, Vigmond EJ. Variability in pulmonary vein electrophysiology and fibrosis determines arrhythmia susceptibility and dynamics. *PLoS Comput Biol* 2018;**14**:e1006166.
25. de Jaegere P, De Santis G, Rodriguez-Olivares R, Bosmans J, Bruining N, Dezutter T, Rahhab Z, El Faquir N, Collas V, Bosmans B, Verheghe B, Ren C, Geleinse M, Schultz C, van Mieghem N, De Beule M, Mortier P. Patient-specific computer modeling to predict aortic regurgitation after transcatheter aortic valve replacement. *JACC Cardiovasc Interv* 2016;**9**:508–512.
26. Gray RA, Pathmanathan P. Patient-specific cardiovascular computational modeling: diversity of personalization and challenges. *J Cardiovasc Transl Res* 2018;**11**:80–88.
27. Morrison TM, Dreher ML, Nagaraja S, Angelone LM, Kainz W. The role of computational modeling and simulation in the total product life cycle of peripheral vascular devices. *J Med Device* 2017;**11**:024503.
28. Dillon-Murphy D, Noorani A, Nordsletten D, Figueroa CA. Multi-modality image-based computational analysis of haemodynamics in aortic dissection. *Biomech Model Mechanobiol* 2016;**15**:857–876.
29. Morris PD, van de Vosse FN, Lawford PV, Hose DR, Gunn JP. "Virtual" (computed) fractional flow reserve. *JACC Cardiovasc Interv* 2015;**8**:1009–1017.

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

30. Xi J, Lamata P, Niederer S, Land S, Shi W, Zhuang X, Ourselin S, Duckett SG, Shetty AK, Rinaldi CA, Rueckert D, Razavi R, Smith NP. The estimation of patient-specific cardiac diastolic functions from clinical measurements. *Med Image Anal* 2013;**17**:133–146.
31. Wang ZJ, Wang VY, Bradley CP, Nash MP, Young AA, Cao JJ. Left ventricular diastolic myocardial stiffness and end-diastolic myofibre stress in human heart failure using personalised biomechanical analysis. *J Cardiovasc Transl Res* 2018;**11**: 346–356.
32. Krittian SBS, Lamata P, Michler C, Nordsletten DA, Bock J, Bradley CP, Pitcher A, Kilner PJ, Markl M, Smith NP. A finite-element approach to the direct computation of relative cardiovascular pressure from time-resolved MR velocity data. *Med Image Anal* 2012;**16**:1029–1037.
33. Donati F, Figueroa CA, Smith NP, Lamata P, Nordsletten DA. Non-invasive pressure difference estimation from PC-MRI using the work-energy equation. *Med Image Anal* 2015;**26**:159–172.
34. Donati F, Myerson S, Bissell MM, Smith NP, Neubauer S, Monaghan MJ, Nordsletten DA, Lamata P. Beyond Bernoulli: improving the accuracy and precision of non-invasive estimation of peak pressure drops. *Circ Cardiovasc Imaging* 2017;**10**:e005207.
35. Nørgaard BL, Leipsic J, Gaur S, Seneviratne S, Ko BS, Ito H, Jensen JM, Mauri L, De Bruyne B, Bezerra H, Osawa K, Marwan M, Naber C, Erglis A, Park S-J, Christiansen EH, Kaltoft A, Lassen JF, Bøtker HE, Achenbach S. Diagnostic performance of noninvasive fractional flow reserve derived from coronary computed tomography angiography in suspected coronary artery disease. *J Am Coll Cardiol* 2014;**63**:1145–1155.
36. Min JK, Taylor CA, Achenbach S, Koo BK, Leipsic J, Nørgaard BL, Pijls NJ, De Bruyne B. Noninvasive fractional flow reserve derived from coronary CT angiography. *JACC Cardiovasc Imaging* 2015;**8**:1209–1222.
37. Rajani R, Modi B, Ntalas I, Curzen N. Non-invasive fractional flow reserve using computed tomographic angiography: where are we now and where are we going? *Heart* 2017;**103**:1216–1222.
38. Morrison TM, Pathmanathan P, Adwan M, Margerrison E. Advancing regulatory science with computational modeling for medical devices at the FDA's Office of Science and Engineering Laboratories. *Front Med* 2018;**5**: doi: 10.3389/fmed.2018.00241.
39. Haissaguerre M, Hocini M, Shah AJ, Derval N, Sacher F, Jais P, Dubois R. Noninvasive panoramic mapping of human atrial fibrillation mechanisms: a feasibility report noninvasive panoramic mapping of human atrial fibrillation mechanisms. Introduction. *J Cardiovasc Electrophysiol* 2013;**24**:711–717.
40. Daubert C, Behar N, Martins RP, Mabo P, Leclercq C,. Avoiding non-responders to cardiac resynchronization therapy: a practical guide. *Eur Heart J* 2016;**38**: 1463–1472.
41. Prakosa A, Arevalo HJ, Deng D, Boyle PM, Nikolov PP, Ashikaga H, Blauer JJE, Ghafoori E, Park CJ, Blake RC, Han FT, MacLeod RS, Halperin HR, Callans DJ, Ranjan R, Chrispin J, Nazarian S, Trayanova NA. Personalized virtual-heart technology for guiding the ablation of infarct-related ventricular tachycardia. *Nat Biomed Eng* 2018;**2**:732–740.
42. Hill YR, Child N, Hanson B, Wallman M, Coronel R, Plank G, Rinaldi CA, Gill J, Smith NP, Taggart P, Bishop MJ. Investigating a novel activation-repolarisation time metric to predict localised Vulnerability to reentry using computational modelling. *PLoS One* 2016;**11**: doi: 10.1371/journal.pone.0149342.
43. Alfonso F, Nihoyannopoulos P, Stewart J, Dickie S, Lemery R, McKenna WJ. Clinical significance of giant negative T waves in hypertrophic cardiomyopathy. *J Am Coll Cardiol* 1990;**15**:965–971.
44. Pelliccia A, Di Paolo FM, Quattrini FM, Basso C, Culasso F, Popoli G, De Luca R, Spataro A, Biffi A, Thiene G, Maron BJ. Outcomes in athletes with marked ECG repolarization abnormalities. *N Engl J Med* 2008;358:152–161.
45. Pelliccia A, Corrado D, Bjørnstad HH, Panhuyzen-Goedkoop N, Urhausen A, Carre F, Anastasakis A, Vanhees L, Arbustini E, Priori S. Recommendations for participation in competitive sport and leisure-time physical activity in individuals with cardiomyopathies, myocarditis and pericarditis. *Eur J Prev Cardiol* 2006;**13**: 876–885.
46. Passini E, Mincholé A, Coppini R, Cerbai E, Rodriguez B, Severi S, Bueno-Orovio A. Mechanisms of pro-arrhythmic abnormalities in ventricular repolarisation and anti-arrhythmic therapies in human hypertrophic cardiomyopathy. *J Mol Cell Cardiol* 2016;**96**:72–81.
47. Lyon A, Bueno-Orovio A, Zacur E, Ariga R, Grau V, Neubauer S, Watkins H, Rodriguez B, Mincholé A. Electrocardiogram phenotypes in hypertrophic cardiomyopathy caused by distinct mechanisms: apico-basal repolarization gradients vs. Purkinje-myocardial coupling abnormalities. *Europace* 2018;**20**: III102–III112.
48. Lyon A, Ariga R, Mincholé A, Mahmod M, Ormondroyd E, Laguna P, de Freitas N, Neubauer S, Watkins H, Rodriguez B. Distinct ECG phenotypes identified in hypertrophic cardiomyopathy using machine learning associate with arrhythmic risk markers. *Front Physiol* 2018;**9**:213.
49. Arevalo HJ, Vadakkumpadan F, Guallar E, Jebb A, Malamas P, Wu KC, Trayanova NA. Arrhythmia risk stratification of patients after myocardial infarction using personalized heart models. *Nat Commun* 2016;**7**: doi: 10.1038/ncomms11437.
50. Dawes TJW, de Marvao A, Shi W, Fletcher T, Watson GMJ, Wharton J, Rhodes CJ, Howard LSGE, Gibbs JSR, Rueckert D, Cook SA, Wilkins MR, O'Regan DP. Machine learning of three-dimensional right ventricular motion enables outcome prediction in pulmonary hypertension: a cardiac MR imaging study. *Radiology* 2017;**283**:381–390.
51. Warriner DR, Jackson T, Zacur E, Sammut E, Sheridan P, Hose DR, Lawford P, Razavi R, Niederer SA, Rinaldi CA, Lamata P. An asymmetric wall-thickening pattern predicts response to cardiac resynchronization therapy. *JACC Cardiovasc Imaging* 2018;**11**:1545–1546.
52. Mirams GR, Pathmanathan P, Gray RA, Challenor P, Clayton RH. Uncertainty and variability in computational and mathematical models of cardiac physiology. *J Physiol* 2016;**594**:6833–6847.
53. Pathmanathan P, Gray RA. Ensuring reliability of safety-critical clinical applications of computational cardiac models. *Front Physiol* 2013;**4**:358.
54. Winslow RL, Trayanova N, Geman D, Miller MI. Computational medicine: translating models to clinical care. *Sci Transl Med* 2012;**4**:158rv11–158rv11.
55. Kurokawa J, Kodama M, Clancy CE, Furukawa T. Sex hormonal regulation of cardiac ion channels in drug-induced QT syndromes. *Pharmacol Ther* 2016;**168**: 23–28.
56. Niemeijer MN, van den Berg ME, Deckers JW, Aarnoudse ALHJ, Hofman A, Franco OH, Uitterlinden AG, Rijnbeek PR, Eijgelsheim M, Stricker BH. ABCB1 gene variants, digoxin and risk of sudden cardiac death in a general population. *Heart* 2015;**101**:1973–1979.
57. Kim DW, Jang HY, Kim KW, Shin Y, Park SH. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. *Korean J Radiol* 2019;**20**:405.
58. Chang KC, Dutta S, Mirams GR, Beattie KA, Sheng J, Tran PN, Wu M, Wu WW, Colatsky T, Strauss DG, Li Z. Uncertainty quantification reveals the importance of data variability and experimental design considerations for in silico proarrhythmia risk assessment. *Front Physiol* 2017;**8**:917.
59. Dey P, Ross JS, Ritchie JD, Desai NR, Bhavnani SP, Krumholz HM. Data sharing and cardiology. *J Am Coll Cardiol* 2017;**70**:3018–3025.
60. Schiltz M. Science without publication paywalls: cOAlition S for the realisation of full and immediate open access. *PLoS Med* 2018;**15**:e1002663.
61. Britton OJ, Bueno-Orovio A, Van Ammel K, Lu HR, Towart R, Gallacher DJ, Rodriguez B. Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proc Natl Acad Sci USA* 2013;**110**:E2098–E2105.
62. Sánchez C, Bueno-Orovio A, Wettwer E, Loose S, Simon J, Ravens U, Pueyo E, Rodriguez B. Inter-subject variability in human atrial action potential in sinus rhythm versus chronic atrial fibrillation. *PLoS One* 2014;**9**:e105897.
63. Passini E, Britton OJ, Lu HR, Rohrbacher J, Hermans AN, Gallacher DJ, Greig RJH, Bueno-Orovio A, Rodriguez. Human in silico drug trials demonstrate higher accuracy than animal models in predicting clinical pro-arrhythmic cardiotoxicity. *Front Physiol* 2017;**8**: doi: 10.3389/fphys.2017.00668.
64. Sánchez C, Bueno-Orovio A, Pueyo E, Rodríguez B. Atrial fibrillation dynamics and ionic block effects in six heterogeneous human 3D virtual atria with distinct repolarization dynamics. *Front Bioeng Biotechnol* 2017;**5**:29.
65. Liang L, Liu M, Martin C, Elefteriades JA, Sun W. A machine learning approach to investigate the relationship between shape features and numerically predicted risk of ascending aortic aneurysm. *Biomech Model Mechanobiol* 2017;**16**: 1519–1533.
66. Sims CR, Delima LR, Calimaran A, Hester R, Pruett WA. Validating the physiologic model HumMod as a substitute for clinical trials involving acute normovolemic hemodilution. *Anesth Analg* 2018;**126**:93–101.
67. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available: table 1. *Am J Epidemiol* 2016;**183**:758–764.
68. Colatsky T, Fermini B, Gintant G, Pierson JB, Sager P, Sekino Y, Strauss DG, Stockbridge N. The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative—update on progress. *J Pharmacol Toxicol Methods* 2016;**81**:15–20.
69. Cavero I, Holzgrefe H. CiPA: ongoing testing, future qualification procedures, and pending issues. *J Pharmacol Toxicol Methods* 2015;**76**:27–37.
70. Pappalardo F, Russo G, Tshinanu FM, Viceconti M. In silico clinical trials: concepts and early adoptions. *Brief Bioinform* 2019;**20**:1699–1708.
71. Viceconti M, Henney A, Morley-Fletcher E. In silico clinical trials: how computer simulation will transform the biomedical industry. *Int J Clin Trials* 2016; **3**:37.
72. Haddad T, Himes A, Thompson L, Irony T, Nair R; MDIC Computer Modeling and Simulation Working Group Participants. Incorporation of stochastic engineering models as prior information in Bayesian medical device trials. *J Biopharm Stat* 2017;**27**:1089–1103.

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

73. Faris O, Shuren J. An FDA viewpoint on unique considerations for medical-device clinical trials. *N Engl J Med* 2017;**376**:1350–1357.
74. Kovatchev BP, Breton M, Man CD, Cobelli C. In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes. *J Diabetes Sci Technol* 2009;**3**: 44–55.
75. Zemzemi N, Bernabeu MO, Saiz J, Cooper J, Pathmanathan P, Mirams GR, Pitt-Francis J, Rodriguez B. Computational assessment of drug-induced effects on the electrocardiogram: from ion channel to body surface potentials. *Br J Pharmacol* 2013;**168**:718–733.
76. Rocatello G, El Faquir N, De Santis G, Iannaccone F, Bosmans J, De Backer O, Sondergaard L, Segers P, De Beule M, de Jaegere P, Mortier P. Patient-specific computer simulation to elucidate the role of contact pressure in the development of new conduction abnormalities after catheter-based implantation of a self-expanding aortic valve. *Circ Cardiovasc Interv* 2018;**11**:e005344.
77. Andreu D, Ortiz-Pérez JT, Fernández-Armenta J, Guiu E, Acosta J, Prat-González S, De Caralt TM, Perea RJ, Garrido C, Mont L, Brugada J, Berruezo A. 3D delayed-enhanced magnetic resonance sequences improve conducting channel delineation prior to ventricular tachycardia ablation. *Europace* 2015;**17**:938–945.
78. Cedilnik N, Duchateau J, Dubois R, Sacher F, Jaïs P, Cochet H,Sermesant M. Fast personalized electrophysiological models from computed tomography images for ventricular tachycardia ablation planning. *Europace* 2018;**20**:iii94–iii101.
79. Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, Forshee R, Walderhaug M, Botsis T. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J Biomed Inform* 2017;**73**:14–29.
80. Politou E, Alepis E, Patsakis C. Forgetting personal data and revoking consent under the GDPR: challenges and proposed solutions. *J Cybersecurity* 2018;**4**: doi: 10.1093/cybsec/tyy001.
81. Darcy AM, Louie AK, Roberts LW. Machine learning and the profession of medicine. *JAMA* 2016;**315**:551.
82. Esfandiari N, Babavalian MR, Moghadam A-ME, Tabar VK. Expert systems with applications knowledge discovery in medicine: current issue and future trend. *Expert Syst Appl* 2014;**41**:4434–4463.
83. Nordling L. A fairer way forward for AI in health care. *Nature* 2019;**573**: S103–S105.
84. Pathmanathan P, Cordeiro JM, Gray RA. Comprehensive uncertainty quantification and sensitivity analysis for cardiac action potential models. *Front Physiol* 2019; **10**:721.
85. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep* 2016;**6**:26094.
86. Land S, Gurev V, Arens S, Augustin CM, Baron L, Blake R, Bradley C, Castro S, Crozier A, Favino M, Fastl TE, Fritz T, Gao H, Gizzi A, Griffith BE, Hurtado DE, Krause R, Luo X, Nash MP, Pezzuto S, Plank G, Rossi S, Ruprecht D, Seemann G, Smith NP, Sundnes J, Rice JJ, Trayanova N, Wang D, Jenny Wang Z, Niederer SA. Verification of cardiac mechanics software: benchmark problems and solutions for testing active and passive material behaviour. *Proc R Soc A Math Phys Eng Sci* 2015;**471**:20150641.
87. Niederer SA, Kerfoot E, Benson AP, Bernabeu MO, Bernus O, Bradley C, Cherry EM, Clayton R, Fenton FH, Garny A, Heidenreich E, Land S, Maleckar M, Pathmanathan P, Plank G, Rodríguez JF, Roy I, Sachse FB, Seemann G, Skavhaug O, Smith NP. Verification of cardiac tissue electrophysiology simulators using an N-version benchmark. *Philos Trans R Soc A Math Phys Eng Sci* 2011;**369**: 4331–4351.
88. Eden C, Johnson KW, Gottesman O, Bottinger EP, Abul-Husn NS. Medical student preparedness for an era of personalized medicine: findings from one US medical school. *Per Med* 2016;**13**:129–141.

Downloaded from https://academic.oup.com/eurheartj/article/41/48/4556/5775673 by Bodleian Law Library user on 12 July 2021

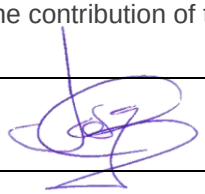
Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	The 'Digital Twin' to enable the vision of precision cardiology
Publication Status	☑ Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	Jorge Corral-Acero, Francesca Margara, Maciej Marciniak, Cristobal Rodero, Filip Loncaric, Yingjing Feng, Andrew Gilbert, Joao F Fernandes, Hassaan A Bukhari, Ali Wajdan, Manuel Villegas Martinez, Mariana Sousa Santos, Mehrdad Shamohammdi, Hongxing Luo, Philip Westphal, Paul Leeson, Paolo DiAchille, Viatcheslav Gurev, Manuel Mayr, Liesbet Geris, Pras Pathmanathan, Tina Morrison, Richard Cornelussen, Frits Prinzen, Tammo Delhaas, Ada Doltra, Marta Sitges, Edward J Vigmond, Ernesto Zacur, Vicente Grau, Blanca Rodriguez, Espen W Remme, Steven Niederer, Peter Mortier, Kristin McLeod, Mark Potse, Esther Pueyo, Alfonso Bueno-Orovio, Pablo Lamata, "The 'Digital Twin' to enable the vision of precision cardiology", *European Heart Journal*, Volume 41, Issue 48, 21 December 2020, Pages 4556–4564, DOI: https://doi.org/10.1093/eurheartj/ehaa159

Student Confirmation

Student Name:	Jorge Corral Acero		
Contribution to the Paper	(1) Co-led the literature review with my fellow Francesca Margara on mechanistic and statistical modelling synergy towards maximising value of available data, building knowledge, accelerating regulatory decision-making and improving clinical decision; (2) was one of the main contributors to the introduction and discussion sections; (3) was one of the main contributors to writing and reviewing the manuscript; (4) led in figure preparation and prepared several figures; and (5) organised and led internal meetings and discussions. The contribution of the 14 PIC Fellows co-authors of this paper shall be acknowledged.		
Signature		Date	12/07/2021

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Vicente Grau		
Supervisor comments: I confirm that the above statement is correct.		
Signature 	Date	12/07/2021

This completed form should be included in the thesis, at the end of the relevant chapter.

CHAPTER 4

Accurate and Fully Automated Estimation of the 3D Shape of the Heart from 2D CMR

Research Objective 1

4. Accurate and Fully Automated Estimation of the 3D Shape of the Heart from 2D CMR

Research Objective 1

4.1. Chapter Publications

- **Paper II**: J. Corral Acero et al. "SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI". *FIMH 2019*. Lecture Notes in Computer Science, vol 11504. Springer, Cham; 2019. p361-69.

- **Paper III**: J. Corral Acero et al. "Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation". *STACOM 2019*. Lecture Notes in Computer Science, vol 12009. Springer, Cham. 2019. p384–94.

- **Paper IV**: J. Corral Acero et al. "A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation". *STACOM 2020*. Lecture Notes in Computer Science. Springer, Cham; 2021. p. 196–207.

4.2. Chapter Outline

CMR imaging is the gold standard for the accurate non-invasive quantification of cardiac function and structure [18], [19], [156]. Its prognostic value in CVD risk management has been repeatedly demonstrated in numerous trials [7]–[9], [152], [199], [363]. As explained in Chapter 2, segmentation is a standard pre-requisite for both the calculation of image biomarkers and to conduct shape analysis. Substantial progress has been made towards automation thanks to the DL advances discussed before (Section 2.4.2). However, CMR analysis mostly remains manual in clinical practice, with the incurred time burden and associated costs [182]–[185]. The main challenges limiting the automation of these tasks include: (a) the anatomical variability of the heart, especially in pathological scenarios; (b) the intrinsic complexity of cardiac dynamics; and (c) the heterogeneity in imaging acquisitions, introduced by different scanners, centres, and protocols [156]. Moreover, DL approaches are

constrained by the limited availability of clinical data [279]. This chapter aims to overcome some of these limitations in order to build a fully automated pipeline to accurately segment 2D CMR images and obtain the personalised 3D shape of a patient's heart.

Paper II, awarded as best imaging paper at FIMH19, introduces a novel augmentation methodology that tackles the challenges of data availability and anatomical variability. The proposed method learns the variability of the heart in order to produce anatomically meaningful deformations to augment the original instances. This anatomical variability is learnt via statistical models of deformation, and hence the method is called SMOD. The underlaying hypothesis is that, by populating the space of anatomical variability, the segmentation NN will converge to a better generalisation during training. This will translate into a performance improvement in the test set. The method is validated against standard augmentation, based on cropping, flipping, translation, and rotation. SMOD reports a quantitative improvement and proves particularly useful when the available training data is critically limited, or the structures to segment are highly variable. This is highlighted in a test study on HCM, as a landmark cardiovascular disease in terms of anatomical variability [136].

This segmentation work is further extended in Paper III, which presents a 2-step DL approach to automatically characterise the LV and calculate the relevant regional metrics. The first NN detects where the heart is located. Based on its output, the cardiac images are pre-processed, that is, oriented, cropped to the ROI, and normalised in intensities within this ROI. This pre-processing step has a double positive effect. Firstly, it leads to smaller images where labels (background vs heart) are more balanced. This simplifies the task of the second NN, and contributes to an improvement in performance. Secondly, it orients the CMR slices to a canonical position, so that LV metrics can be directly regressed from the segmentation of the second NN. The work includes a complete assessment of the impact of the SMOD hyperparameters and the architecture of the second NN on the overall performance. The resulting pipeline was awarded the 1st price at the LV Quantification Challenge (STACOM and MICCAI19), and achieved the highest scores in all its categories [271]. My colleague

Hao Xu shall be acknowledged for co-implementing the VGG architecture, his help in writing-up the manuscript, and his support during the challenge.

The final publication of this chapter, Paper IV, aims to resolve imaging heterogeneity towards multi-centre and multi-vendor applications. To this aim, two domain adaptation techniques (domain adversarial training and domain unlearning) were incorporated into the 2-step segmentation pipeline, in collaboration with Prof Mark Jenkinson's group (Welcome Centre for Integrative Neuroimaging, Oxford Centre for Functional MRI of the Brain, Nuffield Department of Clinical Neurosciences, University of Oxford). The work was benchmarked at the Multi-Centre, Multi-Vendor, Multi-Disease Cardiac Image Segmentation Challenge [272], [364], organised within STACOM 2020. While the work did not win the challenge, there were not statically significant differences between our results and the 1st ranked algorithm. Special credits shall be given to Vaanathi Sundaresan for her help in running the domain adaption approaches, and for building the containers for participating in the challenge. The resulting segmentation pipeline is made publicly available in [291].

To sum up, these three works taken together complement each other and explicitly address the major challenges towards automated segmentation mentioned above. While Paper II contributes to overcome the anatomical variability and the limited training data availability, Paper IV addresses the variability in appearance introduced by different scanners and protocols. Finally, Paper III introduces a 2-step process that eases the fine segmentation by improving label balance and removing variability in heart orientation and position.

Altogether, this results in accurate 2D segmentations of the heart, from which personalised 3D meshes to each patient's heart can be extracted. The construction of the 3D meshes exploited in this Thesis is based on a cubic Hermite variational mesh warping technique [365]. In essence, this technique combines image registration and mesh projection to fit an idealised template of the LV to the myocardial segmentation, arranged in 3D [365]. The proven accuracy, its robustness to SAx slice misalignment and segmentation issues, and the uptake of this meshing methodology in different

clinical studies argue in favour to its adoption [366]–[368]. Moreover, this model-based approach directly provides mesh point correspondence, which facilitates the shape analysis, and further justifies its use in this Thesis. The shape reconstruction along with the shape analysis is detailed in the next chapter (see Paper V).

The AMI case study, presented next in Chapters 5 and 6, is limited to LV structure (i.e. simple and symmetrical circular shapes), rich in training data (100+ full cycle ground truth segmentations) and presents no unseen domains. In such scenario, the implementations presented in Papers II and IV are deemed insignificant and, thus, the segmentation pipeline deployed in Paper III (and not the combination of the three works) is considered for simplicity.

Paper II

SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI

Jorge Corral Acero, Ernesto Zacur, Hao Xu, Rina Ariga, Alfonso Bueno-Orovio, Pablo Lamata, Vicente Grau

In: Coudière Y., Ozenne V., Vigmond E., Zemzemi N. (eds) Functional Imaging and Modeling of the Heart. FIMH 2019. Lecture Notes in Computer Science, vol 11504. Springer, Cham (https://doi.org/10.1007/978-3-030-21949-9_39)



SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI

Jorge Corral Acero[1(✉)], Ernesto Zacur[1], Hao Xu[1], Rina Ariga[2], Alfonso Bueno-Orovio[3], Pablo Lamata[4], and Vicente Grau[1]

[1] Department of Engineering Science, University of Oxford, Oxford, UK
jor.corral@eng.ox.ac.uk
[2] John Radcliffe Hospital, Oxford, UK
[3] Department of Computer Science, University of Oxford, Oxford, UK
[4] Biomedical Engineering Department, King's College London, London, UK

Abstract. Deep learning has revolutionized medical image analysis in recent years. Nevertheless, technical, ethical and financial constraints along with confidentiality issues still limit data availability, and therefore the performance of these approaches. To overcome such limitations, data augmentation has proven crucial. Here we propose SMOD, a novel augmentation methodology based on Statistical Models of Deformations, to segment 2D cine scans in cardiac MRI. In brief, the shape variability of the training set space is modelled so new images with the appearance of the original ones but unseen shapes within the space of plausible realistic shapes are generated. SMOD is compared to standard augmentation providing quantitative improvement, especially when the training data available is very limited or the structures to segment are complex and highly variable. We finally propose a state-of-art, deep learning 2D cardiac MRI segmenter for normal and hypertrophic cardiomyopathy hearts with an epicardium and endocardium mean Dice score of 0.968 in short and long axis.

Keywords: Data augmentation Segmentation Deep learning Models of deformation Cardiac MRI

1 Introduction

The expansion of big data has entered the cardiovascular medicine field in recent years [1, 2]. Machine learning approaches and deep learning techniques have been proven successful in medical imaging, and in particular in the cardiovascular field [3, 4]. This has had a strong impact on a wide range of applications, from classification, object detection and registration to phenotype clustering or risk prediction [3].

Among the uses of deep learning techniques in medical imaging, segmentation has been one of the most successful ones [3]. Since the U-net architecture was introduced by Ronneberger et al. in 2015 [5], a number of variations of that structure using

© Springer Nature Switzerland AG 2019
Y. Coudière et al. (Eds.): FIMH 2019, LNCS 11504, pp. 361–369, 2019.
https://doi.org/10.1007/978-3-030-21949-9_39

convolutional neural networks (CNN) have been explored, reporting excellent performance. A good example of this is the segmentation of both left and right ventricles (LV and RV) in cardiac magnetic resonance imaging (MRI) [6–8].

Despite these promising results, limitations and challenges remain [2], in particular concerning the availability, quantity and quality of training data [1]. Technical, ethical or financial constraints may limit data acquisition and confidentiality issues prevent access to data. Furthermore, manual data annotation is a costly and tedious process. Thus, the size of the training datasets in medical imaging is still a limiting factor. This scenario becomes even worse when dealing with specific diseases or applications. Since the performance of deep learning methods depends heavily on the available data, data augmentation and techniques such as transfer learning have become crucial for a better generalization and to avoid overfitting [9–12].

Image cropping, translation, rotation, scaling, ipping and even shearing are routinely applied as standard practice to enlarge datasets [11–14]. Other more complex approaches such as filtering, gamma correction, addition of Gaussian noise or even training a secondary network to perform the augmentation have been considered [9, 10, 12, 13]. Additionally, non-affine transformations, such as random elastic deformations or diffeomorphic transformations, have also been proposed [15, 16].

Statistical shape and deformation models have been used in a variety of medical image analysis tasks, including cardiac image segmentation but not, to the best of our knowledge, in combination with deep learning approaches [17]. We propose a new augmentation methodology based on models of deformation to segment 2D cine cardiac MRI scans and thus overcome the mentioned dataset size limitations. In brief, the shape variability of the training set space is modelled so new images with the gray level appearance of the original ones but unseen shapes within the space of plausible shapes are generated. As a result, the deformations are non-affine, follow a prior and are anatomically meaningful, in contrast with augmentation approaches based on random deformations [15]. We assess the performance of the proposed augmentation methodology versus standard augmentation approaches depending on data availability and on variability of the structure to be segmented. Finally, we build a state-of-art 2D cine MRI cardiac segmenter for both normal and Hypertrophic Cardiomyopathy (HCM) hearts, a condition that presents the largest amount of cardiac anatomical variability and represents the most common cause of sudden cardiac death in the young.

2 Materials and Methods

2.1 Dataset

We make use of two independent datasets containing cardiac MRIs for the assessment of our proposed augmentation methodology:

STACOM Dataset. This refers to the publicly available MICCAI 2018 LV Full Quantification Challenge STACOM dataset (https://lvquan18.github.io/). It contains 145 cine-MR mid-cavity short axis (SA) cardiac cycle sequences, 20 frames each. They were collected from 3 different hospitals, including diverse pathologies and a wide

range of subject ages. The LV endocardium and epicardium have been manually contoured and revised by two cardiac radiologists in all frames. The original pixel size ranges from 0.68 mm/pixel to 2.08 mm/pixel. Pre-processing to align the images was performed, as described in the challenge.

HCM Dataset. This dataset, published at [18], contains 55 HCM phenotype positive, 18 HCM phenotype negative but genotype positive and 37 normal patients scanned at Oxford John Radcliffe Hospital. Only mid-ventricular SA (to facilitate inter-dataset comparisons), horizontal long axis (HLA) and vertical long axis (VLA) end-diastole instances are considered. The LV endocardium, epicardium and left atrium (LA) along with RV endocardium were contoured in the end-diastole frames and double checked by 2 experts. The pixel size ranges from 1.33 to 2.08 mm/pixel.

2.2 Statistical Models of Deformation

Our registration method consists of an initial rigid registration, minimizing the sum of square distances (SSD) of image intensities, followed by the application of Diffeomorphic Log Demons [19] for non-rigid registration ($\sigma_{uid} = 2$, $\sigma_{diff} = 1.8$ and $\sigma_i/\sigma_x = 0.82$). We use this method to build a model of deformation, in the following way: (a) we generate an atlas representing the specific dataset population; (b) each image is registered to the atlas to calculate one velocity field per image; (c) velocity fields are combined using Principal Component Analysis (PCA) and the deformation model is then used to generate new samples. Methodological details are provided below.

Generation of a Population-Specific Atlas. (1) The atlas is initialized to a randomly selected image among the training set, which we denote as A_0. (2) Each image from the dataset, I, is rigidly registered to A_0 to obtain the registered set I_{T0}. (3) The average of the images in the I_{T0} set is calculated, obtaining the first iteration of the atlas, A_1. (4) I_{T0} is non-rigidly registered to A_1 and averaged to obtain A_2. Step 4 is iterated until convergence is achieved. (5) The resulting atlas, A, is manually segmented to generate A_s.

Set of Velocity Fields to Describe the Shape Variability. The set of segmentations of the images, S, is non-rigidly registered to A_s, obtaining the set of velocity fields, $\{v_i\}$, to diffeomorphically bring each image to the atlas space. The set $\{v_i\}$ encodes the variability of shape of the set of images, I, with respect to the reference A, and can be exploited to generate new deformations, which result in new images.

Dimensionality Reduction and Image Generation. Each velocity of the set $\{v_i\}$ can be organized as the column of a matrix V, to apply PCA and thus reduce the dimensionality. A mode cut capturing 90% variability is considered and the random velocity fields are sampled as follows:

$$v_g \quad V + U \quad x \quad d \qquad (1$$

Where V is the mean of the velocity fields; U, the eigenvectors of the principal modes; d, the eigenvalues of the principal modes; and x, an array with random numbers following the distribution $N(0, \sigma)$. We set the value of σ to 1 using visual inspection to achieve a realistic but varied set of shapes.

Each of the images, i, is brought to the atlas space applying v_i and transformed back to the image space applying the inverse of the random velocity field, v_g (different from displacement fields, inverse transformations are obtained just inverting the sign of the velocity field). Thus, a new image with the appearance of image i but a random shape within the space of variability of the original images is obtained (see Fig. 1).

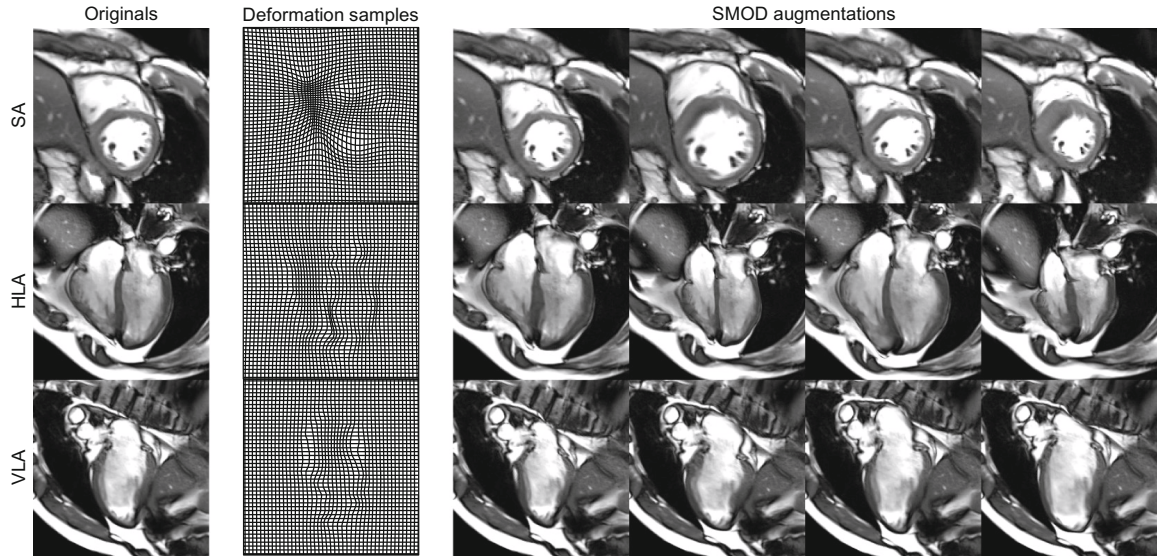


Fig. 1. Left column: Original mid-ventricular SA (top row), HLA (middle row) and VLA (bottom row) end-diastole MRI images. Middle column: Samples of the anatomically meaningful deformations obtained by applying SMOD on originals. Right Column: Augmentation samples from originals based on SMOD.

2.3 Augmentation Strategies

We consider three image augmentation strategies: (a) augmentation based on statistical models of deformations, SMOD, as described above; (b) standard augmentation, STD, based on random rotations (0–360°), ipping and translation (SA: –6 to 6 pixels in x and y, HLA and VLA: –10 to 10 pixels in x and y); and (c) a combination of both, SMOD+. For each cross-validation run, a model of deformations is learnt from scratch based on the particular training set to produce fair results. The augmentation process is "online", meaning that for each epoch the augmented training data fed to the network is different.

2.4 Neural Network

To segment the images, we used a Fully Convolutional Neural Network (FCNN) with skip connections following a U-Net architecture [5], as depicted in Fig. 2. The segmentation output layer has sigmoid activation functions and all the other FCN activations are rectified linear units (ReLU). Two convolutional layers (kernel size 3 3) and two batch normalization (BN) layers form a convolutional block, with the number of filters for each block specified in Fig. 2. Max-pooling and up-sampling layers have a stride of 2 in both dimensions. We used binary cross-entropy as loss function and early

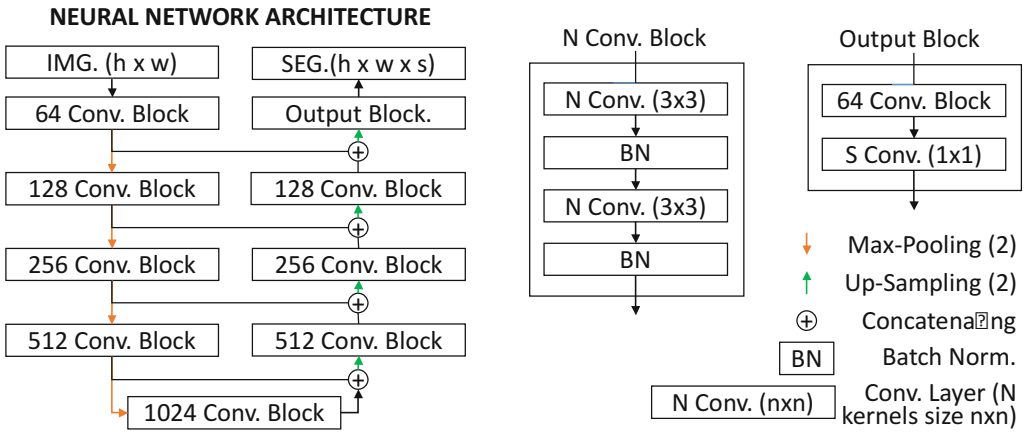


Fig. 2. Graphical overview of the proposed neural network structure.

stopping based on validation loss. The neural networks are implemented in Keras with Tensor ow as backend. The networks are trained with Adadelta optimizer on a NVIDIA GeForce GTX 1080 Ti 10 Gb. The Batch size is 8.

2.5 Assessment

We calculate Dice scores and Hausdorff distances between predicted and original segmentations to assess the performance of the augmentation strategies.

3 Experiments and Results

3.1 STACOM Dataset

We performed three batches of experiments considering different training set sizes to assess the impact of the augmentation strategy depending on data availability. In each batch, the three presented augmentation strategies were tested to segment the LV endocardium and epicardium in SA instances.

In the first batch, a five-fold cross-validation (29 subjects on each fold for testing) is performed. The remaining dataset is randomly split into 100 and 16 subjects, for training and validating data, respectively. In the second and third batches, the split into training/validation/testing data is 10/5/130 and 5/5/135 subjects, respectively. Cross fold validation was not considered in these 2nd and 3rd batches since the testing set, where the assessment is performed, already represents more than the 89% of the whole STACOM dataset in both scenarios. The images are cropped to 80 80 after augmentation and the whole cine sequence is considered in all experiments. That is, a training set of 100 subjects effectively means a training set of 20 100 segmented images.

The Dice score and Hausdorff distance results of the myocardium segmentations, calculated as the region between epicardium and endocardium, are depicted in Fig. 3. Similar results are obtained when evaluating endocardium and epicardium separately.

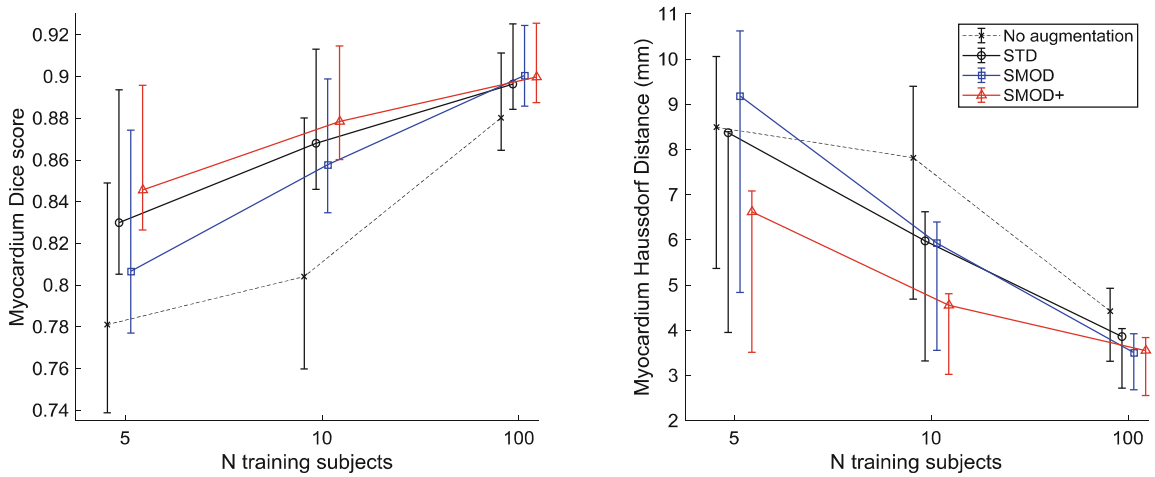


Fig. 3. STACOM dataset results. Mean Dice scores (left) and Hausdorff distances (right) in LV myocardium depending on the number of subjects to train with. Error-bars: 25th and 75th percentiles.

3.2 HCM Dataset

We assess the performance of the augmentation strategies depending on the intrinsic shape variability of the structure to be segmented on the HCM dataset, where not only the LV endocardium and epicardium are segmented but also the LA and the RV endocardium. Experiments to segment each of the three MRI planes were carried out with the three augmentation strategies. Five-fold cross validation is considered in all the experiments. The resulting split, in each of the folds, is: 22 testing, 78 training and 10 validation samples. After augmentation, the images are cropped to 80 80 in SA, and to 224 224 in HLA and VLA.

Figure 4 summarizes the segmentation Dice score results obtained for the different anatomical structures in the different MRI planes training under the three described augmentation strategies. Similar behavior and patterns are observed in Haussdorf distances (not shown). A representative example of the resulting segmentation using SMOD+ is presented in Fig. 5.

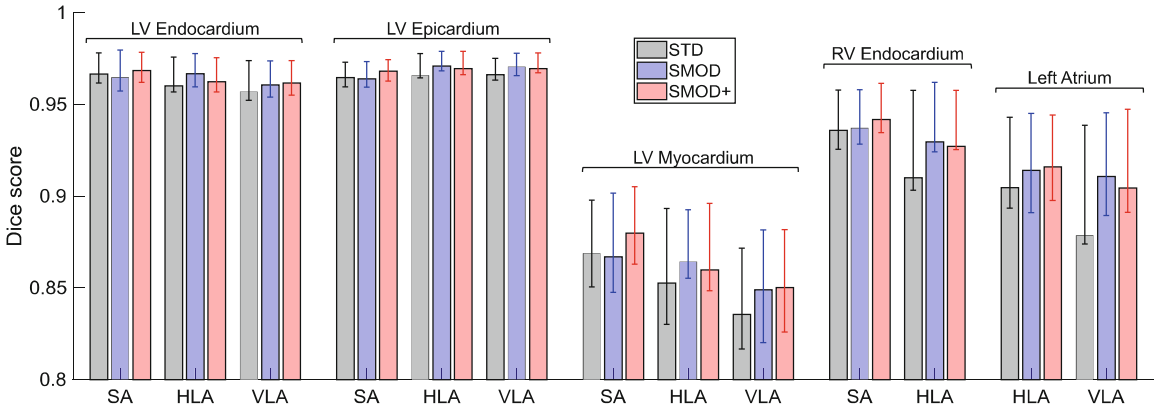


Fig. 4. HCM dataset results. Mean Dice scores and 25th and 75th percentiles grouped by segmented structure and MRI view.

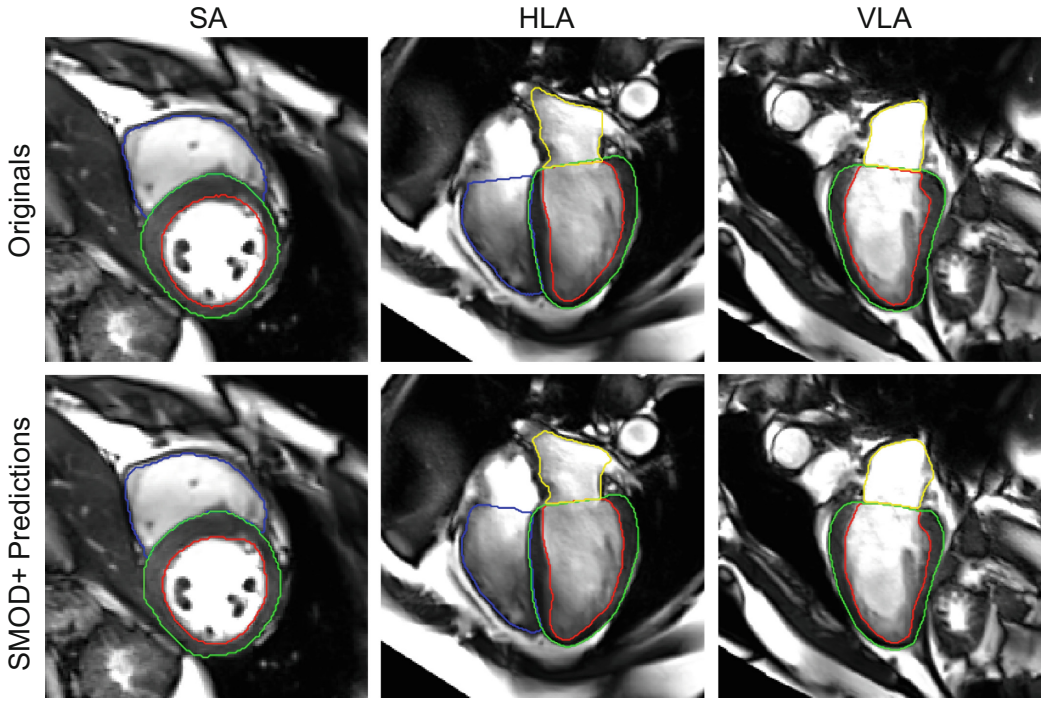


Fig. 5. Top row: LV endocardium (red), LV epicardium (green), LA (yellow) and RV endocardium (blue) ground truth segmentations of mid-ventricular SA (left), HLA (middle) and VLA (right) end-diastole MRI images from HCM dataset. Bottom row: SMOD+ predictions on these unseen instances. (Color figure online)

4 Discussion and Conclusions

Overall, our experiments confirm the importance of data augmentation, as reported in previous literature. Comparing to training with no augmentation, any of the three augmentation strategies significantly improves the results as it can be seen in Fig. 3.

The experiments completed on the STACOM dataset (Fig. 3) show that, while there is little difference among augmentation strategies when training with 100 subjects, SMOD+ outperforms the rest of strategies and becomes more relevant when the number of available subjects to train with decreases. We believe this occurs because the shape of the structures to segment (LV endocardium and epicardium) is reasonably simple and symmetrical, and thus when there are sufficient training samples the space of possible segmentations may be completely described with only rotations, translations and ipping. However, when the number of training subjects drops, standard augmentation is no longer enough to cover the space of variability and the models of deformations, learnt from the few samples, contribute filling the gaps and result in a better training. The observed pattern could be explained because the training network focuses more on local features rather than global ones to perform the segmentation task. The described behavior is confirmed both in Dice scores and Haussdorf distances.

In contrast to the STACOM data results when training with 100 subjects, the experiments on the HCM dataset show a more appreciable improvement in accuracy when using the proposed augmentation method (Fig. 4). SMOD+ reports better results than STD segmenting any structure in any plane, with difference being clearer in LV

myocardium, LA and RV endocardium. We believe this confirms our appreciation as explained above: HCM cases have a larger variation in anatomy due to the disease, and the LA and RV have a less regular shape than the LV. We should also note that, even though the training set size (78 subjects) in the HCM experiments is comparable to STACOM one, only one frame of the cine sequences (end-diastole) is considered in HCM versus 20 instances for the STACOM dataset. Therefore, an alternative explanation for the difference in performance might be the difference in data availability.

A limitation of the work is that only one neural network architecture has been considered. Furthermore, different levels of deformation (σ value, Eq. 1) as well as training in larger and more heterogeneous datasets could be explored. Methods other than PCA, which assumes Gaussianity, could be considered to better describe the distribution of deformations. Moreover, comparison to augmentation based on completely random deformations (no anatomical meaning) could be tested.

We expect the developed augmentation methodology to have more of an impact in 3D since usually fewer samples are available. We also presume that it could be scaled up to other applications and anatomies, particularly those with less regular shapes. Additionally, models of deformation could be learnt from secondary larger datasets and applied to the actual training dataset, enlarging even further the space of plausible shapes and presumably leading to a better generalization. Even though we chose segmentation-to-segmentation registration since less noise is introduced, image-to-image registration could be used to obtain the velocity fields, opening the possibility for learning variability from unsegmented cohorts.

To conclude, (a) a new augmentation methodology based on statistical models of deformation has been developed; (b) it has been proven that it improves the standard augmentation training results when segmenting 2D cardiac MRI slices; (c) the improvement is remarkable especially when the training data available is very limited or the structures to segment are complex and variable; and (d) an accurate state-of-art segmenter for healthy and HCM cine MRI scans has been built.

Acknowledgements. This work was supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie (g.a. 764738) and by the British Heart Foundation (PG/16/75/32383). Authors are financially supported by a Wellcome Trust Senior Research Fellowship (to PL, 209450/Z/17/Z) and a BHF Intermediate Basic Science Research Fellowship (to ABO, FS/17/22/32644).

References

1. Rumsfeld, J.S., Joynt, K.E., Maddox, T.M.: Big data analytics to improve cardiovascular care: promise and challenges. Nat. Rev. Cardiol. **13**(6), 350–359 (2016)
2. Shameer, K., Johnson, K.W., Glicksberg, B.S., Dudley, J.T., Sengupta, P.P.: Machine learning in cardiovascular medicine: are we there yet? Heart **104**(14), 1156–1164 (2018)
3. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
4. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H., Aerts, H.J.W.L.: Artificial intelligence in radiology. Nat. Rev. Cancer **18**(8), 500–510 (2018)

5. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
6. Oktay, O., Ferrante, E., Kamnitsas, K., Heinrich, M., Wai, B., Caballero, J., et al.: Anatomically Constrained Neural Networks (ACNNs): application to cardiac image enhancement and segmentation. IEEE Trans. Med. Imaging **37**(2), 384–395 (2018)
7. Poudel, R.P.K., Lamata, P., Montana, G.: Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. In: Zuluaga, M.A., Bhatia, K., Kainz, B., Moghari, M.H., Pace, D.F. (eds.) RAMBO/HVSMR -2016. LNCS, vol. 10129, pp. 83–94. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-52280-7_8
8. Vigneault, D.M., Xie, W., Ho, C.Y., Bluemke, D.A., Noble, J.A.: Ω-Net (omega-net): fully automatic, multi-view cardiac mr detection, orientation, and segmentation with deep neural networks. Med. Image Anal. **48**, 95–106 (2018)
9. Hussain, Z., Gimenez, F., Yi, D., Rubin, D.: Differential data augmentation techniques for medical imaging classification tasks. In: Proceedings American Medical Informatics Association Annual Symposium, vol. 2017, pp. 979–984 (2017)
10. Lemley, J., Bazrafkan, S., Corcoran, P.: Smart augmentation learning an optimal data augmentation strategy. IEEE Access **5**, 5858–5869 (2017)
11. Asperti, A., Mastronardo, C.: The effectiveness of data augmentation for detection of gastrointestinal diseases from endoscopical images. In: Proceedings 11th International Joint Conference on Biomedical Engineering Systems and Technologies, pp. 199–205 (2018)
12. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning, vol. 13, no. 1. MIT Press, Cambridge (2017)
13. Perez, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning. CoRR, vol. abs/1712.0 (2017)
14. Dong, H., Yang, G., Liu, F., Mo, Y., Guo, Y.: Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In: Valdés Hernández, M., González-Castro, V. (eds.) MIUA 2017. CCIS, vol. 723, pp. 506–517. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60964-5_44
15. Castro, E., Cardoso, J.S., Pereira, J.C.: Elastic deformations for data augmentation in breast cancer mass detection. In: IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), pp. 230–234 (2018)
16. Arteaga, M.O., Sørensen, L., Cardoso, J., Modat, M., et al.: PADDIT: probabilistic augmentation of data using diffeomorphic image transformation, October 2018. arXiv:1810.01928
17. Alba, X., et al.: An algorithm for the segmentation of highly abnormal hearts using a generic stat. shape model. IEEE Trans. Med. Imaging **35**(3), 845–859 (2016)
18. Raman, B., Ariga, R., Spartera, M., Sivalokanathan, S., Chan, K., Dass, S., et al.: Progression of myocardial fibrosis in hypertrophic cardiomyopathy: mechanisms and clinical implications. Eur. Heart J. Cardiovasc. Imaging **20**(2), 157–167 (2019)
19. Lorenzi, M., Ayache, N., Frisoni, G.B., Pennec, X., Alzheimer's Disease Neuroimaging Initiative (ADNI): LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm. Neuroimage **81**, 470–483 (2013)

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI
Publication Status	☑ Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	Jorge Corral Acero, Ernesto Zacur, Hao Xu, Rina Ariga, Alfonso Bueno-Orovio, Pablo Lamata, Vicente Grau, "SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI", In: Coudière Y., Ozenne V., Vigmond E., Zemzemi N. (eds) Functional Imaging and Modeling of the Heart. *FIMH 2019*. Lecture Notes in Computer Science, vol 11504. Springer, Cham; 2019. p361-69. DOI: https://doi.org/10.1007/978-3-030-21949-9_39

Student Confirmation

Student Name:	Jorge Corral Acero		
Contribution to the Paper	(1) Developed and implemented methods; (2) designed and ran experiments; (3) analysed results; (4) wrote the manuscript; and (5) prepared figures and tables.		
Signature		Date	12/07/2021

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. Vicente Grau		
Supervisor comments: I confirm that the above statement is accurate.		
Signature 	Date	12/07/2021

This completed form should be included in the thesis, at the end of the relevant chapter.

Paper III

Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation

Jorge Corral Acero, Hao Xu, Ernesto Zacur, Jurgen E. Schneider, Pablo Lamata, Alfonso Bueno-Orovio, Vicente Grau

In: Pop M. et al. (eds) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. STACOM 2019. Lecture Notes in Computer Science, vol 12009. Springer, Cham.
(https://doi.org/10.1007/978-3-030-39074-7_40)



Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation

Jorge Corral Acero[1(✉)], Hao Xu[1], Ernesto Zacur[1], Jurgen E. Schneider[2], Pablo Lamata[3], Alfonso Bueno-Orovio[4], and Vicente Grau[1]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK
{jor.corral,hao.xu}@eng.ox.ac.uk
[2] Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK
[3] Department of Biomedical Engineering, King's College of London, London, UK
[4] Department of Computer Science, University of Oxford, Oxford, UK

Abstract. Deep learning has been widely applied for left ventricle (LV) analysis, obtaining state of the art results in quantification through image segmentation. When the training datasets are limited, data augmentation becomes critical, but standard augmentation methods do not usually incorporate the natural variation of anatomy. In this paper we propose a pipeline for LV quantification applying our data augmentation methodology based on statistical models of deformations (SMOD) to quantify LV based on segmentation of cardiac MR (CMR) images, and present an in-depth analysis of the effects of deformation parameters in SMOD performance. We trained and evaluated our pipeline on the MICCAI 2019 Left Ventricle Full Quantification Challenge dataset, and achieved average mean absolute error (MAE) for areas, dimensions, regional wall thickness and phase of 106 mm^2, 1.52 mm, 1.01 mm and 8.0% respectively in a 3-fold cross-validation experiment.

Keywords: Deep learning Data augmentation LV quantification

1 Introduction

Automatic quantification of the left ventricle (LV) has been greatly enhanced by the development of deep learning algorithms in the past few years. Convolutional neural networks have shown great accuracy and exibility for LV quantification. Recently, the MICCAI 2018 Left Ventricle Full Quantification Challenge made possible to compare a wide range of deep learning algorithms performing on the same benchmark dataset with both direct regression [1] and segmentation based [2–4] approaches. Direct

Jorge Corral Acero and Hao Xu contributed equally.

© Springer Nature Switzerland AG 2020
M. Pop et al. (Eds.): STACOM 2019, LNCS 12009, pp. 384–394, 2020.
https://doi.org/10.1007/978-3-030-39074-7_40

regression approaches have shown promising results, while segmentation-based approaches were in general, at the time of the challenge, more accurate.

With the development of big databases such as UK Biobank [5], applying deep learning algorithms on big data has become possible in biomedical applications [6, 7], in uencing the choice and design of neural networks. With more training data, deeper networks with more parameters can be trained, which usually results in better performance. However, in clinical practice, especially for pathological cases, it is difficult to acquire such big dataset, and data augmentation becomes important. In this regard, our recently developed augmentation method based on statistical models of deformation has shown promising results on a variety of datasets for segmentation task [8].

The MICCAI 2019 Left Ventricle Full Quantification Challenge has provided a benchmark dataset which, compared to the corresponding 2018 dataset, is closer to real-life clinical conditions, with no pre-processing applied to the images. We propose a segmentation-based quantification pipeline enhanced with statistical models of deformation, developed and evaluated on this dataset.

2 Methods

We propose a complete pipeline for quantifying the LV from cardiac MR (CMR) images, consisting of the following steps. We first build a population-specific atlas, and train an initial neural network to locate the centre of the heart in all the images. We then rigidly register each image to the atlas previously calculated. We build the statistical models of deformation, which we use to augment the images using different strategies. Finally, we train a second neural network to perform the fine segmentation and retrieve the LV metrics from the segmentation results.

2.1 Data

We developed and evaluated our pipeline using the MICCAI 2019 Left Ventricle Full Quantification Challenge dataset, which consists of 56 training subjects and 30 testing subjects. For each subject in the training data, a single short-axis (SAX) CMR sequence consisting of 20 frames was provided together with a set of clinically significant LV indices including regional wall thicknesses, cavity dimensions, cavity areas and myocardium and cardiac phase for each frame. Endocardial and epicardial segmentation binary masks were also made available as reference, and pixel-spacing values were also given for metrics evaluation. For subjects in the testing dataset, only CMR image sequences and pixel-spacing values were provided.

Comparing to MICCAI 2018 Left Ventricle Full Quantification Challenge, which had 145 training subjects and 30 testing subjects, the size of training dataset reduced by 61.4% and the testing dataset remained the same size [10].

2.2 Rigid Registration

Our rigid registration method was based on the maximization of cross-correlation of image intensities. In order to avoid converging to a local minimum, the algorithm was

initialized to different transformations distributed in the space of possible transformations. Diffeomorphic Log Demons [11] was applied for non-rigid registration (σ_{fluid} 2, σ_{diff} 1.8 and σ_i/σ_x 0.82).

2.3 Atlas

In order to build the atlas, the set of images, I, was first rigidly aligned, using only the first frame, and then non-rigidly registered.

For rigid alignment, the atlas was initialized to a randomly selected instance among the training set, which we denote as A_0, and cropped to completely contain the heart. The rest of the instances were first centred, assuming the mass centre of the epicardium reference segmentation as the centre of the LV, and then rigidly registered to A_0, constraining the transformation to rotations only. The obtained transformations for each of the first frames were extended to the other frames to obtain the registered set IT_0. The intensity average of the images in the IT_0 set was calculated to obtain the first iteration of the atlas, A_1.

For non-rigid alignment, the segmentations of IT_0 were non-rigidly registered to the segmentation of A_1, obtaining the transformation set T_1. The transformations T_1 were then applied to IT_0 and the average of intensities calculated to obtain the atlas, A. Since the segmentation masks were used, convergence was achieved in one single step.

2.4 Initial Segmentation and Rigid Registration

To initialize the rigid alignment, we trained a variation of U-Net [9] for epicardial segmentation. We first down-sampled all the images to 256 256 and normalized them by clipping the smallest and largest 5% intensity values. More details of the network are described in Sect. 2.6. Based on the initial epicardium segmentation of the first frames, we centred and oriented the set of images, I, to the atlas, A, as described in Sect. 2.2.

2.5 Statistical Models of Deformation

We implemented the statistical models of deformation following the SMOD+ method in [8]. Once the rigid registration was completed, the set of segmentations of the images, S, is non-rigidly registered to the atlas segmentation, A_s, obtaining the set of velocity fields, $\{v_i\}$, to diffeomorphically bring each image to the atlas space.

This set $\{v_i\}$ intrinsically encodes the shape variability of the set of images, I, with respect to the reference A. Thus, the distribution of $\{v_i\}$ can be sampled to obtain new velocity fields that implicitly lead to anatomically meaningful deformations within the space of plausible shapes, and we built a statistical model of deformations that can be exploited to generate new images.

In order to generate random deformations, v_g, we first reduced the dimensionality of the distribution of velocity fields by applying principal component analysis (PCA) on the residuals. Then, we sampled the relative weights of the main modes of variation with a multivariate Gaussian distribution, centred at 0 and with standard deviation σ. Finally, each of the images, i, was brought to the atlas space applying v_i and

transformed back to the image space applying the inverse of the random velocity field, v_g. Thus, a new image with the appearance of image i but a random shape within the space of variability of the original images was obtained.

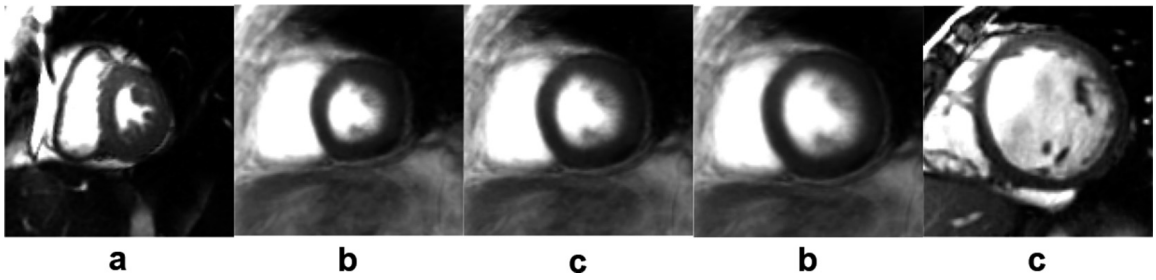


Fig. 1. Atlas and extreme cases of LV shape. The images were input images of the neural networks with the size of 128 128 and pixel-spacing of 1.1 mm. (*a*) and (*e*) are the smallest and largest LVs from the original dataset, respectively; (*b*) to (*d*) are generated by PCA mode 1 with $\sigma =$ 3, $\sigma = 0$ and $\sigma =$ +3.

2.6 Augmentation Strategies

We implemented two augmentations strategies: (1) standard augmentation based on random ipping, rotations (0–360°) and translations (±11 mm in x and y); and (2) augmentation based on SMOD+, which we combined with standard augmentation samples due to the large variability of LV sizes shown in Fig. 1.

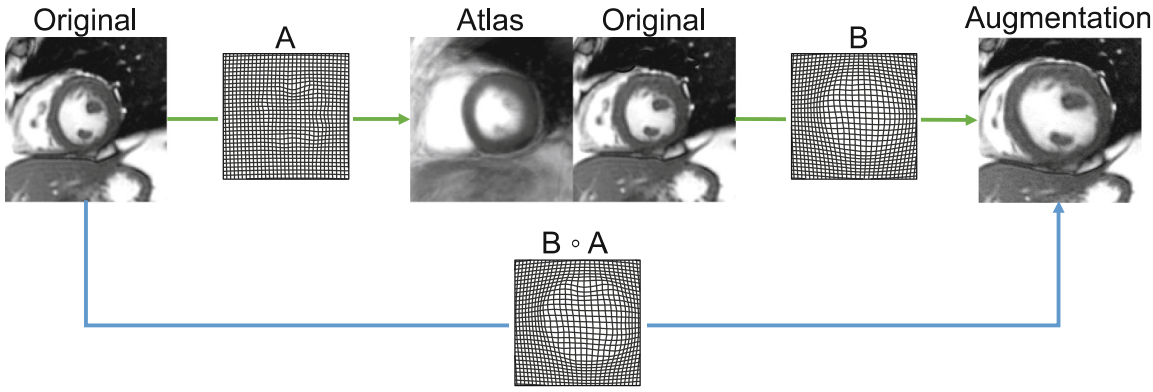


Fig. 2. Combined transformation stages.

The different transformations needed to generate a new image were mathematically combined by convolution as shown in Fig. 2, and therefore the images were interpolated and resized only once at the end of the process. The final resolution used as the neural network input was 128 128, with a pixel-spacing of 1.1 mm.

2.7 Neural Network

We compared two neural networks for the final segmentation task, which were a variation of U-Net [9] and a segmentation network based on VGG-16 [6]. For both networks, there were four 2 2 max-pooling stages with the stride of 2 2, and the number of filters were 64, 128, 256, 512 and 1024 for each stage accordingly. The size of all kernels was 3 3 and the activation functions were ReLU for all layers other than the output layer, which was sigmoid. The key difference between the networks was the up-sampling process, with step-by-step up-sampling stages for the U-Net and concatenated up-samples from each scale for VGG-16. This difference is shown in Fig. 3. We implemented the training with cross-entropy as the loss function and Adadelta as optimizer.

The initial segmentation network introduced in Sect. 2.3 shared the same U-Net architecture, while the input size was 256 256 and the number of filters were decreased to 16, 32, 64, 128 and 256 for efficiency.

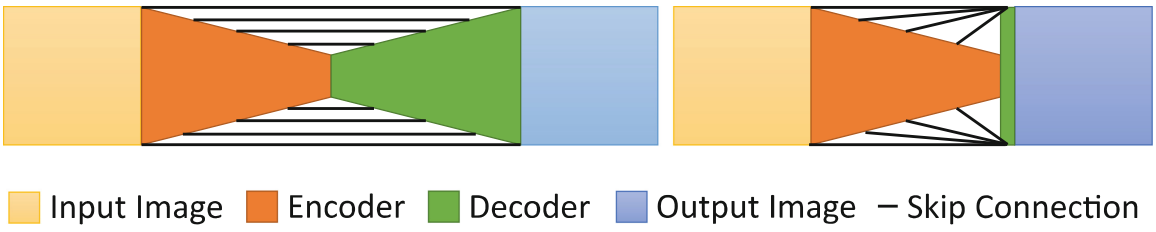


Fig. 3. Schematics of the two neural networks compared in this paper.

2.8 Metrics Evaluation

In the absence of a detailed description of metrics calculation in the challenge, the following approach was adopted. Metrics were calculated from our segmentation results by first converting the neural network outputs to binary masks and then thresholding at 0.5. We extracted the largest object from the binary masks and filled any existing holes. The areas were calculated by multiplying the pixel area times the number of pixels of the region. The dimensions were calculated by averaging the distances between the endocardial contour points and the cavity centroid within the corresponding section. To calculate the regional wall thickness, we first calculated the middle contour of the myocardium and then averaged the closest point-to-point distances between both endocardial and epicardial contours to the middle contour for each middle contour point. The phase estimation was calculated by first defining the frames with maximum and minimum cavity areas to be end-diastolic (ED) and end-systolic (ES) frames, and then assigning linearly interpolated labels to the other frames.

Applying our metrics estimation method to the reference segmentations provided by challenge organisers led to a bias when compared to the set of reference metric values also provided by challenge organisers. Such differences with respect to the provided dataset introduced unnecessary complexity when designing the pipeline and could have been at least partly (and for the areas fully) eliminated with the provision of detailed descriptions of metrics calculation by the organisers. To minimize possible

errors that could be introduced within this stage, we calculated a correction factor λ using the reference area metrics (Ar) and the segmentation estimated areas (As), which was the only metric independent of LV orientation, by minimizing the error of ($Ar - \lambda As$). The square-root of λ was then multiplied to 1D metrics estimated from the segmentation.

3 Experiments and Results

We performed 3-fold cross-validation experiments on the training dataset, with the size of each fold being 18, 19 and 19. The subjects were randomly assigned to one-fold, and for each cross-validation experiment we used 4 subjects as validation set and kept the rest as training set.

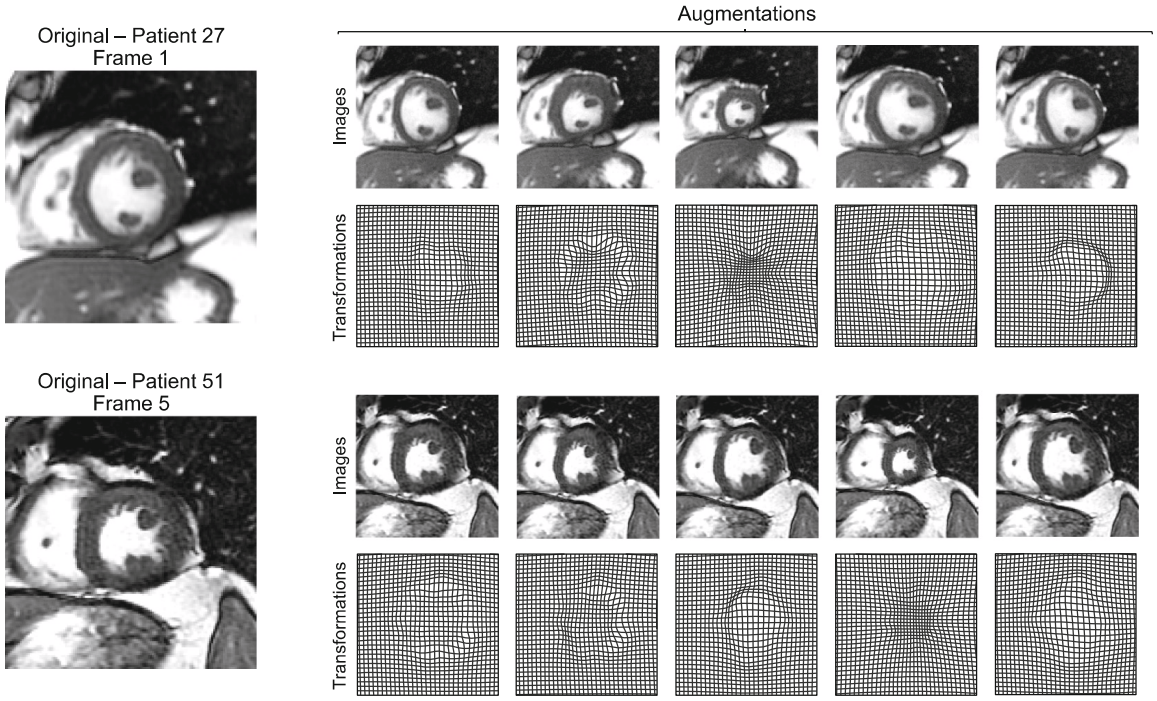


Fig. 4. Example of generated augmentations. Five randomly generated augmentations are shown for each of the two images. The augmented cases varied in size, shape and myocardium thickness of LV.

A model of deformation (described in Sect. 2.4) was learnt for each fold, and the metrics correction factors (described in Sect. 2.7) were also calculated for each fold independently. The network parameters were updated using the training set and model selection was performed using the validation set with early stop. For each training epoch, new instances of training images were randomly generated and used to update the network parameters. Examples of resultant augmented images were shown in Fig. 4, along with the combined transformations.

Table 1. Mean absolute error results.

	Base-line	U-Net Std.	U-Net σ = 1	U-Net σ = 2	U-Net σ = 3	VGG Std.	VGG σ = 1	VGG σ = 2	VGG σ = 3	Test
Dice Endo	1 ±0	0.950 ±0.02	0.949 ±0.033	**0.953** ±0.023	0.950 ±0.026	0.938 ±0.045	0.941 ±0.031	0.943 ±0.032	0.941 ±0.031	N/A
Dice Epi	1 ±0	0.966 ±0.01	0.965 ±0.019	**0.967** ±0.014	0.965 ±0.013	0.957 ±0.026	0.958 ±0.020	0.959 ±0.021	0.959 ±0.020	N/A
A1 (mm^2)	24 ±19	102 ±87	107 ±111	92 ±83	100 ±83	101 ±121	95 ±77	90 ±79	101 ±84	184
A2 (mm^2)	25 ±19	132 ±105	142 ±105	121 ±98	135 ±105	140 ±121	155 ±131	126 ±110	140 ±112	525
Areas (mm^2)	25 ±19	117 ±97	125 ±109	**106** ±91	118 ±96	121 ±122	125 ±111	108 ±97	120 ±101	355
Dim1 (*mm*)	0.64 ±0.72	1.59 ±1.53	1.73 ±2.30	1.46 ±1.28	1.58 ±1.45	1.40 ±1.78	1.40 ±1.16	1.32 ±1.08	1.51 ±1.23	2.59
Dim2 (*mm*)	0.68 ±0.74	1.70 ±1.39	1.81 ± 2.18	1.53 ± 1.31	1.60 ± 1.38	2.17 ± 3.01	1.72 ± 1.48	1.89 ± 1.65	1.96 ±1.77	2.33
Dim3 (*mm*)	0.77 ±0.87	1.65 ± 1.32	1.63 ± 1.59	1.56 ± 1.28	1.64 ± 1.26	2.01 ± 2.54	1.76 ± 1.55	1.83 ± 1.62	1.81 ±1.48	2.40
Dims (*mm*)	0.69 ±0.78	1.65 ±1.42	1.72 ±2.05	**1.52** ±1.29	1.60 ±1.36	1.86 ±2.52	1.63 ±1.42	1.68 ±1.49	1.76 ±1.52	2.44
RWT1 (*mm*)	0.35 ±0.45	1.01 ±1.03	0.98 ±0.89	0.85 ±0.68	0.89 ±0.68	0.89 ±0.75	1.01 ±0.89	0.91 ±0.80	0.91 ±0.84	2.40
RWT2 (*mm*)	0.41 ±0.37	1.23 ±0.85	1.23 ±0.92	1.05 ±0.78	1.18 ±0.86	1.19 ±0.90	1.19 ±0.94	1.15 ±0.84	1.22 ±0.87	2.39
RWT3 (*mm*)	0.33 ±0.27	1.27 ±0.97	1.22 ±0.95	1.10 ±0.87	1.26 ±1.03	1.21 ±0.97	1.21 ±0.97	1.15 ±0.92	1.26 ±1.00	2.20
RWT4 (*mm*)	0.36 ±0.45	1.20 ±0.90	1.27 ±0.97	1.21 ±1.02	1.23 ±1.00	1.22 ±0.99	1.16 ±0.93	1.13 ±0.91	1.29 ±1.10	1.91
RWT5 (*mm*)	0.41 ±0.37	0.91 ±0.75	1.02 ±0.79	1.00 ±0.74	0.97 ±0.78	1.10 ±1.30	0.93 ±0.78	0.99 ±0.89	1.14 ±0.97	1.98
RWT6 (*mm*)	0.45 ±0.43	0.91 ±0.76	0.85 ±0.72	0.84 ±0.66	0.88 ±0.67	1.08 ±1.31	1.02 ±0.78	0.99 ±0.92	0.93 ±0.74	2.21
RWT (*mm*)	0.38 ±0.40	1.09 ±0.9	1.10 ±0.89	**1.01** ±0.81	1.07 ±0.86	1.12 ±1.06	1.09 ±0.89	1.06 ±0.89	1.12 ±0.94	2.18
Phase (%)	2.0	7.9	8.3	8.0	8.1	8.2	**7.8**	8.4	8.2	9.5

Results of the experiments are shown in Table 1. Errors in LV metrics obtained from ideal segmentations are reported in the baseline experiment, which used the reference segmentation provided by challenge organizers after applying the correction factor. For area metrics, after applying the correction factor there was still a mean absolute error (MAE) of 25 mm^2, which is around 25% of the MAE with our best segmentation results. Such an error might have been removed shall we had an accurate description of metric calculations. We could also see a 2% phase estimation error, which is purely dependent on cavity areas and introduced during resampling the images, suggesting the reference phase was sensitive to small noise.

Comparing the two networks, the performance of the U-Net was better than VGG-16 based segmentation network for Dice score, area, dimension and regional wall thickness values. Despite a more accurate estimation of the endocardium using the U-Net, VGG-16 achieved a more accurate phase estimation. This could be caused by the effect of noise we detected in the baseline experiments. From the results we could see that there was a negative effect on the segmentation task by removing multiple stage up-sampling, even though VGG-16 is deeper in the down-sampling stages.

Comparing the two augmentation strategies, our modified SMOD+ approach with $\sigma = 2$ produced the best results. The performance of $\sigma = 1$ and $\sigma = 3$ were limited because the variation of the deformation was either too close to the atlas or far enough to become unrealistic, and for both cases the generalization of the network was disrupted with either unbalanced data or unexpected data. By calculating the p-values, we found significant differences between the two augmentation strategies.

Bland-Altman plots were produced to show the agreement between our best performing network with the reference metrics in Fig. 5. The vast majority of the data points lies within $mean \pm 1.96 \quad std$ suggesting a good agreement between the two measurements.

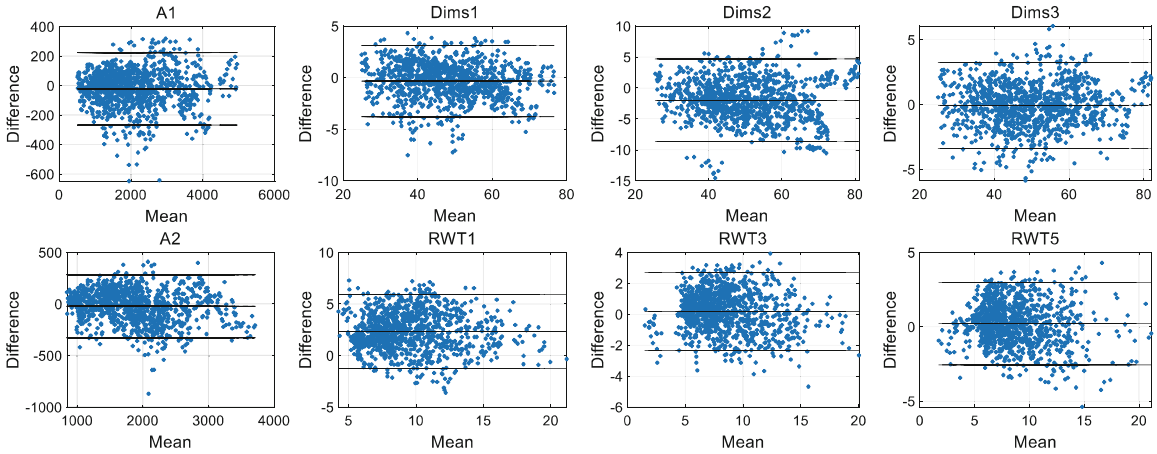


Fig. 5. Bland-Altman plots for U-Net with σ = 2. Areas, dimensions and three regional wall thickness metrics are shown.

We also evaluated qualitatively the segmentation results of our best performing network. Three examples with Dice score from high to low (including the worst case) are shown in Fig. 6. Our segmentation results from neural networks appeared to be consistent with image features, however, the manual reference segmentation contours were comparatively independent from image features. The values of Dice score showed similarity between our segmentation results and the provided references, and larger Dice score represented better similarity between the two.

For the testing dataset we used the entire training dataset to get the model of deformation and the correction factor for better generalization. We used all three networks of U-Net with $\sigma = 2$ and embedded the neural network predictions by averaging before calculating the metrics. The performance of our pipeline on the testing dataset is shown in Table 1. Comparing to the cross-validation experiment results using

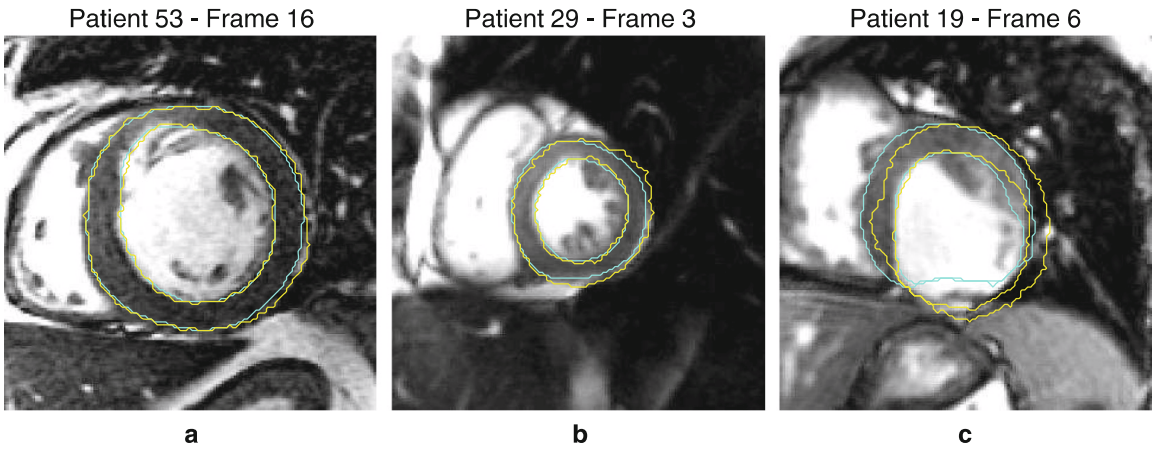


Fig. 6. Segmentation results of training dataset. (*a*) to (*c*) correspond to Dice scores from high to low. The yellow contour is the reference segmentation, and the cyan contour is our proposed U-Net result based on SMOD+ augmentation. (Color figure online)

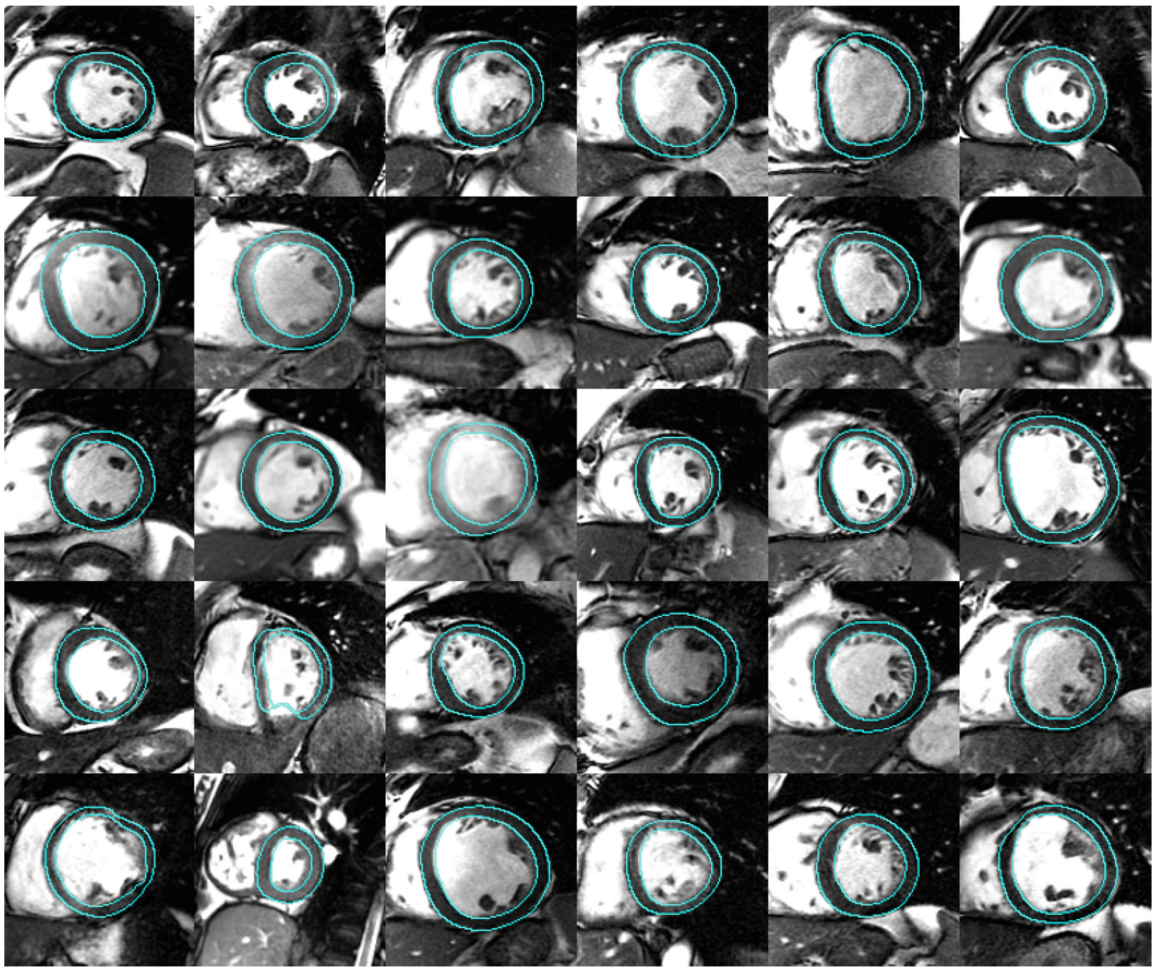


Fig. 7. Segmentation results of the testing dataset. We presented the segmentation result of the first frame of all subjects within the testing dataset. Different from the training dataset, there were no reference segmentation provided, and therefore only the segmentation results from our proposed neural networks are shown.

the training dataset, the testing dataset errors are comparatively larger. For metric A2 (representing myocardium area), the testing result mean absolute error is more than 4 times bigger and reaches 525 mm^2, which is larger than a square with the side of 2 cm.

In order to further investigate this difference between training dataset and testing dataset results, we produced the segmentation result of all the subjects in the testing data for qualitative analysis. Results for the first frame of each subject are shown in Fig. 7. From the visual inspection, the segmentation results of the testing dataset were comparable with the training dataset, and there was no clear evidence suggesting why would the metric evaluation of the testing dataset performed worse than in the training dataset based on the segmentation results.

4 Conclusion

In this paper, we have proposed a full quantification pipeline of the LV using CMR mages, developed and applied to the MICCAI 2019 Left Ventricle Full Quantification Challenge. We performed 3-fold cross-validation experiments on the training dataset, and for all the combinations of network structure and augmentation strategies, U-Net with our modified SMOD+ augmentation achieved the best results within our pipeline, showing the benefits of using multi-stage up-sampling and advanced augmentation strategies.

Compared to MICCAI 2018 Left Ventricle Full Quantification Challenge, the dataset was closer to real-life clinical conditions by removing the pre-processing of the images. At the same time, the size of the training dataset was reduced from 145 to 56 subjects. Both changes made the task significantly more challenging, which steered our focus towards the pre-processing and metrics evaluation stages, as well as the implementation of an anatomically meaningful augmentation method to enhance the neural network performance. Despite the more challenging task, our method achieved comparable results to last year's participants for both cross-validation on the training dataset and the final testing dataset.

The performance of our pipeline on the testing dataset did not reach the level of our cross-validation experiments, and based on the provided qualitative evaluation of the segmentation results the reason of such big differences between the mean absolute errors remains unclear to us. Similar performance drops in testing datasets were also identified in all the best ranking methods in MICCAI 2018 Left Ventricle Full Quantification Challenge [1–4]. It appears to us that this phenomenon is less dependent on the candidate methods, but rather closely related to the distribution of subjects in the training and testing dataset. Additional details on the testing dataset, and an explicit description of metrics calculation, would facilitate the interpretability of these results and improve future challenges.

Acknowledgments. This work was supported by the European Unions Horizon 2020 research and innovation program under the Marie Sklodowska-Curie (g.a. 764738) and by the British Heart Foundation (PG/16/75/32383). Authors are financially supported by a Wellcome Trust Senior Research Fellowship (to PL, 209450/Z/17/Z) and a BHF Intermediate Basic Science Research Fellowship (to ABO, FS/17/22/32644).

References

1. Li, J., Hu, Z.: Left ventricle full quantification using deep layer aggregation based multitask relationship learning. In: Pop, M., et al. (eds.) STACOM 2018. LNCS, vol. 11395, pp. 381–388. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12029-0_41
2. Kerfoot, E., Clough, J., Oksuz, I., Lee, J., King, A.P., Schnabel, J.A.: Left-ventricle quantification using residual U-Net. In: Pop, M., et al. (eds.) STACOM 2018. LNCS, vol. 11395, pp. 371–380. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12029-0_40
3. Xu, H., Schneider, J.E., Grau, V.: Calculation of anatomical and functional metrics using deep learning in cardiac MRI: comparison between direct and segmentation-based estimation. In: Pop, M., et al. (eds.) STACOM 2018. LNCS, vol. 11395, pp. 402–411. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12029-0_43
4. Guo, F., Ng, M., Wright, G.: Cardiac MRI left ventricle segmentation and quantification: a framework combining U-Net and continuous max- ow. In: Pop, M. (ed.) STACOM 2018. LNCS, vol. 11395, pp. 450–458. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12029-0_48
5. UK Biobank. https://www.ukbiobank.ac.uk/
6. Bai, W., et al.: Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. J. Cardiovasc. Magn. Reson. **20**(1), 65 (2018)
7. Petersen, S.E., et al.: The impact of cardiovascular risk factors on cardiac structure and function: Insights from the UK Biobank imaging enhancement study. PLoS ONE **12**(10), e0185114 (2017)
8. Corral Acero, J., et al.: SMOD - data augmentation based on statistical models of deformation to enhance segmentation in 2D cine cardiac MRI. In: Coudière, Y., Ozenne, V., Vigmond, E., Zemzemi, N. (eds.) FIMH 2019. LNCS, vol. 11504, pp. 361–369. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21949-9_39
9. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
10. Xue, W., Brahm, G., Pandey, S., Leung, S., Li, S.: Full left ventricle quantification via deep multitask relationships learning. Med. Image Anal. **43**, 54–65 (2018)
11. Orbes-Arteaga, M., et al.: PADDIT: probabilistic augmentation of data using diffeomorphic image transformation. In: Medical Imaging 2019: Image Processing, vol. 10949, p. 109490S. International Society for Optics and Photonics, March 2019

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation
Publication Status	☑ Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	Jorge Corral Acero, Hao Xu, Ernesto Zacur, Jurgen E. Schneider, Pablo Lamata, Alfonso Bueno-Orovio, Vicente Grau, "Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation", In: Pop M. et al. (eds) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. *STACOM 2019*. Lecture Notes in Computer Science, vol 12009. Springer, Cham; 2019. p384–94. DOI: https://doi.org/10.1007/978-3-030-39074-7_40

Student Confirmation

Student Name:	Jorge Corral Acero		
Contribution to the Paper	(1) Developed and implemented SMOD; (2) co-implemented the deep learning architectures; (2) designed and ran all experiments; (3) analysed results; (4) prepared figures; and (5) prepared the final manuscript, in collaboration with Hao Xu. The contribution of Dr. Hao Xu, co-author of this paper, shall be acknowledged.		
Signature		Date	12/07/2021

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. Vicente Grau		
Supervisor comments: I confirm that the above statement is accurate.		
Signature 	Date	12/07/2021

This completed form should be included in the thesis, at the end of the relevant chapter.

Paper IV

A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation

Jorge Corral Acero, Vaanathi Sundaresan, Nicola Dinsdale, Vicente Grau, Mark Jenkinson

In: Puyol Anton E. et al. (eds) Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges. STACOM 2020. Lecture Notes in Computer Science, vol 12592. Springer, Cham.
(https://doi.org/10.1007/978-3-030-68107-4_20)



A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation

Jorge Corral Acero[1(✉)], Vaanathi Sundaresan[2], Nicola Dinsdale[2], Vicente Grau[1], and Mark Jenkinson[2,3,4]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, UK
jor.corral@eng.ox.ac.uk
[2] Wellcome Centre for Integrative Neuroimaging, Oxford Centre for Functional MRI of the Brain, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK
[3] Australian Institute of Machine Learning, Department of Computer Science, University of Adelaide, Adelaide, SA 5005, Australia
[4] South Australian Health and Medical Research Institute (SAHMRI), North Terrace, Adelaide, SA 5000, Australia

Abstract. Segmentation of anatomical structures from Cardiac Magnetic Resonance (CMR) is central to the non-invasive quantitative assessment of cardiac function and structure. Anatomical variability, imaging heterogeneity and cardiac dynamics challenge the automation of this task. Deep learning (DL) approaches have taken over the field of automatic segmentation in recent years, however they are limited by data availability and the additional variability introduced by differences in scanners and protocols. In this work, we propose a 2-step fully automated pipeline to segment CMR images, based on DL encoder-decoder frameworks, and we explore two domain adaptation techniques, domain adversarial training and iterative domain unlearning, to overcome the imaging heterogeneity limitations. We evaluate our methods on the MICCAI 2020 Multi-Centre, Multi-Vendor & Multi-Disease Cardiac Image Segmentation Challenge training and validation datasets. The results show the improvement in performance produced by domain adaptation models, especially among the seen vendors. Finally, we build an ensemble of baseline and domain adapted networks, that reported state-of-art mean Dice scores of 0.912, 0.857 and 0.861 for left ventricle (LV) cavity, LV myocardium and right ventricle cavity, respectively, on the externally validated Challenge dataset, including several unseen vendors, centers and cardiac pathologies.

Keywords: Segmentation · Cardiac magnetic resonance · Deep learning · Domain adaptation · Data harmonization

J. Corral Acero and V. Sundaresan–Equally contributed to the work.
V. Grau and M. Jenkinson–Equally co-supervised the work.

© Springer Nature Switzerland AG 2021
E. Puyol Anton et al. (Eds.): STACOM 2020, LNCS 12592, pp. 196–207, 2021.
https://doi.org/10.1007/978-3-030-68107-4_20

1 Introduction

Cardiac magnetic resonance (CMR) imaging allows for an accurate non-invasive quantification of cardiac function and structure [1–3]. The valuable information that it provides in cardiovascular disease management has been repeatedly shown [4–9]. Nevertheless, the anatomical variability and the intrinsic complexity of cardiac dynamics and geometry represent a challenge, and CMR analysis remains manual in clinical practice [3, 10–12]. Deep learning (DL) has revolutionized medical image analysis in recent years, progressing towards automatic segmentation. However, these approaches are challenged by (1) the limited data availability due to technical, ethical and financial constraints along with confidentiality issues, especially for specific pathologies and (2) the imaging heterogeneity introduced by anatomical variability and the use of different scanners and protocols [13–16].

Various techniques have been proposed for improved and robust performance of DL models under limited training data. The most common technique is data augmentation, including affine and non-affine transformations to populate the space of shape variability [17–20]. Other techniques incorporate modifications in the network architecture such as reducing the number of network parameters to avoid overfitting or using a 2-step segmentation pipeline that improves class balance by zooming into the anatomical region of interest (ROI) [21, 22]. Scanner-induced variation in datasets echoes 'domain shift' which leads to the performance of models trained on data from one scanner (the source domain) degrading when applied to another (the target domain). Techniques such as transfer learning [23] have proven successful. However, fine-tuning for every unseen domain is still required. Solutions based on domain adaptation (DA) techniques, which aim to create a single feature representation for all domains which is invariant to domain but discriminative for the task of interest [24], have been proposed to overcome this limitation. One of the successful DA approaches is *domain adversarial training of neural networks* (DANN) [25]. DANN assumes that predictions must be based on domain-invariant features, and jointly optimizes the underlying features to simultaneously minimize the loss of the label predictor and maximize the loss of a domain predictor. An alternative approach was proposed using an iterative framework [26] of *domain unlearning* (DU) for adversarial adaptation, creating a classifier which is more uniformly uninformative across domains [27]. This framework has been successfully applied to harmonization of MRI for brain tissue segmentation [28].

The MICCAI 2020 Multi-Centre, Multi-Vendor & Multi-Disease Cardiac Image Segmentation Challenge (M&Ms 2020), including a dataset with a wide variety in centers and scanner vendors and granting external validation, provides a benchmark to assess the generalizability of the segmentation algorithms [29]. We propose a fully automated pipeline to segment CMR short-axis (SAx) stacks, based on the 2-step DL framework that was awarded 1st prize in the LVQuan19 Challenge [22, 30]. We evaluate the segmentation performance of the proposed method, and explore two DA techniques, DANN and DU, for multi-vendor and multi-center applications on the M&Ms 2020 Challenge. The results show that the DA techniques contribute to performance improvement and suggest more experimentation for improvement on unseen vendors applications. A final model ensemble was built, achieving state-of-art performance for both seen and unseen vendors.

2 Materials and Methods

In essence, the proposed pipeline first locates where the heart is (1st neural network - NN) and then focuses on that ROI to produce a fine segmentation (2nd NN). Three alternative implementations are proposed for this latter step. The pipeline is coupled with pre-processing and post-processing stages, as illustrated on Fig. 1.

2.1 Data

We deployed and evaluated our methods on the publicly available M&Ms 2020 Challenge dataset [29], which involves 4 acquisition centers and consists of 150 annotated SAx images from two different MRI vendors, A and B (75 each), and 25 unannotated images from a third vendor, C. The annotated images are segmented only at end-diastole (ED) and end-systole (ES), including left ventricular (LV) cavity, LV endocardium and right ventricle (RV) cavity masks. The image resolution varies from 192 × 240 to 384 × 384 pixels, and the pixel spacing, from 0.977 to 1.625 mm. An additional testing set of 200 cases from 6 acquisition centers, including vendors A, B, C and a fourth unseen vendor, D, in equal proportions, is held by the organizers for external validation (20% of the set – 40 cases) and final challenge results (80% of the set). Details on the acquisition and annotation protocols can be checked in [29].

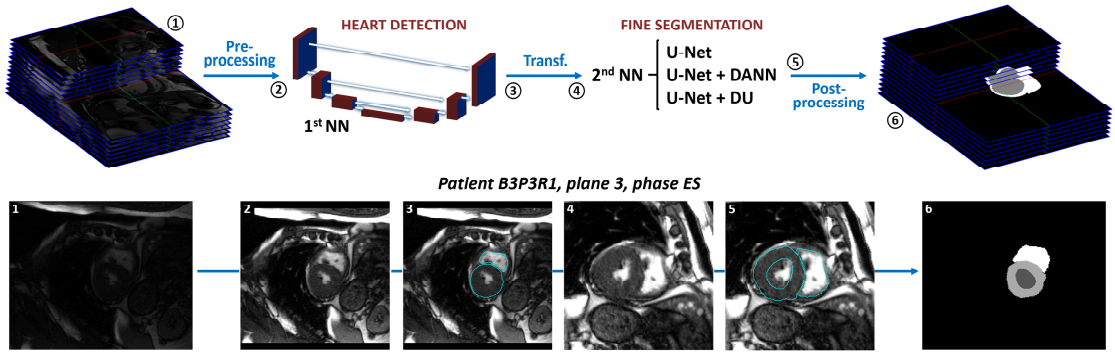


Fig. 1. Scheme of the proposed segmentation pipeline (top) along with a step-by step explanatory illustration of its application on patient B3P3R1 (bottom).

2.2 1st NN: Pre-processing, Heart Detection and Transformation

The 3 most central ES and ED slices of each patient are normalized in intensity and resolution (pre-processing – Fig. 1, step 1–2) and fed to the 1stNN that segments the LV epicardium and RV endocardium (heart detection – Fig. 1, step 2–3), so that a ROI can be defined (transformation – Fig. 1, step 3–4). The pre-processing consists of a linear interpolation to a 2D pre-defined template (256 × 256 pixels with symmetrical 1.12 mm pixel spacing) centered in the image, followed by an intensity clipping (10th and 96th percentiles, selected empirically to avoid clipping intensities in important structures) and a minmax normalization to 0–255 (uint8). The 2D mass center coordinates of the

LV and RV are provided by the 6 prediction outputs of the 1st NN, and a transformation is calculated from these to align the images to a smaller pre-defined template (144 × 144 pixels, 1.32 mm pixel spacing – balance between small size and high resolution while image large enough to cover the heart) centered in the heart. This same transformation is applied to each slice of the ES and ED SAx stacks. Details of the architecture and implementation of the 1st NN, based on a standard U-Net, can be checked in [22].

2.3 2nd NN: Fine Segmentation and Postprocessing

The transformed images, linear interpolated to the smaller template and normalized in intensities (5th and 93th percentiles clipping – cutoffs empirically adjusted to the intensities of the structures in the ROI space), are fed to the 2nd NN for a fine segmentation of the LV endo- and epicardium and the RV endocardium (Fig. 1, step 4–5). The LV myocardium is computed as the region between LV endo and epicardium, and the predictions are back interpolated to the original resolution and rearranged into the original 3D setting in a post-processing step (Fig. 1, step 5–6). This final step also accounts for segmentation quality enhancement (binarizing predictions, filling holes and removing stray clusters). Three architectures are proposed for the 2nd NN as described below:

U-Net. Baseline Method
A U-Net, fine-tuned for cardiac segmentations is applied for baseline comparisons. Details are available in [19].

Domain Adversarial Training of Neural Networks (DANN)
The DANN model, proposed by [25], consists of a feature extractor network with a domain predictor and label predictor. The gradient-reversal layer, placed between the feature extractor and the domain predictor, reverses the gradient direction during backpropagation and maximizes the domain prediction loss, thus minimizing the shift between the domains, while simultaneously making the model discriminative towards the main task of label prediction (Fig. 2).

Domain Unlearning (DU)
The DU model[1] [28] is based on the iterative unlearning framework [26] for adversarial adaptation, which rather than using a gradient reversal layer, optimizes two opposing loss functions in a sequential manner: one to maximize the performance of a label predictor given the fixed feature representation, and another to update the feature representation in order to maximally confuse the domain classifier (Fig. 2).

2.4 Implementation Details for DANN and DU

For the label predictor of the DANN and DU models, we used the Adam optimizer, with the same parameters as used in the baseline U-Net method. For the domain predictor in DANN and DU, we trained with the Momentum optimizer (with a momentum value of 0.9) and the Adam optimizer respectively. We used a batch size of 16, with a learning

[1] DU code available at: https://github.com/nkdinsdale/Unlearning_for_MRI_harmonisation.

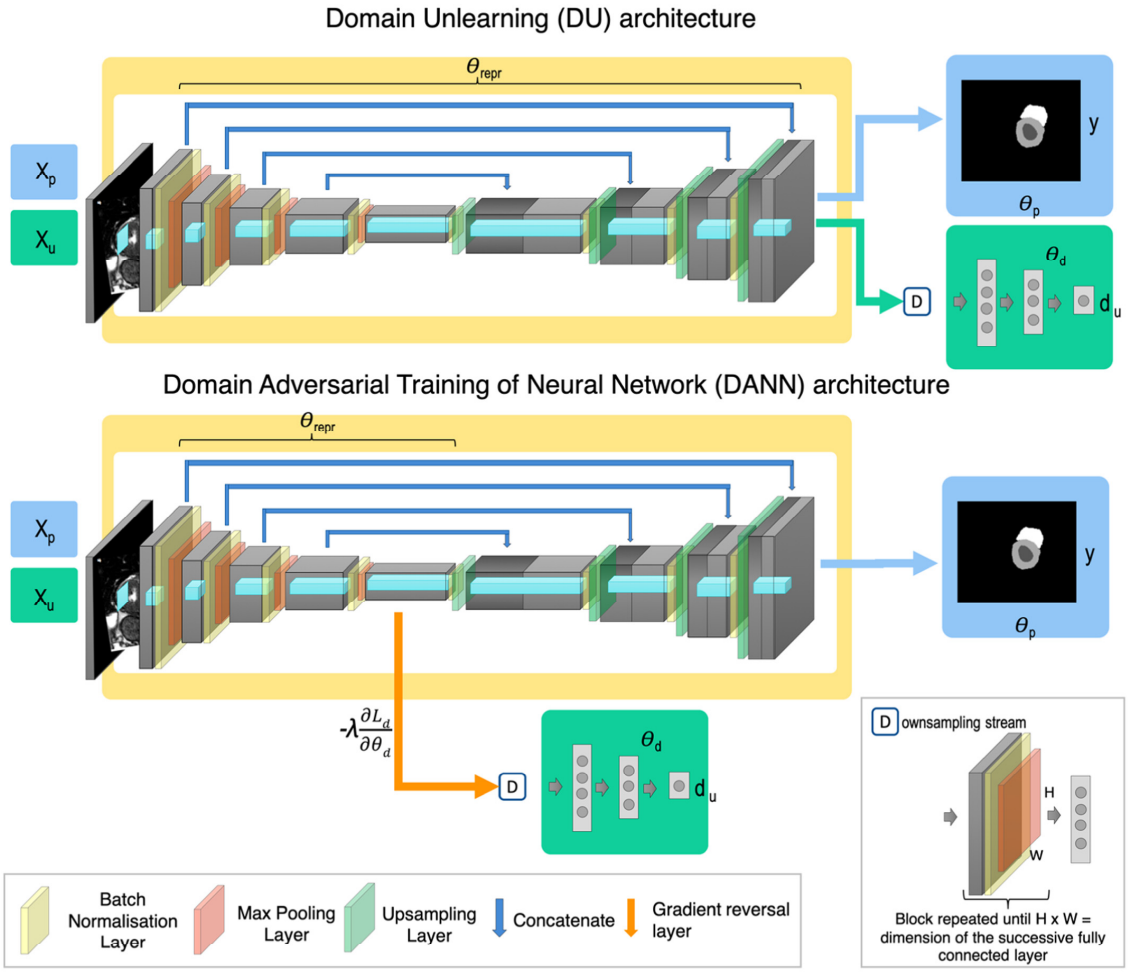


Fig. 2. Architectures of DU (top) and DANN (bottom) models. The models consist of the feature extractor, label predictor and domain predictor with corresponding training parameters θ_{repr}, θ_p and θ_d. The models take input features X_p and input domain information X_u and predicts output labels y, while unlearning output domains d_u. The DU model updates the label predictor, feature extractor and domain predictor in a sequential manner, while in DANN, label prediction and domain unlearning occur simultaneously.

rate of 10^{-3} and 10^{-5} for DANN and DU respectively. In addition, in DU we used a beta value of 1000 (a factor used for weighting the domain confusion a). These training hyper parameters were chosen empirically, and a criterion based on a patience (number of epochs to wait for progress on validation set) value of 25 epochs was used to determine model convergence (early stopping). Both models were run on an NVIDIA Tesla V100 GPU. The DANN and DU models took around 50 s and 45 s per epoch respectively for training.

2.5 Data Augmentation

Standard data augmentation based on random rotations (0 to 360°), translation (1st NN: ±20 mm. 2nd NN: ±8 mm) and flipping was applied during the training of any of the architectures. For each epoch, all the original training images are randomly transformed

as described prior to be fed to the NN. Thus, the originals are never used directly, and each instance fed to the NN is only seen once.

2.6 Performance Evaluation

Volumetric Dice scores between predicted and original segmentations in the native image space are calculated to assess the fine segmentation performance of the three proposed approaches.

3 Experiments and Results

The 1st NN was trained on the 150 labeled images (A and B vendors) following a 5-fold cross-validation strategy with a training-validation-testing split ratio of 107-13-30 images. The 5 resulting models were used in combination (majority voting) for the prediction of unseen test subjects.

To assess the quantitative performance of the 3 proposed methods on the 2nd NN fine segmentation, for seen and unseen vendors, we carried out 2 experiments:

- Training on the 150 labeled images from vendors A and B (A+B) and cross-validation evaluation on the same set, A+B (seen vendors assessment).
- Training on vendors A and C (A+C), and evaluation on vendor B (unseen vendor assessment). The baseline method was only trained on A, since unlabeled data (vendor C) cannot be incorporated into training, as done in the other 2 methods.

Table 1. Volumetric Dice results from the two proposed experiments. Data presented as median (interquartile range). The best performing models are highlighted in bold. Significant differences ($p < 0.05$) of mean values with respect to the baseline are marked with an asterisk (*).

Train	Test	Method	LV cavity	LV myocardium	RV cavity
A+B	5-fold validation A+B	BASELINE	0.927 (0.886–0.954)	0.845 (0.819–0.879)	0.889 (0.833–0.927)
		DANN	0.948 (0.918–0.964)*	0.873 (0.849–0.896)*	0.926 (0.897–0.943)*
		DU	**0.953 (0.927–0.965)***	**0.878 (0.855–0.897)***	**0.928 (0.898–0.945)***
A+C	B	BASELINE	0.913 (0.869–0.956)	0.829 (0.800–0.853)	**0.895 (0.841–0.930)**
		DANN	**0.927 (0.875–0.957)***	**0.838 (0.813–0.866)***	0.891 (0.845–0.936)*
		DU	0.914 (0.864–0.956)	0.835 (0.811–0.862)*	0.892 (0.842–0.934)

In both cases, and for all the 3 methods, we followed a 5-fold cross-validation approach with the split ratio described above. The results are shown in Table 1. While both DA techniques significantly improved the segmentation performance for seen vendors, the improvement in performance decreased for unseen vendors and only DANN reports a significantly better performance on average (*t*-test) than the baseline in segmenting any structure. Indeed, baseline is superior in RV median Dice under unseen vendors.

Finally, the 3 models were trained on A+B+C (only A+B for baseline), following the 5-fold training and split ratio described, and submitted to the M&Ms 2020 Challenge for external validation on unseen samples from vendors A, B, C and D. In addition, an ensemble with the combination in majority voting of the 15-resulting baseline, DANN and DU models (5folds x 3methods) was submitted. The external results, shown in Table 2, are consistent with our experiments and confirm the improvement in performance created by DA, especially for unseen vendors. The differences were not significant due to the

Table 2. Volumetric Dice results from the external validation set, stratified by vendor (median and interquartile range) and aggregated (average). The best performing models are highlighted in bold.

Method	Vendor	LV cavity	LV myocardium	RV cavity
Baseline	A	0.916 (0.857–0.957)	0.835 (0.808–0.879)	0.862 (0.713–0.896)
	B	0.951 (0.922–0.964)	0.864 (0.852–0.902)	0.942 (0.860–0.951)
	C	0.909 (0.828–0.952)	0.860 (0.827–0.888)	**0.870 (0.837–0.921)**
	D	0.911 (0.880–0.935)	0.828 (0.793–0.860)	0.871 (0.803–0.897)
	Average	0.904	0.844	0.852
DANN	A	**0.936 (0.876–0.961)**	**0.863 (0.827–0.889)**	0.878 (0.823–0.909)
	B	0.951 (0.923–0.968)	**0.885 (0.856–0.906)**	0.942 (0.895–0.957)
	C	0.903 (0.810–0.956)	0.860 (0.828–0.886)	0.844 (0.803–0.894)
	D	0.914 (0.874–0.944)	0.843 (0.798–0.869)	0.873 (0.816–0.908)
	Average	0.908	0.856	0.853
DU	A	0.918 (0.876–0.957)	0.857 (0.826–0.875)	0.878 (0.822–0.909)
	B	0.954 (0.922–0.970)	0.879 (0.853–0.902)	0.939 (0.890–0.951)
	C	0.898 (0.828–0.955)	0.862 (0.821–0.891)	0.856 (0.832–0.908)
	D	**0.926 (0.863–0.947)**	0.840 (0.807–0.870)	**0.882 (0.811–0.909)**
	Average	0.906	0.854	0.855
Ensemble	A	0.923 (0.866–0.961)	0.856 (0.825–0.886)	**0.886 (0.807–0.920)**
	B	**0.954 (0.923–0.967)**	0.882 (0.855–0.907)	**0.946 (0.884–0.956)**
	C	**0.911 (0.834–0.956)**	**0.865 (0.839–0.896)**	0.863 (0.833–0.920)
	D	0.922 (0.894–0.944)	**0.844 (0.801–0.867)**	0.878 (0.820–0.904)
	Average	**0.912**	**0.857**	**0.861**

size of the sets (only 10 patients per vendor). The best performance on average was obtained with the ensemble model in any of the structures.

Based on the previous results, the ensemble was submitted for final testing and M&Ms 2020 Challenge participation (remaining 160 unseen cases). Results along with a best and a worst case, provided by the Challenge organizers, in comparison with the rest of the participants are depicted below in Table 3 and Figs. 3 and 4. A drop in performance in comparison to the external validation set is observed, especially in vendor B.

Table 3. Volumetric Dice results from the final testing set, stratified by vendor (median and interquartile range) and aggregated (average). Significant differences ($p < 0.05$) of mean values with respect to the ensemble external validation results are marked with an asterisk (*).

Method	Vendor	LV cavity	LV myocardium	RV cavity
Ensemble	A	0.929 (0.887–0.960)	0.853 (0.810–0.875)	0.892 (0.832–0.925)
	B	0.933 (0.887–0.957)	0.866 (0.838–0.893)	0.903 (0.854–0.936)*
	C	0.924 (0.882–0.946)	0.845 (0.806–0.868)	0.887 (0.824–0.929)
	D	0.914 (0.875–0.929)	0.831 (0.784–0.867)	0.892 (0.847–0.922)
	Average	**0.902**	**0.833***	**0.863**

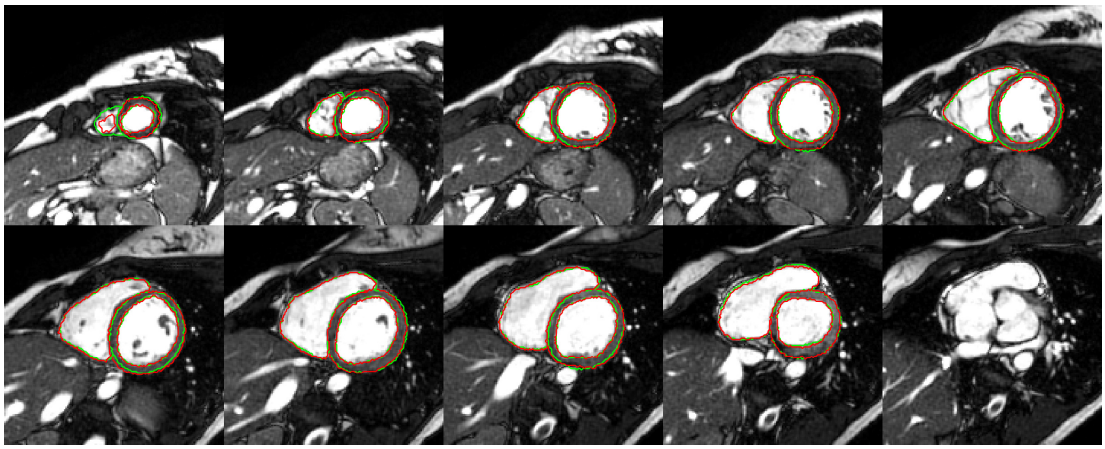


Fig. 3. Best case sample from the test set. Patient G1K1V3, vendor B, ED phase. Ground truth in green, ensemble prediction in red. Dice results of 0.979, 0.891 and 0.954 for LV cavity, LV myocardium and RV cavity, respectively. (Color figure online)

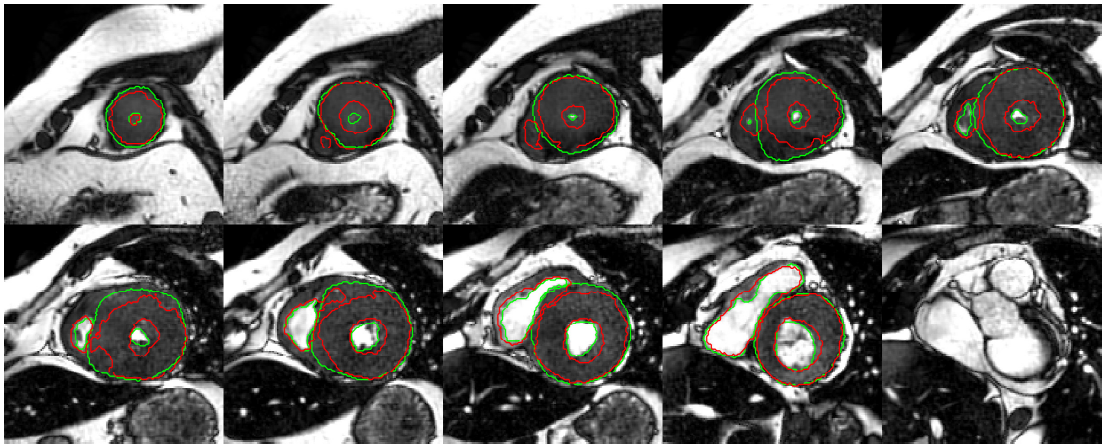


Fig. 4. Worst case sample from the test set. Patient L8N7Z0, vendor B, ES phase. Ground truth in green, ensemble prediction in red. Dice results of 0.653, 0.795 and 0.771 for LV cavity, LV myocardium and RV cavity, respectively. (Color figure online)

4 Discussion and Conclusions

Our contributions in this work are as follows: (1) we have developed a 2-step DL-based completely automatic pipeline for segmenting LV cavity, LV myocardium and RV cavity on CMR images; (2) we have implemented two DA techniques, domain adversarial training and domain unlearning, to overcome image heterogeneity and scanner-related variations; and (3) we have externally validated our methods on a challenging dataset, including unseen scanners, centers and pathologies, obtaining mean Dice scores of 0.902, 0.833 and 0.863 for LV cavity, LV myocardium and RV cavity, respectively.

The results show that the implemented DA methods significantly improve segmentation performance on seen labelled vendors. However, the 2 proposed DA methods only achieved comparable or slightly superior performance to baseline on the unseen vendors, and so the potential improvement that these 2 methods offer remains inconclusive for multi-vendor applications.

The DA model improvements for seen vendors, especially with DU, illustrated by the A+B experiment in Table 1 can be explained by the fact that both DA models learn a generic feature representation (by unlearning domain-specific information) enabling information from both domains to be incorporated. This essentially increases the size of the training dataset and its variability, improving the label prediction. We believe that the DU model provided better results compared to DANN since the minmax optimization (with gradient reversal) could occasionally become unstable and get trapped in local maxima. On the other hand, feature representations are updated to unlearn domain-specific information with each iteration in DU models. Thus, the label predictor is consistently improved using these feature representations, making the DU model comparatively more stable.

However, despite training DANN and DU models with data from vendor C, there was a lack of improvement on that vendor when compared to baseline results, according to the external validation experiments. A plausible explanation is the additional variability introduced by unseen pathological cases and centers within the same vendor domain. Further investigation is required, after the test data is released, to understand this behavior

and to propose changes accordingly via domain shift analysis and hyperparameter fine-tuning.

Comparing the A+B and A+C experiment results, the relatively small drop in performance in the baseline models on the prediction of vendor B suggests the robustness of the baseline method for unseen domain prediction. This is further confirmed by the external validation experiments (see predictions on vendor C and D, Table 2). Credit for this robustness should be given to the 2-step implementation, which normalizes for orientation, resolution and appearance and crops the original images to the ROI, levering the label imbalance (background vs structure) towards segmentation enhancement.

In an ensemble model, different models are trained with different initializations and each of them learn different aspects of the data. For instance, while the baseline model learns the salient features for label prediction, DANN and DU learn domain invariant features, still retaining task-specific features. Hence, we achieved better overall results in the external validation dataset by combining the predictions using an ensemble model compared to any of the individual models. The dice scores obtained by this model ensemble in the validation dataset (see Table 2) are comparable to those reported in [21], trained on 4875 theoretically healthy subjects, while we only used 150 labelled subjects for training and evaluated on data from multiple centers and vendors, including pathological cases.

While according to the M&Ms 2020 Challenge description the external validation (40 cases) and the final testing set (160 cases) come from the same test set, including samples from different vendors, centers and pathologies in the same proportion, significant differences in performance are obtained for the myocardium structure and for vendor B. A plausible explanation is that the size of the validation set might not be large enough to cover the variability of the test set. We believe that the discrepancies between the testing and validation set across participants is worth exploring to check this hypothesis and to improve the representativeness of the sets in future editions of the Challenge.

To sum up, we have evaluated our method and explored various DA techniques on heterogeneous data from multiple vendors and centers, achieving the best performance with the ensemble of DA and baseline models. We have made our methods publicly available as Singularity containers that may serve as an independent testing tool for the community[2]. Future directions include further exploration of DA models for improving model generalizability.

Acknowledgements. The authors of this paper declare that the segmentation method they implemented for participation in the M&Ms challenge has not used any pre-trained models nor additional MRI datasets other than those provided by the organizers. This work was supported by the EU's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie (g.a. 764738 to JCA). VS is supported by Wellcome Centre for Integrative Neuroimaging, which has core funding from the Wellcome Trust (203139/Z/16/Z). ND is supported by the Engineering and Physical Sciences Research Council and Medical Research Council (EP/L016052/1). MJ is

[2] The CMR SAx Segmentation Models presented in this paper are available at: https://figshare.com/articles/software/A_2-step_Deep_Learning_method_with_Domain_Adaptation_for_Multi-Centre_Multi-Vendor_and_Multi-Disease_Cardiac_Magnetic_Resonance_Segmentation/13140554.

supported by the National Institute for Health Research (NIHR), Oxford Biomedical Research Centre (BRC) and Wellcome Trust (215573/Z/19/Z). The computational aspects of this research were supported by the Wellcome Trust Core Award (203141/Z/16/Z) and the NIHR Oxford BRC. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Stokes, M.B., Roberts-Thomson, R.: The role of cardiac imaging in clinical practice. Aust. Prescr. **40**(4), 151–155 (2017)
2. Peterzan, M.A., Rider, O.J., Anderson, L.J.: The role of cardiovascular magnetic resonance imaging in heart failure. Card. Fail. Rev. **2**(2), 115–122 (2016)
3. Peng, P., Lekadir, K., Gooya, A., Shao, L., Petersen, S.E., Frangi, A.F.: A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. Magn. Reson. Mater. Phys., Biol. Med. **29**(2), 155–195 (2016). https://doi.org/10.1007/s10334-015-0521-4
4. Dall'Armellina, E.: From recognized to novel quantitative CMR biomarkers of LV recovery. JACC Cardiovasc. Imaging **10**(9), 1000–1002 (2017)
5. Desch, S., et al.: Cardiac magnetic resonance imaging parameters as surrogate endpoints in clinical trials of acute myocardial infarction. Trials **12**, 204 (2011)
6. Reindl, M., Eitel, I., Reinstadler, S.J.: Role of cardiac magnetic resonance to improve risk prediction following acute ST-elevation myocardial infarction. J. Clin. Med. **9**(4), 1041 (2020)
7. Bulluck, H., Dharmakumar, R., Arai, A.E., Berry, C., Hausenloy, D.J.: Cardiovascular magnetic resonance in acute ST-segment–elevation myocardial infarction. Circulation **137**(18), 1949–1964 (2018)
8. Anderson, J.L., Morrow, D.A.: Acute myocardial infarction. N. Engl. J. Med. **376**(21), 2053–2064 (2017)
9. Zhang, S., Li, L., Guo, W., Wang, L., Yan, X., Zhang, C.: Observation on left ventricular remodeling in acute myocardial infarction. Zhonghua nei ke za zhi **38**(2), 107–109 (1999)
10. Petersen, S.E., et al.: Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort. J. Cardiovasc. Magn. Reson. **19**(1), 18 (2017)
11. Suinesiaputra, A., et al.: Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours. J. Cardiovasc. Magn. Reson. **17**(1), 63 (2015)
12. Ibanez, B., et al.: 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: the task force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Socie. Eur. Heart J. **39**(2), 119–177 (2017)
13. Rumsfeld, J.S., Joynt, K.E., Maddox, T.M.: Big data analytics to improve cardiovascular care: promise and challenges. Nat. Rev. Cardiol. **13**(6), 350–359 (2016)
14. Shameer, K., Johnson, K.W., Glicksberg, B.S., Dudley, J.T., Sengupta, P.P.: Machine learning in cardiovascular medicine: are we there yet? Heart **104**(14), 1156–1164 (2018)
15. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
16. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H., Aerts, H.J.W.L.: Artificial intelligence in radiology. Nat. Rev. Cancer **18**(8), 500–510 (2018)

17. Hussain, Z., Gimenez, F., Yi, D., Rubin, D.: Differential data augmentation techniques for medical imaging classification tasks. In: AMIA Symposium, vol. 2017, pp. 979–984 (2017). Accessed 20 Jan 2019
18. Perez, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning. CoRR, vol. abs/1712.0 (2017)
19. Corral Acero, J., Zacur, E., Xu, H., Ariga, R., Bueno-Orovio, A., Lamata, P., Grau, V.: SMOD - data augmentation based on statistical models of deformation to enhance segmentation in 2D cine cardiac MRI. In: Coudière, Y., Ozenne, V., Vigmond, E., Zemzemi, N. (eds.) FIMH 2019. LNCS, vol. 11504, pp. 361–369. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21949-9_39
20. Lemley, J., Bazrafkan, S., Corcoran, P.: Smart augmentation learning an optimal data augmentation strategy. IEEE Access **5**, 5858–5869 (2017)
21. Bai, W., et al.: Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. J. Cardiovasc. Magn. Reson. **20**(1), 65 (2018)
22. Corral Acero, J., et al.: Left ventricle quantification with cardiac MRI: deep learning meets statistical models of deformation. In: Pop, M., et al. (eds.) STACOM 2019. LNCS, vol. 12009, pp. 384–394. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39074-7_40
23. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2010)
24. Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., Vaughan, J.W.: A theory of learning from different domains. Mach. Learn. **79**(1–2), 151–175 (2010)
25. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., et al.: Domain-adversarial training of neural networks. J. Mach. Learn. Res. (2015)
26. Hoffman, J., Tzeng, E., Darrell, T., Saenko, K.: Simultaneous deep transfer across domains and tasks. In: Csurka, G. (ed.) Domain Adaptation in Computer Vision Applications. ACVPR, pp. 173–187. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58347-1_9
27. Alvi, M., Zisserman, A., Nellåker, C.: Turning a blind eye: explicit removal of biases and variation from deep neural network embeddings. In: Leal-Taixé, L., Roth, S. (eds.) ECCV 2018. LNCS, vol. 11129, pp. 556–572. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11009-3_34
28. Dinsdale, N.K., Jenkinson, M., Namburete, A.I.L.: Unlearning Scanner Bias for MRI Harmonisation in Medical Image Segmentation, pp. 15–25. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59713-9_36
29. Campello, V.M.: Multi-centre, multi-vendor & multi-disease cardiac image segmentation challenge (M&Ms) MICCAI 2020. https://www.ub.edu/mnms/. Accessed 1 July 2020
30. Yang, G., Hua, T., Xue, W., Shuo, L.: Left ventricle full quantification challenge MICCAI 2019. https://lvquan19.github.io/. Accessed 1 July 2020

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation
Publication Status	☑ Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	Jorge Corral Acero, Vaanathi Sundaresan, Nicola Dinsdale, Vicente Grau, Mark Jenkinson, "A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation", In: Puyol Anton E. et al. (eds) Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges. *STACOM 2020*. Lecture Notes in Computer Science, vol 12592. Springer, Cham; 2021. p. 196–207. DOI: https://doi.org/10.1007/978-3-030-68107-4_20

Student Confirmation

Student Name:	Jorge Corral Acero		
Contribution to the Paper	(1) Designed the experiments; (2) Developed and implemented the 2-step architecture and the baseline methods; (3) pre-processed the data; (4) designed and implemented the postprocessing of the segmentations; (5) co-ran the experiments; (5) analysed the results; (6) prepared figures and tables; and (6) led the write-up of the manuscript. The contribution of Dr. Vaanathi Sundaresan, co-author of this manuscript, shall be acknowledged.		
Signature		Date	12/07/2021

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. Vicente Grau		
Supervisor comments: I confirm that the above statement is accurate.		
Signature 	Date	12/07/2021

This completed form should be included in the thesis, at the end of the relevant chapter.

CHAPTER 5

Analysis of the Heart Anatomy from 3D Meshes: Impact on the AMI Case Study

Research Objective 2

5. Analysis of the Heart Anatomy from 3D Meshes: Impact on the AMI Case Study

Research Objective 2

5.1. Chapter Publications

- **Paper V**: J. Corral Acero et al. "Understanding and Improving Risk Assessment after Myocardial Infarction using Automated Left Ventricular Shape Analysis". Under 3rd revision in *JACC Cardiovasc Imaging.*

5.2. Chapter Outline

This chapter explores the potential of 3D shape analysis towards gaining a better understanding of the anatomy of the heart, and how risk is accordingly modulated in CVD. To do so, the chapter focuses on AMI, the deadliest cardiac disease in the world. Acute changes after an infarct have proven crucial to early prognosis [8], [11], [12], [151]. CMR is the gold-standard for the assessment of acute remodelling in AMI, as discussed in Chapter 2. LVEF and ESV are the most established imaging biomarkers in current clinical practice for post-infarction risk stratification and therapeutic decision-making. Nevertheless, they are limited to global function and fail to capture local abnormalities [7], [11], [199], [207]. To overcome these limitations, Paper V proposes the disentanglement of these stand-alone markers into 3D alternative descriptions of contraction and ES shape, respectively.

The fully automated pipeline developed in Chapter 4 was executed on a large AMI cohort to obtain the personalised mesh of each of the patients. Point correspondence between 3D meshes is directly obtained thanks to the meshing methodology deployed. This design principle allows PCA to be directly applied in order to condense the anatomical information of the meshes (several thousand of points) into a reduced number of variables (PCA modes). These modes are interpretable and account for the main shape variations observed within the population. This is explained in detail in Appendix A. The same applies to 3D contraction, approximated as the displacement between ED and ES. By

finding what PCA modes are prognostic (i.e., by means of linear discriminant analysis or Cox assessment), we are implicitly identifying 3D shape and contraction patterns related to risk.

The analysis described above resulted in the identification of three ES shape and three 3D contraction features associated with major adverse cardiac events (MACE), prescribed as the clinical endpoints of the longitudinal study. The prognostic information of these features proved superior to ESV and LVEF, respectively. Moreover, these novel markers demonstrated complementary to other biomarkers. They contributed to a significant prognostic improvement in a multivariable setting, including CMR markers, cardiovascular risk factors, and basic clinical patient characteristics. This performance gain, as well as the robustness of the automated pipeline, was further validated by a reproducibility analysis. The improvement in AMI stratification, along with an automation of the analysis that completely removes the manual burden, argues for the impact and clinical relevance of the work conducted in this chapter.

The study concludes with the investigation of the interplay between the identified 3D LV patterns and other clinical parameters to further understand post-infarction remodelling. Three of these novel biomarkers were found consistent with the three predominant LV remodelling descriptors according to the literature: ESV, LVEF, and ES myocardial thickness. Another two features showed alternative manifestations of anterior impairment linked to MACE, suggesting a perfusion impairment in the left anterior descending artery and LV akinesia. Finally, the sixth feature pointed to contraction impairment in the basal region as highly related to adverse events. It is hypothesised that this could be caused by circumferential fibre damage, in contrast to the endocardial vertical fibre injury that lead to decreased LVEF, as one of the other modes showed. Moreover, the relation between these myocardial 3D shape changes and micro-damage was investigated, showing a strong correlation for some of the modes.

As additional output from the study, a shape and contraction atlas has been built from 1000+ AMI patients. It is worth mentioning the potential of this atlas for applications ranging from alternative

AMI shape studies and computer simulations up to synthetic patient data generation for training algorithms or educational use.

Paper V is currently under 3rd revision in *JACC Cardiovascular Imaging*. The feedback received during the revision process has contributed to further (a) improve the reporting of the results, (b) strengthen the validation of the findings, (c) ease the interpretability and (d) enrich the discussion on limitations and future work. The supplemental material of this manuscript can be found in Appendix A, where the methods are further detailed, and additional results are shown.

Paper V

Understanding and Improving Risk Assessment after Myocardial Infarction using Automated Left Ventricular Shape Analysis

Jorge Corral Acero, Andreas Schuster, Ernesto Zacur, Torben Lange, Thomas Stiermaier, Sören J. Backhaus, Holger Thiele, Alfonso Bueno-Orovio, Pablo Lamata, Ingo Eitel, Vicente Grau

Submitted to *JACC Cardiovasc Imaging* on 26th May 2021

TITLE PAGE

Title

Understanding and Improving Risk Assessment after Myocardial Infarction using Automated Left Ventricular Shape Analysis.

Brief title

LV 3D-analysis for post-infarction assessment

Authors

Jorge Corral Acero[a,✉], Andreas Schuster[b], Ernesto Zacur[a], Torben Lange[b], Thomas Stiermaier[c], Sören J. Backhaus[b], Holger Thiele[d], Alfonso Bueno-Orovio[e], Pablo Lamata[f], Ingo Eitel[c*], Vicente Grau[a*].

(*) Equally co-supervised the work.

(✉) Corresponding author.

Affiliations

a. Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK.

b. University Medical Centre Göttingen, Department of Cardiology and Pneumology, Georg-August University, Göttingen, Germany; and German Centre for Cardiovascular Research (DZHK), partner site Göttingen, Göttingen, Germany.

c. University Heart Centre Lübeck, Medical Clinic II, Cardiology, Angiology and Intensive Care Medicine, Lübeck, Germany; University Hospital Schleswig-Holstein, Lübeck, Germany; and German Centre for Cardiovascular Research (DZHK), partner site Hamburg/Kiel/Lübeck, Lübeck, Germany.

d. Heart Centre Leipzig at University of Leipzig, Department of Internal Medicine and Cardiology, Leipzig, Germany; and Leipzig Heart Institute, Leipzig, Germany.

e. Department of Computer Science, University of Oxford, Oxford, UK.

f. Dept. of Biomedical Engineering, King's College London, London, UK.

Address for correspondence

Full name: Jorge Corral Acero
Email address: jor.corral@eng.ox.ac.uk
Telephone number: +34646346222
Postal address: University of Oxford, Old Road Campus Research Building, Headington, Oxford OX3 7DQ (UK).

Funding

This work was supported by the EU's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie (g.a. 764738), by the DZHK (German Center for Cardiovascular Research) and by the British Heart Foundation (PG/16/75/32383, FS/17/22/32644), and Wellcome Trust (209450/Z/17).

Conflict of interest: none declared.

Article specifications

- ✓ Article length: 3,978 words
- ✓ Article length including body text, references and figure legends: 5,050 words
- ✓ Number of references: 24
- ✓ Number of figures and tables: 5

ABSTRACT AND KEYWORDS

Objectives

This study aims at the identification of novel 3D imaging end-systolic (ES) shape and contraction descriptors towards risk-related features and superior prognosis in acute myocardial infarction (AMI).

Background

Left ventricular ejection fraction (LVEF) and end-systolic volume (ESV) remain the main imaging biomarkers for post-AMI risk stratification. However, they are limited to global systolic function and fail to capture functional and anatomical regional abnormalities, hindering their performance in risk stratification.

Methods

A multicenter cohort of AMI survivors (n=1021; median age: 63 years; 74.5% male) who underwent cardiac magnetic resonance (CMR) at a median of 3 days after infarct was considered for this study. The clinical endpoint was the 12-month rate of major adverse cardiac events (MACE, n=73), consisting of all-cause death, reinfarction, and new congestive heart failure. A fully automated pipeline was developed to segment CMR images, build 3D statistical models of shape and contraction in AMI, and find the 3D patterns related to MACE occurrence.

Results

The novel ES Shape markers proved superior to ESV (0.681 (0.679 - 0.684) vs 0.600 (0.598 - 0.602), P<0.001); and 3D contraction, to LVEF (0.716 (0.714 - 0.718) vs 0.681 (0.679 - 0.684), P<0.001) in MACE occurrence prediction (cross-validated AUC). They also contributed to a significant improvement in a multivariable setting including CMR markers, cardiovascular risk factors and basic patient characteristics (0.747 (0.745 - 0.749), P<0.001). Based on these novel 3D descriptors, three impairments caused by AMI were identified: global, anterior, and basal, being the latter the most complementary signature to already known predictors.

Conclusions

The quantification of 3D differences in ES shape and contraction, enabled by a fully automated pipeline, improves post-AMI risk prediction, and identifies shape and contraction patterns related to MACE occurrence.

Clinical Trial

The clinical trials that this study is based on, AIDA-STEMI (NCT00712101) and TATORT-NSTEMI (NCT01612312), were registered with ClinicalTrials.gov and approved by the ethics committees of the participating sites. Registry URLs: AIDA-STEMI: https://clinicaltrials.gov/ct2/show/NCT00712101. TATORT-NSTEMI: https://clinicaltrials.gov/ct2/show/NCT01612312.

Keywords

Cardiovascular magnetic resonance, acute myocardial infarction, risk assessment, 3D shape analysis, left ventricle, fully automated analysis.

CONDENSED ABSTRACT

This multicenter CMR study, including 1000+ acute myocardial infarction patients, investigates the role of left ventricular (LV) 3D shape and contraction in modulating post-infarction risk. The main contributions are three-fold. Firstly, it proves the superior prognostic information of 3D CMR analysis against the most-established post-infarction risk predictors, ejection fraction and end-systolic volume. Secondly, it identifies the 3D shape and contraction signatures related to risk. Finally, it proposes a fully automated pipeline for accurate CMR segmentation and quantification of LV indices (both established and proposed 3D-based descriptors), overcoming the need of manual annotation.

ABBREVIATIONS LIST

Abbreviations and acronyms

AMI	Acute Myocardial Infarction
CMR	Cardiac Magnetic Resonance
ED	End-diastole
ES	End-systole
IS	Infarct Size
LV	Left Ventricle
LVEF	Left Ventricular Ejection Fraction
MACE	Major Adverse Cardiac Events
MVO	Microvascular Obstruction
PCA	Principal Component Analysis
r_s	Spearman correlation coefficient
SAx	Short Axis Stacks

INTRODUCTION

Preventive and personalized medicine is crucial to improve both the efficiency and efficacy of healthcare systems (1). Advances have significantly improved the prognosis of acute myocardial infarction (AMI) patients worldwide (2). Nevertheless, AMI survivors are at risk of recurrent cardiovascular events, and mortality within 6 months sits at 12% (3).

Cardiac magnetic resonance (CMR) imaging has proven uniquely suitable to assess morphological and functional myocardial alterations, including left ventricular (LV) remodeling, central to AMI early prognosis prediction (4). LV micro-myocardial injury, assessed by late gadolinium enhancement CMR and quantified as infarct size (IS) and microvascular obstruction (MVO), has emerged as a robust measure in AMI risk assessment (4,5). Nonetheless, LV macro-function, typically quantified as LV ejection fraction (LVEF), remains the preferred image biomarker in AMI guidelines (3,4). Patients with pumping function deterioration, at higher risk of severe cardiac events, also exhibit an increase in end-systolic volume (ESV), reported in several studies as superior to LVEF in predictive power (6); while

patients with early revascularization and preserved pumping function are characterized by an increase in LV wall thickening in both the infarcted and remote regions (7).

LVEF and ESV are limited to global systolic function and fail to capture functional and anatomical regional abnormalities (4,6,7). Moreover, LVEF is mainly preserved or only moderately reduced in most AMI survivors and therefore recurrent adverse events often occur in patients at a theoretical low risk, which further evidences the need of stratification improvements (8). To overcome such limitations, unravelling of the global LV remodeling changes at the different regions of the ventricle, i.e. 17-segment AHA model, has been proposed (9). High dimensional 3D shape analysis is also used in cardiac research to describe subtle focal changes (10).

This study aims at the disentanglement of ESV and LVEF into 3D shape and contraction using only standard CMR short axis stacks (SAx) to (a) gain a better understanding of how 3D remodeling patterns modulate risk and the interplay between macro-remodeling and micro-damage, and (b) improve AMI risk stratification. We present a fully automated shape analysis pipeline including state-of-the-art artificial intelligence and 3D meshing and its application on a large multicenter population including ST-segment–elevation myocardial infarction (STEMI) and non-STEMI (NSTEMI) patients.

METHODS

Study population

Our retrospective population consists of 1235 AMI survivors (8) from two multicenter randomized trials, AIDA-STEMI (Abciximab Intracoronary versus intravenously Drug Application in ST-Elevation Myocardial Infarction) and TATORT-NSTEMI (Thrombus Aspiration in Thrombus Containing Culprit Lesions in Non-ST-Elevation Myocardial Infarction) (11,12). In both studies, reperfusion therapy with primary percutaneous coronary intervention and postinfarction medical treatment were supplied according to state-of-the-art guideline recommendations (3). The size of the study is argued in Supplemental Material VIII, based on (13).

AIDA-STEMI (NCT00712101) and TATORT-NSTEMI (NCT01612312) were registered with ClinicalTrials.gov, approved by the ethics committees and complied with the Declaration of Helsinki, including written informed consent. Data supporting our findings is available upon reasonable request.

CMR imaging protocol and manual analysis

The same standardized protocol, including ECG-gated balanced steady-state free precession sequences and T1-weighted late gadolinium enhanced images, was followed for all AMI patients on 1.5T or 3T clinical scanners (11,12). Horizontal and vertical long-axis views as well as continuous stacks of SAx slices capturing the whole LV were acquired in all sequences. Infarct characteristics, ventricular volumes and LVEF were determined in sequential SAx by blinded clinicians using certified evaluation software (cmr42; Circle Cardiovascular Imaging) (12). IS and MVO were assessed in (12), applying standard thresholding techniques.

Study endpoints

The 12-month rate of MACE, consisting of reinfarction, new congestive heart failure and all-cause death, was the predefined clinical endpoint of the study as detailed in (8,11,12). The events were adjudicated by a blinded clinical committee based on the data collected in the study sites. Only one contribution per patient to the endpoint composite was considered in case of multiple MACE events per patient (death>reinfarction>congestive heart failure).

LV fully automated shape analysis

The LV myocardium was segmented in the SAx images and two personalized 3D LV meshes were built for each patient, at end-diastole (ED) and end-systole (ES). The displacements between corresponding points in them constituted the 3D contraction. 3D shape and contraction descriptors were obtained by unsupervised construction of statistical shape models, and used as potential predictive biomarkers. The process (Fig. 1) is fully automated.

LV Segmentation

A 2-step deep learning approach was applied to segment the LV myocardium, based on a UNet architecture with enhanced pre-processing (14,15), that reached the best performance in the 2019 LV Full Quantification Challenge (15). ES and ED phases are detected as those of minimal and maximal LV blood pool area at the mid-ventricular slice. Implementations details are available in Supplemental Material I.

Mesh generation

The construction of 3D meshes from segmentations contours used a solution based on smooth cubic Hermite interpolation, shown to be accurate and robust to SAx slice misalignment and segmentation errors and fully described in (16). In brief, an idealized LV template mesh was fitted to the 3D myocardium mask by combining image registration and mesh projection techniques. Meshing results were evaluated for accuracy using the 3D distance between segmentation contours and mesh surfaces. The anatomical correspondence between meshes was achieved by orienting the LV meshes according to the SAx canonical position.

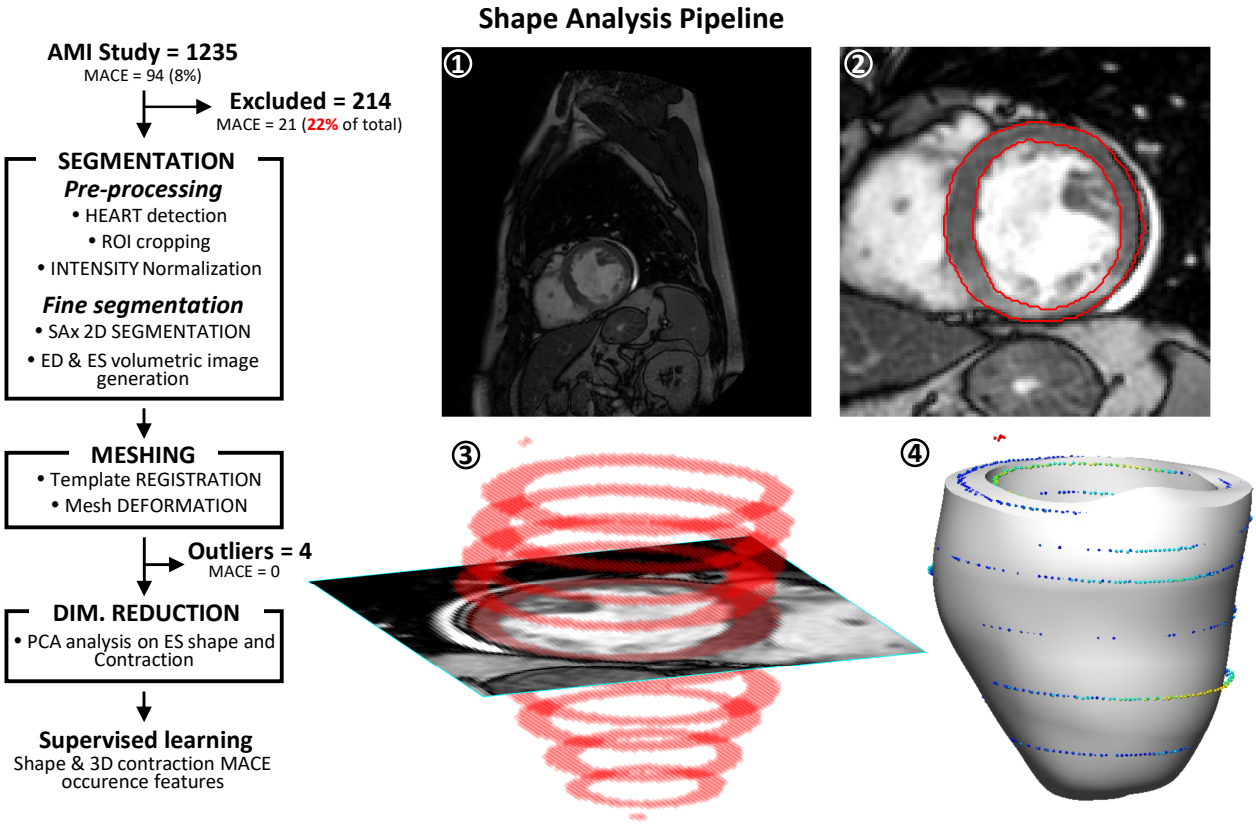


Fig. 1 - **Shape analysis pipeline.** Shape analysis pipeline (left) along sample case (right), consisting of the following steps: (1) SAx stack pre-processing (intensity normalization, heart detection, cropping of region of interest); (2) fine segmentation of LV myocardium; (3) volumetric segmentation at ES and ED phases; (4) personalized mesh generation. 3D contraction is then calculated from the ES and ED meshes, and PCA dimensionality reduction is applied on the meshes to facilitate the supervise learning of MACE occurrence.

Dimensionality reduction

The direct analysis of the 3D meshes, discretized into 2,450 nodes with their corresponding 3D spatial coordinates, is impractical and ill-posed. This motivates the use of principal component analysis (PCA) for dimensionality reduction (10). In short, PCA inspects the data to find the directions of change that maximize the variance observed in the population. The resulting main directions represent the 3D shape patterns of change with respect to the mean shape, which are called anatomical 'PCA modes'. Conceptually, the LV of each subject is decomposed into the mean shape (average ventricle) plus the anatomical modes (shape variations – i.e. thickening, scaling, lengthening, etc.) times a coefficient. Each of the modes is therefore a continuous variable that accounts for a particular pattern of 3D shape variation, which has a certain value for each patient and becomes the potential biomarker to predict MACE (See Fig. 2). The formulation details are explained in the Supplemental Material II.

Three statistical shape models were built with this PCA technique, studying the 3D shape at ED, at ES, and the displacement between them (contraction). PCA coefficients of ED were found to contain no predictive value and are therefore not reported.

Supervised learning

The relevant PCA modes to predict MACE were found by a stepwise multivariable Fisher linear discriminant analysis (LDA), as detailed below. The search was performed on the ES shape and on the contraction modes needed to reconstruct 95% of the total variance of the population. This selection was further refined with another LDA applied on the combination of ES and contraction significant modes found in the previous step.

Evaluation experiments

The predictive value of the novel ES shape and contraction PCA modes was evaluated against standard clinical biomarkers, and in combination with them. Conceptually, ES and contraction PCA modes are the 3D extension of the ESV and LVEF biomarkers, and a first test compared them. Multivariable combinations were next investigated, considering only CMR biomarkers (CMR) and all the clinical variables of the study (ALL).

Comparisons were based on the prediction performance resulting from both a multivariable LDA (binary classification – MACE occurrence at 12 months) and a multivariable Cox analysis (time to MACE). A backward stepwise strategy (unbiased to univariable associations (17)) was followed in both analyses to find inter-variable synergies and address collinearity: all variables of interest were initially included, and the less significant in the model iteratively removed, until a 0.05 *p*-value threshold was met by all of them. Significance in the LDA was approximated as significance in the generalized linear binomial sigmoid regression model. To account for both specificity and sensitivity and class imbalance, the performance of the resulting configurations was assessed via AUC (area under the receiver-operator characteristic curve) for binary classifications and concordance index (C-index) for time-dependent curves.

The generality of findings was evaluated by comparison of performance in the resubstitution (apparent performance - learning and testing with the entire cohort) and cross-validation (optimism corrected performance - 10 cross-fold-validation repeated for a hundred random data splits). To further strengthen this generalization, an additional experiment trained the models only on AIDA-STEMI patients and tested them on the TATORT-NSTEMI independent cohort (see Supplemental Material IV). A reproducibility analysis, where the ED and ES frames were removed to test the robustness of the new methodology, is included in the Supplemental Material VI.

Statistical analysis

The variables of the study (baseline characteristics, cardiovascular risk factors, CMR biomarkers, and significant ES and contraction PCA coefficients) are described according to MACE occurrence. Continuous variables are presented as median and interquartile range (IQR), since they were not normally distributed in a Shapiro–Wilk test except for the modes of variation (Gaussian distributed by definition). Categorical variables are provided as frequencies and percentages. The MACE and No MACE distributions, as well as the cross-validated AUC distributions, were compared by the non-parametric Wilcoxon rank sum test. Univariate Cox regression analyses were also performed. Hazard ratios, 95% confidence intervals and the predictor significance are presented in the results. All analyses were implemented in Matlab R2019b. The study follows the guidelines of transparent reporting of multivariable prediction model for individual prognosis or diagnosis (TRIPOD, see Supplemental Material VIII) and Prediction model Risk Of Bias ASsessment Tool (PROBAST) for model development (17,18).

RESULTS

Patients

A total of 1021 AMI patients (STEMI: n=723; NSTEMI: n=298) of the initial 1235 cohort had both original CMR scans and 12-month follow-up data available and were included in this study (No CMR: n=126; incomplete protocol: n=86; no follow-up: n=2) (8,11,12). They underwent CMR a median of 3 (IQR: 2–4) days after infarction and presented a total of 73 MACEs (congestive heart failure: n=20; reinfarction: n=21; death: n=32) (11,12).

CMR and baseline clinical characteristics are described in Table 1. As reported in (8,11,12), the population was predominantly male, with a median age of 63, LVEF of 50.5% and IS of 13.4% of LV mass, with minor MVO. The MACE subgroup was significantly older, with higher body surface area, a lower percentage of smokers, and larger prevalence of hypertension and diabetes mellitus. They also presented larger ESV, IS, MVO, Killip class on admission and number of diseased vessels; and lower LVEF.

Shape analysis

The automated segmentation of the SAx, including all the planes and phases, resulted in a median Dice score of 0.971 (IQR: 0.945-0.981) and 0.975 (IQR: 0.959-0.982) for LV epicardium and endocardium, respectively. The Spearman correlation coefficients (r_s) between the volumes calculated from automated and manual segmentations were 0.916, 0.919 and 0.842 for ESV, EDV and LVEF, respectively. The median distances between mesh surfaces and contours were 0.981 mm (IQR: 0.806-1.22) for ES and 1.02 mm (IQR: 0.828-1.31) for ED. The r_s coefficients between the manual volumes and those integrated from the mesh were 0.905, 0.914 and 0.792 for ESV, EDV and LVEF, correspondingly. The resulting LVEF calculated from the meshes was negative for 4 patients because of noise in the mid-ventricular segmentation that led to an incorrect phase selection. They were excluded from the statistical analysis (Fig. 1).

95% of the variance was captured by the first 13 ES shape and the 21 first contraction modes of variation in PCA analyses, respectively. The LDA stepwise analysis on the combination of ES and contraction significant modes converged to ES modes 1, 5, 6 and Contraction modes 3, 5, 16 (Fig. 2 and Supplemental Videos I and II). These modes are interpreted as basal (C16), anterior (ES5, C5) and global impairment, in the form of increased ESV (ES1), decreased LVEF (C3) and impaired thickening (ES6) (see Discussion).

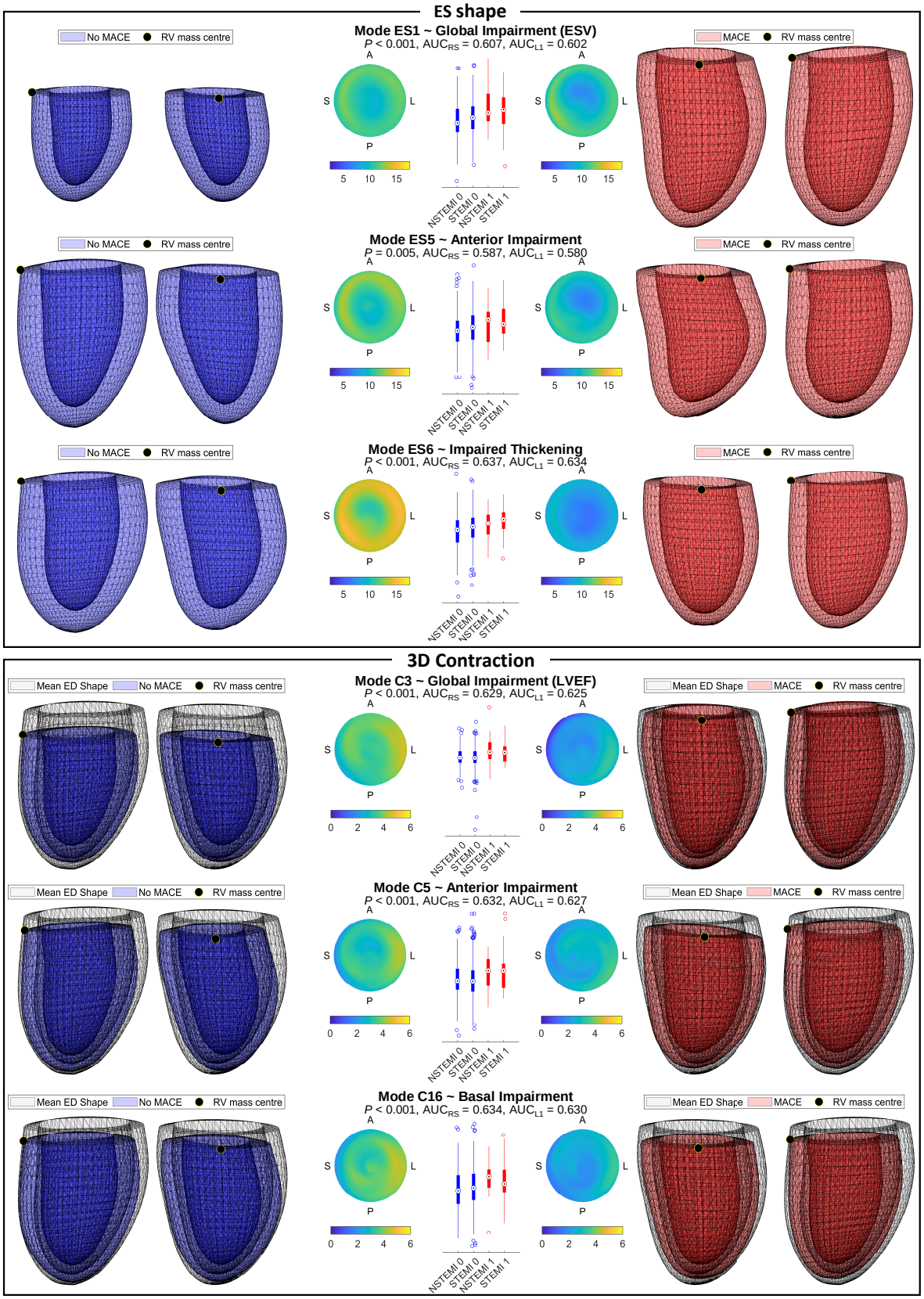


Fig. 2 - **3D LV shape and contraction risk-related patterns.** The 3 main ES shape features (ES1, ES5, ES6) and contraction patterns (C3, C5, C16) related to MACE occurrence are illustrated. Meshes shown in anterior and septal views, and as differential thickness maps (ED-ES thickness) on polar plots of the AHA model. To facilitate comparisons, the contractions are applied on the mean ED shape (reference transparent surface) and visualized as resulting ES shapes. No MACE (blue, class 0) and MACE (red, class 1) representations correspond to the 5th and 95th percentiles in the LDA direction. *P* values and AUCs shown along MACE and no MACE distributions, further stratified into infarct aetiology (STEMI and NSTEMI).

Table 1 - Baseline characteristics, cardiovascular risk factors and CMR biomarkers.

Variable	ALL Patients	MACE (n = 73)	No MACE (n = 948)	AUC_k	P value	HR	HR P value
Age	63 (52 - 72)	72 (61 - 77)	63 (52 - 72)	0.659	<0.001	1.80 (1.39 - 2.32)	<0.001
Sex	753/1011 (74.5)	46/71 (64.8)	707/940 (75.2)	0.505	0.052	0.81 (0.65 - 1.00)	0.050
Height, cm	2 (1 - 2)	2 (1 - 3)	1 (1 - 2)	0.597	0.003	0.69 (0.55 - 0.86)	0.001
Weight, kg	81 (72 - 90)	76 (70 - 86)	82 (73 - 90)	0.568	0.035	0.82 (0.64 - 1.05)	0.110
Cardiovascular risk factors							
Current smoking	405/935 (43.3)	19/63 (30.2)	386/872 (44.3)	0.522	0.029	0.75 (0.57 - 0.97)	0.032
Hypertension	716/1010 (70.9)	61/71 (85.9)	655/939 (69.8)	0.545	0.004	1.53 (1.13 - 2.08)	0.006
Hyperlipoproteinemia	624/1005 (62.1)	45/71 (63.4)	579/934 (62.0)	<0.5	0.816	1.03 (0.81 - 1.30)	0.824
Diabetes mellitus	231/1010 (22.9)	26/71 (36.6)	205/939 (21.8)	0.526	0.004	1.34 (1.09 - 1.64)	0.005
Body mass index, kg/m^2	27.4 (25.0 - 30.4)	27.0 (25.2 - 30.8)	27.4 (25.0 - 30.3)	<0.5	0.959	1.03 (0.82 - 1.30)	0.773
Body surface area, m2	1.95 (1.83 - 2.08)	1.88 (1.76 - 2.00)	1.96 (1.83 - 2.08)	0.593	0.006	0.74 (0.59 - 0.94)	0.014
Killip class on admission				0.573	<0.001	0.52 (0.41-0.65)	<0.001
1	899/1011 (88.9)	49/71 (69.0)	850/940 (90.4)				
2	76/1011 (7.5)	13/71 (18.3)	63/940 (6.7)				
3	20/1011 (2.0)	4/71 (5.6)	16/940 (1.7)				
4	16/1011 (1.6)	5/71 (7.0)	11/940 (1.2)				
Nr. of diseased vessels				0.567	0.003	1.40 (1.12 - 1.75)	0.003
1	502/1011 (49.7)	25/71 (35.2)	477/940 (50.7)				
2	310/1011 (30.7)	23/71 (32.4)	287/940 (30.5)				
3	199/1011 (19.7)	23/71 (32.4)	176/940 (18.7)				
TIMI flow grade post-PCI				<0.5	0.318	0.95 (0.77 - 1.16)	0.598
0	19/1011 (1.9)	1/71 (1.4)	18/940 (1.9)				
1	21/1011 (2.1)	2/71 (2.8)	19/940 (2.0)				
2	78/1011 (7.7)	8/71 (11.3)	70/940 (7.4)				
3	893/1011 (88.3)	60/71 (84.5)	833/940 (88.6)				
CMR biomarkers							
LV ESV, mL	70 (53 - 91)	86 (60 - 110)	69 (53 - 90)	0.599	0.004	1.43 (1.18 - 1.73)	<0.001
LV EDV, mL	144 (117 - 171)	145 (121 - 170)	144 (117 - 172)	<0.5	0.987	1.05 (0.83 - 1.33)	0.679
LVEF (%)	50.5 (43.3 - 57.3)	40.6 (33.1 - 52.2)	50.8 (44.0 - 57.5)	0.683	<0.001	0.80 (0.74 - 0.87)	<0.001
Infarct size, mL	17.2 (6.4 - 30.2)	24.6 (9.7 - 36.4)	16.7 (6.0 - 29.9)	0.591	0.006	1.29 (1.08 - 1.53)	0.005
Infarct size (% LV mass)	13.4 (5.4 - 21.8)	20.3 (9.6 - 28.9)	13.1 (5.3 - 21.4)	0.609	0.001	1.44 (1.18 - 1.76)	<0.001
MVO, mL	0.00 (0.00 - 1.90)	0.40 (0.00 - 3.00)	0.00 (0.00 - 1.80)	0.543	0.060	1.27 (1.11 - 1.45)	<0.001
MVO (% LV mass)	0.00 (0.00 - 1.39)	0.32 (0.00 - 2.15)	0.00 (0.00 - 1.27)	0.547	0.044	1.26 (1.09 - 1.46)	0.002
ES Shape				0.680			
Mode 1	-5 (-135 - 126)	62 (-61 - 233)	-10 (-136 - 119)		0.002	1.49 (1.18 - 1.87)	<0.001
Mode 5	0 (-34 - 36)	-11 (-51 - 22)	1 (-33 - 37)		0.014	0.73 (0.58 - 0.91)	0.006
Mode 6	-3 (-30 - 28)	-21 (-44 - 3)	-2 (-29 - 30)		<0.001	0.64 (0.50 - 0.81)	<0.001
Contraction Displacement				0.716			
Mode 3	-4 (-54 - 62)	49 (-29 - 106)	-6 (-55 - 59)		<0.001	1.70 (1.36 - 2.12)	<0.001
Mode 5	3 (-37 - 39)	-30 (-57 - 23)	4 (-33 - 40)		<0.001	0.63 (0.51 - 0.78)	<0.001
Mode 16	1 (-15 - 14)	-9 (-19 - 5)	1 (-14 - 15)		<0.001	0.65 (0.52 - 0.82)	<0.001

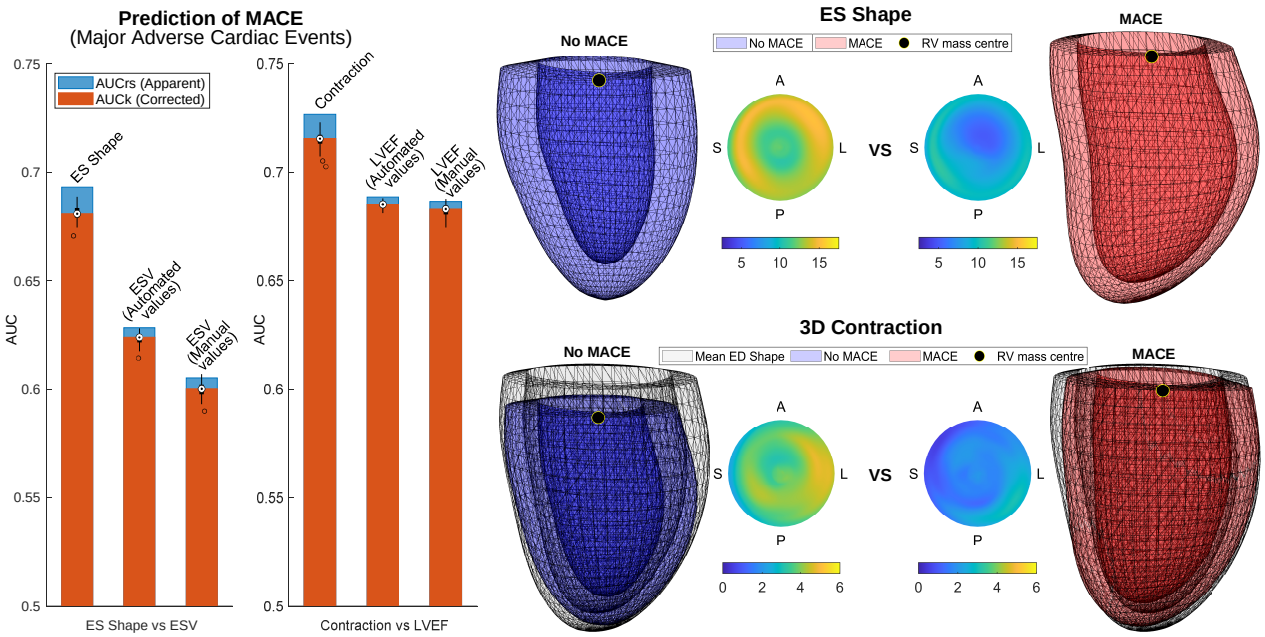
Data presented as n/N (%) or median (interquartile range). *P* values calculated between MACE/No-MACE groups. Hazard ratios (HR) presented with 95% confidence intervals and predictor significance. AUC_k provides the predictive power of each biomarker, assessed via LDA (median AUC, 10-cross-fold validated, 100 random data splits). MACE: major adverse cardiac events; PCI: percutaneous coronary intervention; TIMI: Thrombolysis in Myocardial Infarction.

Endpoints prediction

The automated volumes calculated from the mesh led to a moderate but significant improvement in performance ($P<0.001$) when compared to manual contours (See Central Illustration).

The most predictive variables of the study are ES shape and Contraction, followed by LVEF, age, IS and ESV (Table 1). The LDA on ES shape, described as the combination of ES modes 1, 5 and 6, improves the cross-validated AUC (AUC_k) of endpoints prediction from 0.60 to 0.68 ($P<0.001$) with respect to ESV; the LDA on 3D contraction, described by the contraction modes 3,5,16, increases from 0.68 to 0.72 ($P<0.001$) compared to LVEF (Central Illustration).

The combination of the 3D descriptors with only CMR biomarkers (AUC_k = 0.738), and also including cardiovascular factors and basic patient characteristics (AUC_k = 0.747), led to a further improvement in comparison to the clinical baselines (P<0.001). Similar results were obtained for C-indices (see Table 2). The LDA and Cox analyses converged to the same selection of variables except for a deviation in one variable. The differences in performance hold in the independent testing experiment (Supplemental Material IV). The resulting Receiver-Operator Characteristic curves, Kaplan–Meier estimates and additional multivariable model experiments are included in Supplemental Material III.



Central Illustration - **Prognostic contribution and representative risk-related shape and contraction modes**. **LEFT**: MACE prediction comparison between ESV and LVEF and the proposed 3D disentanglement into ES shape and contraction. Assessment based on AUC resubstitution (AUC_{rs}, blue) and 10-cross-fold validation (AUC_k, orange), repeated for 100 random data splits (black distributions). All prediction differences were significant (P<0.001). Y-axis origin set to 0.5 (random classifier). **RIGHT**: Anterior views of representative No MACE (blue) vs MACE (red) ES shape and 3D contraction, and differential thickness maps (ED-ES thickness) in AHA model representation. Contractions are applied on the mean ED shape (transparent surface) and visualized as resulting ES shapes to ease comparisons.

Table 2 - Risk prediction results.

TYPE	MODEL	LINEAR SELECTION	AUC_K	AUC_{RS}	COX SELECTION	$C\text{-index}_K$	$C\text{-index}_{RS}$
Clinical Baseline	ESV	ESV	0.600 (0.598 - 0.602)	0.605	ESV	0.611 (0.610 - 0.612)	0.614
	LVEF	LVEF	0.682 (0.681 - 0.685)	0.687	LVEF	0.669 (0.668 - 0.670)	0.671
3D analysis	ES Shape	ES1, ES5, ES6	0.681 (0.679 - 0.684)	0.693	ES1, ES5, ES6	0.667 (0.665 - 0.669)	0.677
	3D Contraction	C3, C5, C16	0.716 (0.714 - 0.718)	0.727	C3, C5, C16	0.700 (0.698 - 0.702)	0.709
Multivariable	CMR	ESV, EDV, C5, C16	0.738 (0.736 - 0.740)	0.750	ESV, EDV, C5, C16	0.728 (0.727 - 0.730)	0.736
	ALL	ESV, EDV, Age, Killip, C5, C16	0.747 (0.745 - 0.749)	0.766	ESV, EDV, Age, C5, C16	0.741 (0.739 - 0.744)	0.753

Comparative analysis (LDA and Cox) of clinical baseline, 3D descriptors and multivariable models, with only CMR vs all clinical variables (ALL). ESV, EDV and LVEF calculated from automated volumes (results with manual volumes shown in Table SIII.4). Final selections of variables are reported along their predictive performances, expressed as AUC and C-index, under cross-validation (k=10) and re-substitution (RS). AUC_k and $C\text{-index}_k$ are presented as median (interquartile range). All differences were statistically significant (P<0.001). Killip: Killip class on admission.

DISCUSSION

In this study, we present the first large-sized multicenter CMR study to date that comprehensively analyses the LV 3D shape and contraction after AMI for risk assessment. We have successfully (a) combined AI and computational anatomy technologies to develop a fully automated pipeline that segments the cine CMR SAx stack and identifies the 3D signatures of AMI that predict risks along the first year of follow-up; (b) identified these LV shape and contraction macro-features related to AMI prognosis and built a 3D reference atlas from 1000+ AMI patients; and (c) proven that a 3D enhancement of ESV and LVEF not only outperforms the stand-alone versions, but also contributes to a significant overall risk prediction improvement in a multivariable setting including CMR markers, cardiovascular risk factors and basic patient characteristics.

AMI under the lens of automated shape analysis

Our automated method takes a CMR stack, segments the LV, and summarizes its shape and contraction related to AMI into 6 biomarkers. It reduces the time of the analysis to seconds, removing the burden of manual segmentations. The mesh fitting step (16) ensures 3D-spatial consistency, smoothing the 2D segmentations (i.e. misalignment, segmentation errors) and leading to more accurate volumes. This explains the improvement in MACE prediction driven by the automated volumes in comparison to the manual ones (see Central Illustration).

The method, which runs on a standard laptop, is fully deterministic: given a scan, it always returns the same volumes and scores. This removes intra- and inter-observer variability producing a robust diagnosis. The large size of the cohort, the small gap between re-substitution and cross-validation metrics, and the narrow variance for the 100 random data splits suggest the robustness of the method for MACE prediction. The independent testing experiment confirms the generality of the findings (Supplemental Material IV). This is further supported by the reproducibility analysis studying the effect of deleting the ED and ES frames, which highlights the robustness of the automated volumes (EDV: R2=0.987; ESV: R2=0.990) and confirms a good reproducibility of the method (multivariable ALL: R2=0.901; see Supplemental Material VI), comparable to CMR state-of-art inter-study reproducibility studies (19).

As a result of the automated shape analysis, we have built a 3D LV reference atlas from 1000+ AMI patients that captures the average shape and contraction post-infarct alongside the main variations (PCA modes). These data and the resulting risk models, made publicly available (https://doi.org/10.6084/m9.figshare.13665941.v1), will allow further AMI shape studies, computer mechanistic simulations, or synthetic patient data generation for training algorithms or educational use. The segmenter, specifically modified to cope with different protocols and scanners (as described in (20)), and the meshing pipeline will be likewise available, opening the scope to other cardiac diseases and applications.

Impact and interpretability of the novel methodology

LVEF and ESV are the most established markers for post-infarction risk assessment. While there are data suggesting that these parameters can be fully automatically determined using commercially available software (21), our proposed methodology is further capable of disentangling them into their enhanced 3D versions. It is expected that proposed 3D biomarkers capture more information and achieve additional prognostic value in comparison to their unidimensional version, as shown in Central Illustration (the implication of these AUC improvements is illustrated in Fig. SIII.4). Indeed, although the main motivation of the study was not prediction performance but rather understanding how the 3D features modulates risk, the prognostic value of these 3D biomarkers is superior to any of the variables of the study (Table 1), and, even, to the combination of all the tissue CMR biomarkers (Fig. SIII.3).

Besides, these 3D patterns are complementary to the other variables included in the study (Table 1), as it is evidenced by their significant contribution in any of the considered multivariable approaches (see Table 2) - Unlike, for instance, IS, which is predictive according to the univariate analysis but does not significantly add value to the multivariable models (8), and is therefore not included in the final selection of variables. These results demonstrate the prognostic value of the novel 3D metrics in AMI post-infarction risk assessment, significantly contributing to multivariable models that improve MACE prediction from 0.68 to 0.75 (cross-validated AUC), in comparison to LVEF, the clinical baseline (Table 2).

While the severity of the infarct is associated with morphological and functional alterations, and certain acute changes have proven important to AMI prognosis, their interplay in modulating MACE occurrence remains unsolved (4,6,7). In this study, we have identified 3 ES shape variations and 3 contraction patterns relevant to MACE (Fig. 2 and Supplemental Videos I and II), that suggest 3 possible impairments caused by AMI: global (C3, ES1, ES6), anterior (C5, ES5), and basal (C16). Since the ED shape was not predictive, the changes in morphology seen in the ES modes shall be interpreted as functional and not remodeling. These novel biomarkers are consistent with the predominant LV remodeling descriptors according to the literature: ES1 is strongly correlated to ESV (r_s = 0.868); C3 is related to LVEF (r_s = -0.563); and ES6 to ES myocardial thickness (3,6,7). Indeed, any of the prognostic contraction modes (C3, C5, C16) captures some degree of the latter thickening impairment. ES5 and C5 are alternative (i.e. are not correlated) manifestations of anterior impairment, where the posterior wall pulls the weakened anterior wall and causes a shift in the LV vertical axis. This suggests a perfusion impairment in the left anterior descending (LAD) artery, in agreement to Ortiz-Perez et al. (9). Besides, C16 points to basal impairment as an additional risk factor (see Fig. SX.2, where differences are highlighted), which could be caused by twist differences due to circumferential fiber damage (resulting in larger basal diameters), in contrast to the endocardial vertical and fibers injury that leads to decreased LVEF (elongation and wall thickening, respectively). C16, a basal contraction impermeant, is the signature most complementary to already known predictors.

The derived risk score model, however, does not include the contraction patterns C2 and C14, which could be interpreted as related to injury in the left circumflex and right coronary artery (9) (see Supplemental Material IX and X). These patterns were not significant in the stepwise analysis, suggesting that they could be implicitly described by linear combinations of the selected MACE signatures.

Myocardial tissue death is central to pumping ability and prognosis, however quantifying its contribution to acute remodeling and dysfunction is not straightforward because of hemodynamic reflexes (22). There is an interplay between myocardial 3D shape changes and micro-damage (IS, MVO) that we have detected in our data by the correlation of ES1 to IS (r_s = 0.461) and MVO (r_s = 0.345). Risk-related shape and contraction variations (except for C16) are also significant in stratifying IS and MVO into low/high myocardial damage, using the median values of the AMI population as threshold (see Table 1 and Supplemental Material VII). This explains why IS and MVO are not included in the multivariable risk assessment models: the myocardial damage information is implicit to the identified shape and contraction modes.

The analysis of endpoints prediction stratifying by gender, infarct aetiology and LVEF is discussed in Supplemental Material V.

Towards clinical translation

CMR-based risk models have not found yet their role in clinical practice despite multiple trials showing their incremental prognostic information in AMI management (8). This is partly explained by the more common availability of echocardiography, the complexity of CMR multi-parameter models, and the requirement for significant manual interaction (3,8). Nonetheless, CMR availability has significantly

increased in recent years, CMR post-infarction protocols have been shortened, and complex CMR prognostic markers can now be combined into simple risk score models (8), as we have shown. Our work further contributes to facilitating the adoption of CMR post-infarction risk management in clinical routine by removing the burden of manual segmentation, along with the intra- and interobserver variability, and boosting the MACE prediction with enhanced 3D CMR metrics. While further trials shall be considered to validate the findings of this study, the scientific basis to assume prognosis improvement is solid.

Limitations

Patients were imaged within 10 days after infarction in the absence of optimal post-infarction CMR imaging time recommendations (8). The effect of this post-infarction imaging time on the proposed 3D modes of variations has not been assessed. It is however hypothesized that we are capturing acute injury, whose extent and phenotype can predict outcome, and therefore long-term myocardial adaptions. The main objective of this work is evaluating this hypothesis in a clinical cohort, and not optimizing MACE prediction. For the latter, advanced non-linear classifiers could be explored. The additional prognostic information of later chronic remodeling (23), could be straightforwardly incorporated in our pipeline (i.e. scan at 3 months).

The study focuses only on two instances of the cardiac cycle (ED and ES). This enables for fair and direct comparisons between traditional metrics (ESV, LVEF) and their 3D version. Future work will explore the time dimension and benefit from automated segmentations and 3D reconstruction across the cardiac cycle, presumably finding an even superior marker. The temporal CMR strains, reported to be prognostic in post-infarction management (8), shall be then considered.

The PCA geometric features are limited to the Gaussian assumption, yet they cover satisfactorily the observed variability in the cohort. While other approaches (24) could be considered for dimensionality reduction, they usually come at the cost of interpretability. The PCA selection of features is unsupervised, not biased to outcomes prediction but entirely based on variance, which intrinsically avoids noise and eases interpretability.

CMR long-axis views were excluded from the 3D shape regression for simplicity, which although not evidenced in our results may result in shape lengthening noise. Finally, the results are limited to AMI patients that can undergo a CMR scan.

CONCLUSION

This multicenter CMR study evidences the prognostic value of novel LV 3D shape and contraction metrics in AMI risk assessment. Besides, it further proves the feasibility of fully automated CMR analysis and the significance of multivariable CMR models in AMI stratification.

PERSPECTIVES

Competency in Medical Knowledge

This large size multicenter CMR study identifies the 3D LV patterns related to post-infarction risk and proves their superior prognostic information compared to the most-established predictors, ejection fraction and end-systolic volume. Moreover, it highlights the value of multivariable CMR models in AMI stratification and further supports the feasibility of fully automated CMR analysis.

Translational Outlook

The study demonstrates the potential of 3D computational models to improve myocardial infarction management. This is attained by building a refined understanding on the functional and structural interplay that modulates risk, and by bringing the robustness of state-of-art tools, which ease the clinical workflow and pave the road towards multivariable models and CMR imaging use. To further accelerate this vision and facilitate future contributions, the 3D shape and contraction atlases and shape models resulting from the study are made available.

REFERENCES

1. Corral Acero J, Margara F, Marciniak M, et al. The 'Digital Twin' to enable the vision of precision cardiology. Eur Heart J 2020; 41:4556–64.

2. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016; 388:1545–602.

3. Ibanez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). Eur Heart J 2017; 39:119–77

4. Reindl M, Eitel I, Reinstadler SJ. Role of Cardiac Magnetic Resonance to Improve Risk Prediction following Acute ST-elevation Myocardial Infarction. J Clin Med 2020; 9:1041.

5. Plein S, Younger JF, Sparrow P, Ridgway JP, Ball SG, Greenwood JP. Cardiovascular magnetic resonance of scar and ischemia burden early after acute ST elevation and non-ST elevation myocardial infarction. J Cardiovasc Magn Reson 2008; 10:47

6. Cohn JN, Ferrari R, Sharpe N. Cardiac remodeling-concepts and clinical implications: A consensus paper from an International Forum on Cardiac Remodeling. J Am Coll Cardiol 2000; 35:569–82.

7. Hassell ME, Vlastra W, Robbers L, et al. Long-term left ventricular remodelling after revascularisation for ST-segment elevation myocardial infarction as assessed by cardiac magnetic resonance imaging. Open heart 2017; 4:e000569.

8. Stiermaier T, Backhaus SJ, Lange T, et al. Cardiac Magnetic Resonance Left Ventricular Mechanical Uniformity Alterations for Risk Assessment After Acute Myocardial Infarction. J Am Heart Assoc 2019; 8

9. Ortiz-Pérez JT, Rodríguez J, Meyers SN, Lee DC, Davidson C, Wu E. Correspondence Between the 17-Segment Model and Coronary Arterial Anatomy Using Contrast-Enhanced Cardiac Magnetic Resonance Imaging. JACC Cardiovasc Imaging 2008; 1:282–93.

10. Varela M, Bisbal F, Zacur E, et al. Novel computational analysis of left atrial anatomy improves prediction of atrial fibrillation recurrence after ablation. Front Physiol 2017; 8:68.

11. Thiele H, de Waha S, Zeymer U, et al. Effect of aspiration thrombectomy on microvascular obstruction in NSTEMI patients: The TATORT-NSTEMI trial. J Am Coll Cardiol 2014; 64:1117–24.

12. Eitel I, Wöhrle J, Suenkel H, et al. Intracoronary compared with intravenous bolus abciximab application during primary percutaneous coronary intervention in ST-segment elevation myocardial infarction: Cardiac magnetic resonance substudy of the AIDA STEMI trial. J Am Coll Cardiol 2013; 61:1447–54.

13. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020;368.

14. Corral Acero J, Zacur E, Xu H, et al. SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI. In: Coudière Y. et al. (eds). FIMH 2019. LNCS. Springer, Cham; 2019. p361-69

15. Corral Acero J, Xu H, Zacur E, et al. Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation. In: Pop M. et al. (eds). STACOM 2019. LNCS. Springer, Cham. 2019. p384–94.

16. Lamata P, Sinclair M, Kerfoot E, et al. An automatic service for the personalization of ventricular cardiac meshes. J R Soc Interface 2014; 11:20131023.

17. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 2015; 162:W1

18. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. Ann Intern Med 2019; 170:W1-W33.

19. Grothues F, Smith GC, Moon JCC, et al. Comparison of interstudy reproducibility of cardiovascular magnetic resonance with two-dimensional echocardiography in normal subjects and in patients with heart failure or left ventricular hypertrophy. Am J Cardiol 2002; 90:29–34.

20. Corral Acero J, Sundaresan V, Dinsdale N, Grau V, Jenkinson M. A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation. In: Puyol E. et al. (eds) STACOM 2020. LNCS. Springer, Cham; 2021. p196–207.

21. Schuster A, Lange T, Backhaus SJ, et al. Fully Automated Cardiac Assessment for Diagnostic and Prognostic Stratification Following Myocardial Infarction. J Am Heart Assoc 2020; 9:e016612.

22. Richardson WJ, Clarke SA, Alexander Quinn T, Holmes JW. Physiological implications of myocardial scar structure. Compr Physiol 2015; 5:1877–909

23. Suinesiaputra A, Ablin P, Alba X, et al. Statistical Shape Modeling of the Left Ventricle: Myocardial Infarct Classification Challenge. IEEE J Biomed Health Inform 2018; 22:503–15.

24. Bello GA, Dawes TJW, Duan J, et al. Deep-learning cardiac motion analysis for human survival prediction. Nat. Mach. Intell. 2019; 1(2):95–104.

VIDEO LEGENDS

Supplemental Video I

Interpretability of the ES shape patterns found to be related to AMI prognosis.

Supplemental Video II

Interpretability of the 3D contraction patterns found to be related to AMI prognosis.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Understanding and Improving Risk Assessment after Myocardial Infarction using Automated Left Ventricular Shape Analysis.
Publication Status	□ Published □ Accepted for Publication ☑ Submitted for Publication □ Unpublished and unsubmitted work written in a manuscript style
Publication Details	Jorge Corral Acero, Andreas Schuster, Ernesto Zacur, Torben Lange, Thomas Stiermaier, Sören J. Backhaus, Holger Thiele, Alfonso Bueno-Orovio, Pablo Lamata, Ingo Eitel, Vicente Grau, "Understanding and Improving Risk Assessment after Myocardial Infarction using Automated Left Ventricular Shape Analysis". Submitted to *JACC Cardiovasc Imaging* on 26th May 2021.

Student Confirmation

Student Name:	Jorge Corral Acero		
Contribution to the Paper	(1) Pre-processed the data; (2) developed and implemented methods; (3) designed and ran the experiments; (4) analysed the results; (5) wrote the manuscript; and (6) coordinated the collaboration.		
Signature		Date	12/07/2021

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. Vicente Grau		
Supervisor comments: I confirm that the above statement is accurate.		
Signature 	Date	12/07/2021

This completed form should be included in the thesis, at the end of the relevant chapter.

CHAPTER 6

Enhanced Cardiac Function Characterization Using CMR: Impact on the AMI Case Study

Research Objective 3

6. Enhanced Cardiac Function Characterization Using CMR: Impact on the AMI Case Study

Research Objective 3

6.1. Chapter publications

- **Paper VI**: J. Corral Acero et al. "Comprehensive Characterization of Cardiac Function for Improved Post-Infarction Risk Assessment." Under 1st revision in *Circulation*.

6.2. Chapter Outline

This chapter builds on the two previous research objectives, in order to enhance the characterisation of cardiac function, and complement the structural and functional analysis in CVD using CMR. The methodology herein developed is further showcased on the AMI cohort presented in the previous chapter. As already mentioned, LVEF and ESV are the established biomarkers for AMI assessment in clinical practice [11]. However, they are limited to only two instances of the cycle, ES and ED. In this chapter, three directions of enhancement are explored to further characterise cardiac function in AMI patients. The first direction towards a better description of cardiac contraction aims at characterising the tissue at a bulk level, via CMR-FT strain analysis. This provides granularity in the longitudinal, radial, and circumferential directions, together with information about the uniformity (or lack of it) of myocardial contraction. The second direction of enhancement targets the temporal dynamics of contraction, which is characterised using LV volume transients, proven to be a robust and useful biomarker for this purpose [369]. The third direction is based on the 3D LV patterns already discussed in Chapter 5, and accounts for the 3D detail of cardiac contraction.

The automated pipeline presented in Chapter 4 is extended towards these aims to automatically obtain the LV volume transients from the 2D CMR stacks. The duration, velocities, and contribution to SV of the different phases of the cardiac cycle are likewise regressed automatically from the LV

transients. Theses indexes are referred to as the conventional LV transient markers in Paper VI. Moreover, a PCA approach, similar to the one presented in the previous chapter, is applied to the LV volume transients. This allows to identify interpretable patterns of contraction related to AMI adverse events. These patterns are synthetised and described with just a reduced number of PCA modes (named as AI-derived LV transient metrics in the manuscript). Four contraction patterns are identified in this way to provide additional prognostic value. The CMR-FT analysis focuses on calculating the global strains and quantifying the uniformity of the contraction, proven as prognostic in previous studies (see Section 2.3.4).

The prognostic value of the proposed characterisation of cardiac function is assessed on top of other CMR markers, cardiovascular risk factors, and basic patient characteristics. Results show that any of the three proposed directions of enhancement significantly contributes to an improvement in performance in multivariate risk models. Therefore, these three approaches to characterise the pumping function of the heart provide additional prognostic information for post-infarction risk assessment. A final risk model explores their synergistic combination on top of the other variables of the study. This model reported an outstanding 13% AUC improvement in MACE prediction compared to the risk model recommended in AMI clinical guidelines, based on LVEF.

The four AI-derived LV transient patterns that were identified in the AMI cohort proved significantly superior to the conventional LV transient metrics. As in the previous chapter, the study concludes with the analysis of these patterns to derive an intuitive explanation of their main characteristics and thus advance our understanding of this deadly disease. The first two patterns are interpreted as R-R interval length and average systolic velocity, respectively. Since patients at risk exhibit a decreased LVEF, their passive and active contributions are accordingly reduced and, in consequence, their R-R interval shortened to increase cardiac output. The other two patterns of contraction showing prognostic value are related to diastolic dysfunction. They can be interpreted as the balance between the active and passive contributions, and as diastasis once corrected for R-R

interval length, respectively. It is hypothesised that, in patients at risk, R-R interval shortening compromises ventricular relaxation. This may come at the expense of diastasis and/or faster dynamics, with a swift in the relative filling contributions towards the passive phase, which explains these two latter patterns.

The supplemental material of the manuscript that forms this chapter can be found in Appendix B.

Paper VI

Comprehensive Characterization of Cardiac Function for Improved Post-Infarction Risk Assessment

Jorge Corral Acero, Pablo Lamata, Ingo Eitel, Ernesto Zacur, Ruben Evertz, Torben Lange, Sören J. Backhaus, Thomas Stiermaier, Holger Thiele, Alfonso Bueno-Orovio, Andreas Schuster, Vicente Grau.

Submitted to *Circulation* on 5th October 2021.

TITLE PAGE

Title

Comprehensive Characterization of Cardiac Function for Improved Post-Infarction Risk Assessment.

Authors

Jorge Corral Acero[1,✉], Pablo Lamata[2], Ingo Eitel[3], Ernesto Zacur[1], Ruben Evertz[4], Torben Lange[4], Sören J. Backhaus[4], Thomas Stiermaier[3], Holger Thiele[5], Alfonso Bueno-Orovio[6], Andreas Schuster[4,*], Vicente Grau[1,*].

(*) Equally co-supervised the work.

(✉) Corresponding author: jor.corral@eng.ox.ac.uk / +34646346222 / University of Oxford, Old Road Campus Research Building, Headington, Oxford, OX3 7DQ, UK

Affiliations

1. Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK.

2. Dept. of Biomedical Engineering, King's College London, London, UK.

3. University Heart Centre Lübeck, Medical Clinic II, Cardiology, Angiology and Intensive Care Medicine, Lübeck, Germany; University Hospital Schleswig-Holstein, Lübeck, Germany; and German Centre for Cardiovascular Research (DZHK), partner site Hamburg/Kiel/Lübeck, Lübeck, Germany.

4. University Medical Centre Göttingen, Department of Cardiology and Pneumology, Georg-August University, Göttingen, Germany; and German Centre for Cardiovascular Research (DZHK), partner site Göttingen, Göttingen, Germany.

5. Heart Centre Leipzig at University of Leipzig, Department of Internal Medicine and Cardiology, Leipzig, Germany; and Leipzig Heart Institute, Leipzig, Germany.

6. Department of Computer Science, University of Oxford, Oxford, UK.

Article specifications

- ✓ Article length: 4,510 words
- ✓ Article length including title page, abstract, body text, references and figure legends: 7,133 words
- ✓ Number of references: 43
- ✓ Number of tables: 3
- ✓ Number of figures: 4

ABSTRACT AND KEYWORDS

Background

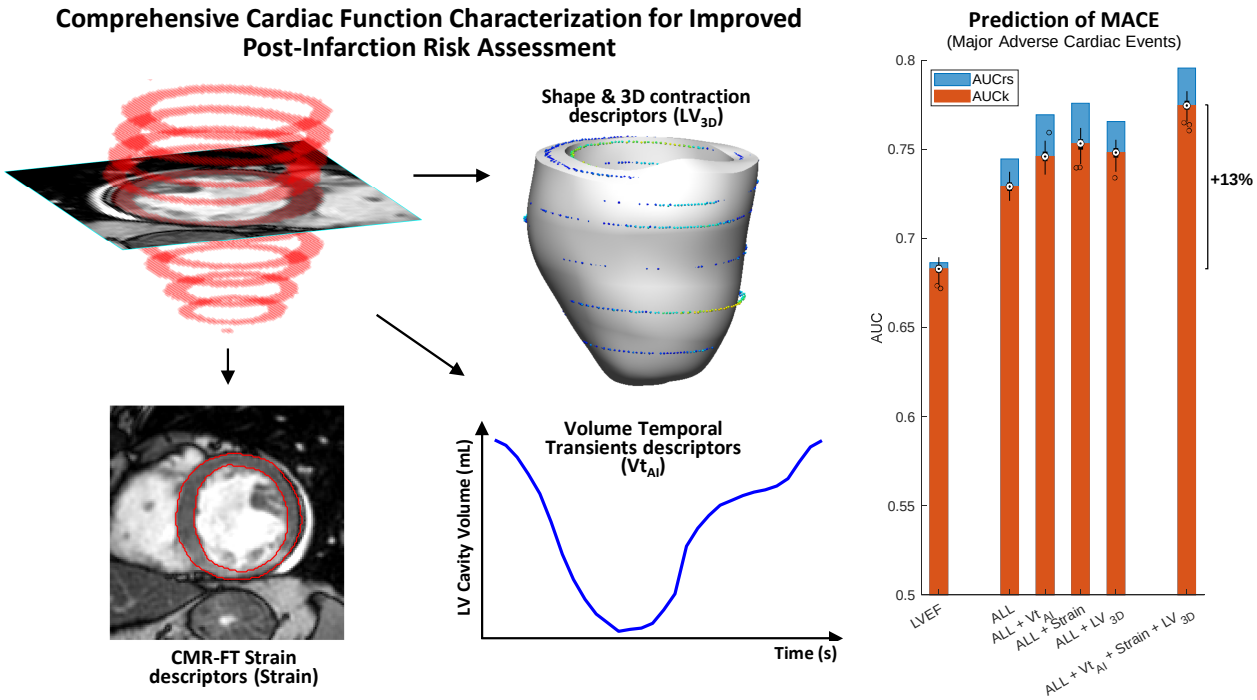
Left ventricular ejection fraction (LVEF), a stand-alone global cardiac function index, is the established image marker for postinfarct risk stratification in acute myocardial infarction (AMI) patients. This study aims at improving risk prediction and identifying risk-related patterns by enhancing the characterization of cardiac function.

Methods

A multicenter cohort of AMI survivors (n=1021; median age: 63 years; 74.5% male) who underwent cardiac magnetic resonance (CMR) at a median of 3 days after AMI was considered for this study. The clinical endpoint was the 12-month rate of major adverse cardiac events (MACE, n=73), consisting of all-cause death, reinfarction, and new congestive heart failure. Cardiac function was characterized from CMR by (1) feature tracking (FT) strain analysis (i.e. bulk tissue contraction); (2) volume temporal transients (i.e. contraction dynamics); and (3) left ventricular (LV) shape analysis (i.e. 3D contraction detail). A fully automated pipeline was developed to extract conventional and novel artificial-intelligence-derived metrics of cardiac contraction, and their relationship with MACE was investigated. Comparisons were based on the prediction performance resulting from both multivariable linear discriminant (LDA – MACE occurrence) and Cox (time to MACE) analyses.

Results

Any of the 3 proposed directions to enhance the description of contraction demonstrated its additional prognostic value on top of established CMR global indexes, myocardial injury markers, basic characteristics, and cardiovascular risk factors (P<0.001). The combination of these 3 directions of enhancement towards a final CMR risk model improved MACE prediction by 13% compared to LVEF (0.774 (0.771 - 0.777) vs 0.683 (0.681 - 0.685) cross-validated AUC, P<0.001), significantly outperforming advanced baseline multivariable models as well (0.729 (0.727 - 0.733) , P<0.001). The automated pipeline reduced the analysis time to seconds, removing the burden of manual annotations and intra- and interobserver variability.



Graphical Abstract - **LEFT**: The 3 proposed directions of enhancement of cardiac function characterization, enabled by a fully automated pipeline, for improved post-infarction management beyond left ventricular ejection fraction (LVEF), the established marker. **RIGHT**: Additional prognostic value of each of the proposed directions of enhancement and the combination of them on top of established CMR global indexes, myocardial injury markers, basic characteristics, and cardiovascular risk factors (ALL). Assessment is based on multivariate backward stepwise linear discriminant analysis, and results are expressed in AUC resubstitution (RS, blue) and 10-fold cross validation (K10, orange), repeated for a hundred random data splits (black distributions). All differences in performance were significant (P<0.001). Note the remarkable improvement of the combination of the proposed approaches with respect to LVEF and the advance multivariate baseline (ALL).

Conclusion

Characterization of cardiac contraction patterns via CMR-FT strains, LV volume transient and shape analyses, enabled by a fully automated pipeline, significantly improves post-AMI risk prediction.

Keywords

Cardiovascular magnetic resonance, acute myocardial infarction, risk assessment, cardiac function characterization, fully automated analysis.

INTRODUCTION

Cardiovascular diseases are the worldwide leading cause of death, with a toll further climbing each decade (1). Acute myocardial infarction (AMI) plays a major role among them, owing to its prevalence, counted in millions, yearly, and its high case-fatality (2). Despite the significant advances in medical treatment towards personalized and preventive medicine (2,3), the mortality of AMI survivors within 6 months after infarct sits at 12% (4). Early and improved risk assessment is crucial to reduce this burden.

Left ventricular (LV) remodeling is central to AMI early prognosis prediction (5). Consequently, LV macro-function, quantified as LV ejection fraction (LVEF), is the preferred image biomarker for therapeutic decision-making and clinical risk stratification, according to AMI guidelines (4,6). Nonetheless, LVFE fails to capture contraction dynamics as well as functional and anatomical regional abnormalities (7,8). Moreover, LVEF is mainly preserved or only moderately reduced in most AMI survivors and therefore recurrent adverse events often occur in patients at a theoretical low risk, which further evidences the need of stratification improvements (9,10).

Cardiac magnetic resonance (CMR) imaging has proven uniquely suitable to assess morphological and functional myocardial alterations and provides incremental prognostic information over and above LVEF (11–13). LV micro-myocardial injury, assessed by late gadolinium enhancement (LGE) CMR and usually quantified as infarct size (IS) and microvascular obstruction (MVO), has emerged as a robust outcome measure in AMI risk assessment (6,11,13,14). Estimation of local LV remodeling using CMR has been proposed to improve prognostic information (15,16). Likewise, shape analysis of the LV to identify 3D patterns related to risk has demonstrated its additional contribution to risk assessment (17). CMR myocardial feature tracking (CMR-FT) has been successfully applied for quantification of LV deformation indexes and mechanical uniformity alterations, associated with hemodynamic alterations, adverse LV remodeling, and clinical outcomes (7,9,13,18,19). Recent advances in deep learning enable the extraction of CMR-derived volume temporal transients, which have proven to be a robust and useful biomarker to characterize cardiac function (20).

Building on these advances, and further automating and enhancing the descriptors of contraction, this work explores the ability to detect novel signatures that predict post-infarction risks at bulk tissue level, temporal dynamics, and 3D detail. The prognostic contribution of this enhanced cardiac function characterization using CMR-FT strains, LV volume transients and 3D LV patterns is assessed on a large multicenter study, including ST-segment–elevation myocardial infarction (STEMI) and non-STEMI (NSTEMI) patients. The study demonstrates the potential of the herein proposed fully automated analysis for AMI risk management and identifies risk-related contraction patterns, expanding the understanding of post-infarction remodeling.

METHODS

Study Population

A total of 1235 AMI survivors from the AIDA-STEMI (NCT00712101) and TATORT-NSTEMI (NCT01612312) multicenter randomized trials, registered with ClinicalTrials.gov and approved by the ethics committees, were considered for the study (21–23). This size of the study is justified in Supplemental Material VIII, based on (24). Reperfusion therapy with primary percutaneous coronary intervention and postinfarction medical treatment were supplied according to state-of-the-art guideline recommendations (4). The study was conducted according to the Declaration of Helsinki and written informed consent was requested for study participation. The data that supports the findings is available upon reasonable request.

CMR Imaging Protocol

The patients underwent CMR imaging on 1.5T or 3T clinical scanners within 10 days after infarct following a standardized protocol (6,21,22), that includes ECG-gated balanced steady-state free precession sequences and T1-weighted LGE images. All sequences were acquired in horizontal and vertical long-axis views as well as continuous short-axis (SAx) stacks capturing the whole LV. Ventricular volumes and infarct characteristics were determined in sequential SAx by blinded clinicians via dedicated software (cmr42; Circle Cardiovascular Imaging) (6,22). Standard thresholding techniques were applied to assess IS and MVO, as explained in (6,22).

Study Endpoints

The predefined clinical endpoint of the study was the 12-month rate of major adverse cardiac events (MACE), consisting of new congestive heart failure, reinfarction and all-cause death, as detailed in (21–23). In case of multiple MACE only one contribution per patient to the composite of endpoints was considered (death>reinfarction>congestive heart failure). The events were adjudicated by a blinded clinical committee based on the data collected in the study sites.

CMR-FT Strain Analysis (Strain)

CMR-FT was conducted at the University Medical Centre Göttingen core laboratory using certified software (2D CPA MR, Cardiac Performance Analysis v1.1.2; TomTec Imaging Systems (25)). Final values were based on the average of 3 independent analyses (26). Circumferential and radial strains were determined at basal, midventricular, and apical locations, as described in (9,19). Longitudinal strains were derived at both 2- and 4-chamber long-axis views. Global longitudinal, circumferential, and radial strains (GLS, GCR and GRS, respectively) as well as circumferential and radial uniformity ratio estimates (CURE and RURE), were reported (27). CURE and RURE range between 0, complete nonuniform contraction, and 1, perfect uniformity, that is, equal strain across the myocardium at any given time point. Calculation details can be found in (9).

LV Fully Automated Volume Transients Analysis

The process, illustrated in Fig. 1, is fully automated and consist of the following steps: (1) LV myocardium segmentation from SAx images and LV volume temporal transient reconstruction from the cavity volume integration along the cardiac cycle; (2) Computation of conventional metrics from LV volume transient; (3) Definition of Artificial Intelligence (AI) based metrics of LV volume transients, by unsupervised construction of statistical models and the subsequent selection of the features with best prognostic value.

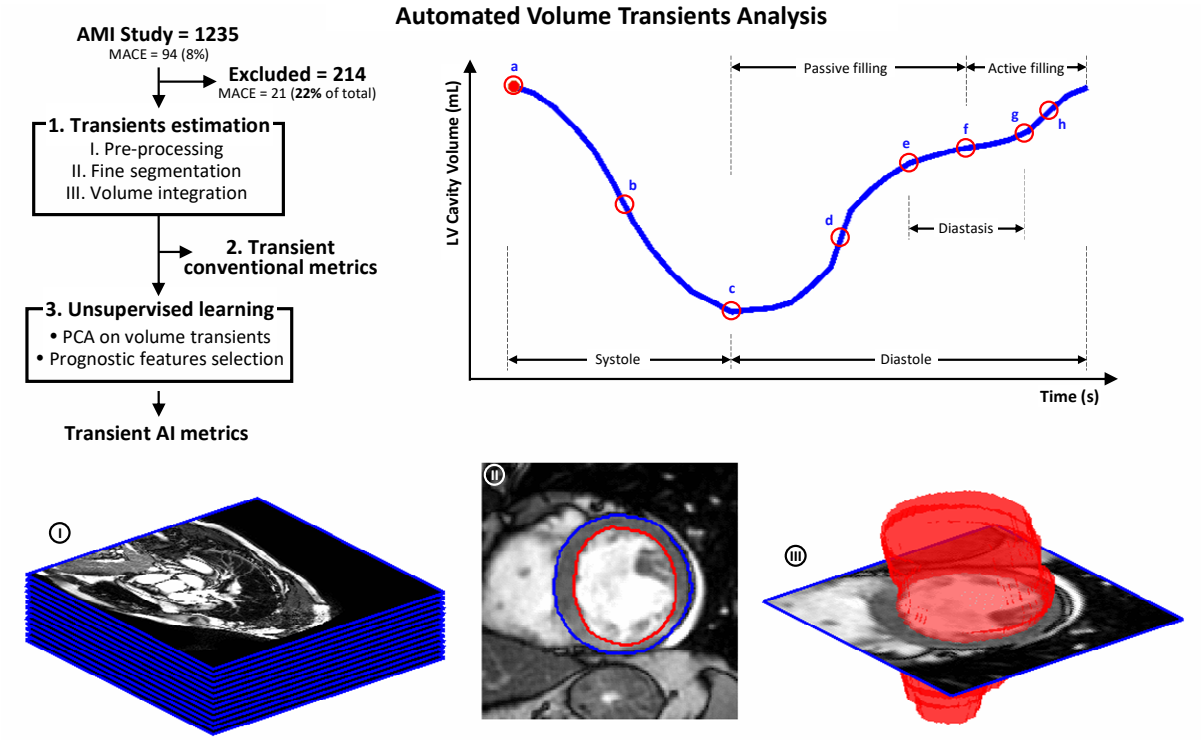


Fig. 1 - LV fully automated volume transient analysis pipeline (top-left), along with patient S692 sample case, consisting of the following steps: (1.I.) the SAx stack (image I) is pre-processed normalizing for intensity, resolution and orientation and cropped to the region of interest, following the prior detection of the heart; (1.II.) the LV myocardium is fine segmented in the SAx stack (image II); (1.III) a volumetric segmentation is synthesized (image III) and the LV volume transient derived; (2) the conventional metrics of the volume transient are estimated (top-right); and (3) PCA analysis is applied on the transients to facilitate the supervise learning of MACE occurrence related features (AI-derived transient metrics). The red circle landmarks on the representative volume transient (top-right) correspond to: a, end-diastole; b, systolic peak velocity; c, end-systole; d, passive filling peak velocity; e, beginning of diastasis; f, mid-diastasis; g, end-diastasis; and g, active filling peak velocity. This figure is available in video format (see Supplemental Videos).

LV Segmentation and Volume Transients Estimation

A 2-step deep learning approach was applied to segment the LV cavity, based on a UNet architecture with enhanced pre-processing (28,29), that reached the best performance in the 2019 LV Full Quantification Challenge (29,30). Implementation details are available in Supplemental Material I. Resulting segmentations were arranged in binary volumetric images, and the LV cavity volume was derived for each frame via trapezoidal integration.

Volume Transient Conventional Metrics (Vt)

The duration, average and peak velocity of the cardiac phases shown in Fig. 1 were derived (20). The transients were resampled using splines to increase resolution and better capture the slope-sensitive metrics. Diastasis was estimated as the mid-diastolic region with a velocity inferior to 80% stroke volume (SV) per second (empirical threshold). The time to diastasis, measured from end-systole and mid-diastasis, was set as the boundary between the passive and active filling contributions. Passive and active filling metrics were not reported in those cases where diastasis was not reliably found (diastasis = 0s, n = 212). The average RR-interval and its variability, measured as highest minus lowest values, as well as the frame trigger times, were directly retrieved from the patient scans metadata.

Volume Transient AI Metrics (Vt_{AI})

Principal Component Analysis (PCA) is an unsupervised machine learning technique that inspects the data to find the directions of change that maximize the variance observed in the population. Independently from the above calculations on raw volume transients, PCA was applied to the volume transients (Formulation details and a more detailed explanation is available in Supplemental Material

III), minmax normalized to standardize in ventricle size and SV (see implications in Supplemental Material II). The resulting directions are the so called 'PCA modes', and represent patterns of change with respect to the mean volume transients (e.g. RR-interval length, passive vs active contribution, etc. - See Fig. 2, panel headings representing mode interpretation). Thus, the volume transient of each subject can be represented by the average transient of the population plus a weighted sum of these modes of variation. The weights used are, therefore, continuous variables which have a unique value for each patient and become potential biomarkers to predict MACE. Among them, those with highest prognostic value were identified by a stepwise multivariable Fisher linear discriminant analysis (LDA) (31), as detailed below, on the modes that reconstruct 95% of the total variance of the population.

LV Fully Automated 3D Shape Analysis (LV_{3D})

The analysis is explained in (17). In brief, an automated pipeline was developed to automatically obtain the personalized 3D LV mesh at end-systole (ES) and end-diastole (ED); and a similar AI methodology to the explained above is applied to identify the MACE-related ES shape and 3D contraction patterns. The contraction was approximated as ED to ES displacements. Table 1 summarizes the three ES shape and three contraction descriptors found, along with their interpretation (17). The 3D LV patterns are illustrated in Fig. SIV.1.

Prognostic Value Assessment

The additional prognostic contribution of each of the 3 proposed directions of enhancement, CMR-FT strains (strains), volume temporal transients (conventional and AI-derived - Vt and Vt_{AI}) and LV 3D patterns (LV_{3D}) markers, was assessed considering only CMR biomarkers (CMR) and all the clinical variables (ALL) of the study (see Table 1). An additional analysis evaluated the additional prognostic value of their combination.

Comparisons were based on the prediction performance resulting from both a multivariable LDA (binary classification – MACE occurrence at 12 months) and a multivariable Cox analysis (time to MACE occurrence). A backward stepwise strategy (unbiased to univariable associations (32)) was followed in both analyses to find inter-variable synergies and address collinearity: all variables of interest were initially included, and the less significant in the model iteratively removed, until a 0.05 *p*-value threshold was met by all of them. Significance in the LDA was approximated as significance in the generalized linear binomial sigmoid regression model.

The performance of the resulting configurations was assessed via area under the receiver-operator characteristic curve (AUC), for binary classifications, and concordance index (C-index) (33), for time-dependent curves, to account for both specificity and sensitivity, given the MACE class imbalance. Performance in both resubstitution (apparent performance - learning and testing with the entire cohort) and cross-validation (optimism corrected performance - 10 cross-fold-validation repeated for a hundred random data splits) is reported to ensure generality of findings.

Table 1 - Baseline CMR, Baseline ALL and 3D End-systolic shape and Contraction variables.

CMR - Baseline CMR variables	**No.**
End-systolic volume (mL), End-diastolic volume (mL), Left ventricular ejection fraction (%), Infarct size (mL), Infarct size (%LV mass), Microvascular obstruction (mL)	6
ALL - Baseline ALL variables	**No.**
Age (y), Body mass index (Kg/m2), Body surface area (m2), Current smoking, Diabetes, Height (cm), Hyperlipoproteinemia, Hypertension, Killip class on admission, Number of disease vessels, Sex, TIMI flow grade post-PCI, Weight (kg) and the 6 CMR variables indicated above.	19
LV_{3D} - 3D Left-ventricular End-systolic shape and Contraction variables	**No.**
Mode ES1 (~Global Impairment, ESV), Mode ES5 (~Anterior Impairment), Mode ES6 (~Impaired Thickening), Mode C3 (~Global Impairment, LVEF), Mode C5 (~Anterior Impairment), Mode C16 (~Basal Impairment)	6

List and number of variables included in the baseline CMR (CMR), baseline ALL (ALL) and LV 3D patterns (LV_{3D}) for prognostic assessment experiments. Interpretation of each of the LV_{3D} variables is indicated in parenthesis. See (17) for more details.

Statistical Analysis

The variables of the study are described according to MACE occurrence (Table 2). Continuous variables were not normally distributed in a Shapiro–Wilk test, except for the Vt_{AI} modes of variation (Gaussian distributed by definition), and are therefore presented as median and interquartile range (IQR). Distributions (variable stratifications, cross-validated AUC model results, etc.) were compared by the non-parametric Wilcoxon rank sum test. The resulting *p*-value of the MACE vs No MACE comparisons, as well as the univariate-LDA cross-validated performance, was reported for each variable. Univariate Cox regression analyses were likewise completed. Hazard ratios, 95% confidence intervals and the predictor significance are presented in the results. Correlation between variables was assessed via Pearson (r_p) and Spearman (r_s) coefficients. All analyses were implemented in Matlab R2019b. The study follows the guidelines of transparent reporting of multivariable prediction model for individual prognosis or diagnosis (TRIPOD, see Supplemental Material VIII) and Prediction model Risk Of Bias ASsessment Tool (PROBAST) for model development (32,34).

RESULTS

Patients

In total, 1021 patients (STEMI: n=723; NSTEMI: n=298) of the 1235 cohort were included in the study (No CMR: n=126; incomplete protocol: n=86; no follow-up: n=2). Out of them, 73 patients suffered from MACEs (congestive heart failure: n=20; reinfarction: n=21; death: n=32) (21–23). The population was predominantly male (74.5%) with a median age of 63 (IQR: 52–72) years, LVEF of 50.5% (IQR: 43.3-57.3) and IS of 13.4% of LV mass (IQR: 5.4-21.8). Their CMR and baseline clinical characteristics have been reported and discussed previously (21–23) and are reproduced in Supplemental Material IV.

CMR-FT Strain Analysis (Strain)

The CMR-FT strain results are described in Table 2. The MACE subgroup exhibited significantly lower global longitudinal, circumferential, and radial strains (P<0.001) and a less uniform radial and circumferential contraction (P<0.001).

Automated Volume Transient AI Analysis

LV Segmentation and Volume Calculation

The automated segmentation of the SAx resulted in a median Dice score of 0.971 (IQR: 0.945-0.981) for the LV cavity and 0.975 (IQR: 0.959-0.982) for the LV epicardium. The correlation, r_s, between the volumes calculated from the automated and manual segmentations were 0.916 and 0.919, for end-systolic and end-diastolic volume (ESV and EDV), respectively.

Volume Transient Conventional Metrics (Vt)

Results are detailed in Table 2. The MACE subpopulation was characterized by a significantly shorter and more variable RR-interval (P<0.001). Their systolic (P=0.038), diastolic (P=0.025), diastasis (P=0.037) and time to diastasis (P=0.013) was accordingly reduced. Nevertheless, only diastasis was significantly smaller when normalizing in time (P=0.032). The passive (P=0.002) and active (P<0.001) contributions to LV filling were significantly compromised in MACE patients, as ESV and EF were significant (P<0.001 – See Table SIV.1), but not when normalizing by SV. Regarding velocities, the MACE cohort presented significantly lower systolic peak (P=0.029), and active filling average (P=0.008) and peak (P=0.003) velocities. However, it was the passive filling average velocity the only one significant to MACE (P=0.013) when normalizing by SV. Fig. SIV.2 shows the average MACE and No MACE volume transients normalized by SV.

Volume Transient AI analysis (Vt_{AI})

95% of the variance was explained by the first 6 unsupervised modes of variation. Among them, modes 1, 2, 3 and 5 ($Vt_{AI}1$, $Vt_{AI}2$, $Vt_{AI}3$ and $Vt_{AI}5$) were determined the most relevant to MACE by the LDA stepwise analysis. These modes are interpreted as RR-interval ($Vt_{AI}1$), systolic velocity ($Vt_{AI}2$), passive vs active filling ($Vt_{AI}3$) and diastasis ($Vt_{AI}5$) (see Discussion). Fig. 2 illustrates these modes and summarizes their significance, classification performance and distributions, stratified by MACE occurrence and infarct aetiology. The correlation between these AI metrics and the conventional transient ones is summarized in Fig. 3 (Pearson) and Supplemental Material VI (Spearman and R^2).

Table 2 - CMR-FT strain, conventional and AI-derived automated volume transient metrics.

Variable	ALL Patients	MACE (n = 73)	No MACE (n = 948)	AUC_k	*P*val	HR	HR *Pval*
Strain - Strain markers:							
Global longitudinal strain (%)	-16.5 (-20.2 - -12.5)	-11.5 (-16.8 - -8.4)	-16.7 (-20.4 - -13.0)	0.710	<0.001	2.13 (1.68 - 2.69)	<0.001
Global circumferential strain (%)	-24.2 (-28.9 - -19.2)	-18.6 (-24.2 - -14.6)	-24.4 (-29.1 - -19.7)	0.674	<0.001	1.72 (1.39 - 2.12)	<0.001
Global radial strain (%)	20.4 (15.7 - 25.9)	16.4 (12.6 - 22.6)	20.5 (16.0 - 26.0)	0.626	<0.001	0.64 (0.50 - 0.83)	<0.001
CURE	0.78 (0.71 - 0.84)	0.71 (00.7 - 0.80)	0.79 (0.72 - 0.84)	0.637	<0.001	0.62 (0.50 - 0.76)	<0.001
RURE	0.75 (0.67 - 0.82)	0.69 (00.6 - 0.78)	0.76 (0.67 - 0.83)	0.629	<0.001	0.66 (0.54 - 0.82)	<0.001
Vt - Volume transient markers:							
RR-interval (ms)	845 (741 - 952)	780 (674 - 848)	845 (750 - 952)	0.644	<0.001	0.64 (0.51 - 0.81)	<0.001
RR-interval variability (%RR)	19.2 (13.1 - 26.3)	24.6 (14.2 - 36.7)	18.9 (13.0 - 25.6)	0.606	<0.001	1.28 (1.13 - 1.46)	<0.001
Systolic phase:							
Systolic time (ms)	297 (267 - 328)	284 (265 - 321)	298 (268 - 328)	0.562	0.106	0.70 (0.51 - 0.98)	0.038
Systolic velocity avg (mL/s)	251 (209 - 301)	227 (184 - 271)	252 (212 - 303)	0.592	0.238	0.61 (0.34 - 1.11)	0.105
Systolic velocity avg (%SV/s)	337 (305 - 375)	352 (312 - 377)	336 (305 - 373)	0.554	0.314	1.08 (0.93 - 1.25)	0.315
Systolic velocity peak (mL/s)	570 (451 - 704)	514 (396 - 650)	573 (456 - 711)	0.573	0.030	0.77 (0.60 - 0.97)	0.029
Diastolic phase:							
Diastolic time (ms)	513 (431 - 602)	462 (394 - 523)	518 (435 - 606)	0.611	0.028	0.76 (0.60 - 0.97)	0.025
Diastasis (ms)	38 (0 - 118)	24 (0 - 83)	44 (0 - 119)	0.569	0.034	0.74 (0.56 - 0.98)	0.037
Diastasis (%)	10.8 (0.0 - 21.6)	5.4 (0.0 - 16.2)	10.8 (0.0 - 21.6)	0.558	0.030	0.74 (0.56 - 0.97)	0.032
Time to diastasis (ms)	354 (307 - 402)	327 (267 - 366)	356 (309 - 405)	0.605	0.017	0.70 (0.53 - 0.93)	0.013
Time to diastasis (%)	64.9 (62.2 - 70.3)	64.9 (56.8 - 70.9)	64.9 (62.2 - 70.3)	0.512	0.192	0.82 (0.62 - 1.09)	0.177
Passive filling (mL)	55.6 (43.0 - 67.2)	48.1 (37.6 - 60.1)	56.0 (43.5 - 67.9)	0.630	0.002	0.64 (0.48 - 0.85)	0.002
Passive filling (%SV)	66.5 (58.2 - 73.5)	67.3 (57.8 - 78.9)	66.4 (58.2 - 73.0)	<0.5	0.767	1.05 (0.77 - 1.42)	0.776
Active filling (mL)	26.2 (19.9 - 33.6)	20.6 (12.7 - 28.5)	26.7 (20.2 - 33.9)	0.647	<0.001	0.71 (0.59 - 0.87)	<0.001
Active filling (%SV)	31.5 (23.8 - 39.8)	26.6 (17.8 - 37.7)	31.7 (24.1 - 39.9)	0.559	0.110	0.81 (0.63 - 1.04)	0.094
Diastolic velocity avg (mL/s)	143 (121 - 170)	141 (109 - 168)	144 (122 - 170)	0.535	0.458	0.20 (0.04 - 0.97)	0.046
Diastolic velocity avg (%SV/s)	195 (166 - 232)	216 (191 - 254)	193 (165 - 230)	<0.5	0.982	1.00 (0.79 - 1.26)	0.982
Passive filling velocity avg (mL/s)	155 (122 - 187)	151 (118 - 180)	155 (122 - 188)	0.516	0.291	0.87 (0.69 - 1.11)	0.274
Passive filling velocity avg (%SV/s)	185 (153 - 215)	205 (170 - 239)	184 (153 - 213)	0.592	0.015	1.35 (1.07 - 1.71)	0.013
Passive filling velocity peak (mL/s)	427 (307 - 573)	384 (271 - 502)	429 (310 - 575)	0.553	0.143	0.80 (0.59 - 1.08)	0.143
Active filling velocity avg (mL/s)	128 (92 - 169)	104 (70 - 144)	130 (94 - 169)	0.610	0.011	0.76 (0.62 - 0.93)	0.008
Active filling velocity avg (%SV/s)	154 (106 - 205)	142 (94 - 222)	154 (107 - 204)	0.521	0.277	0.85 (0.65 - 1.12)	0.260
Active filling velocity peak (mL/s)	315 (222 - 456)	246 (176 - 382)	323 (227 - 464)	0.629	0.003	0.64 (0.47 - 0.85)	0.003
Vt_{AI} - Volume Transient modes:							
$Vt_{AI}1$	0.03 (-0.45 - 0.54)	0.39 (-00.2 - 0.66)	0.02 (-0.47 - 0.51)	0.570	0.112	1.22 (0.96 - 1.56)	0.103
$Vt_{AI}2$	-0.08 (-0.17 - 0.04)	0.00 (-00.1 - 0.13)	-0.08 (-0.18 - 0.04)	0.598	<0.001	1.27 (1.12 - 1.44)	<0.001
$Vt_{AI}3$	0.00 (-0.22 - 0.22)	0.11 (-00.1 - 0.33)	-0.01 (-0.22 - 0.21)	0.588	0.025	1.34 (1.05 - 1.71)	0.020
$Vt_{AI}5$	-0.02 (-0.10 - 0.09)	0.06 (-00.1 - 0.17)	-0.02 (-0.10 - 0.08)	0.611	<0.001	1.35 (1.14 - 1.59)	<0.001

Data presented as median (interquartile range). *P*-values were calculated for the comparison between patients with and without MACE. Univariate Cox regression analyses are presented as HR, Hazard ratios (95% confidence interval), and HR *P*-value, predictor significance. The predictive power of each biomarker is assessed via LDA and presented as median AUC 10-cross-fold validated, repeated for a hundred random data splits, AUC_k. MACE indicates major adverse cardiac events; and avg, average. The passive and active filling metrics were not reported in those cases where diastasis was not found (diastasis = 0s, n = 212).

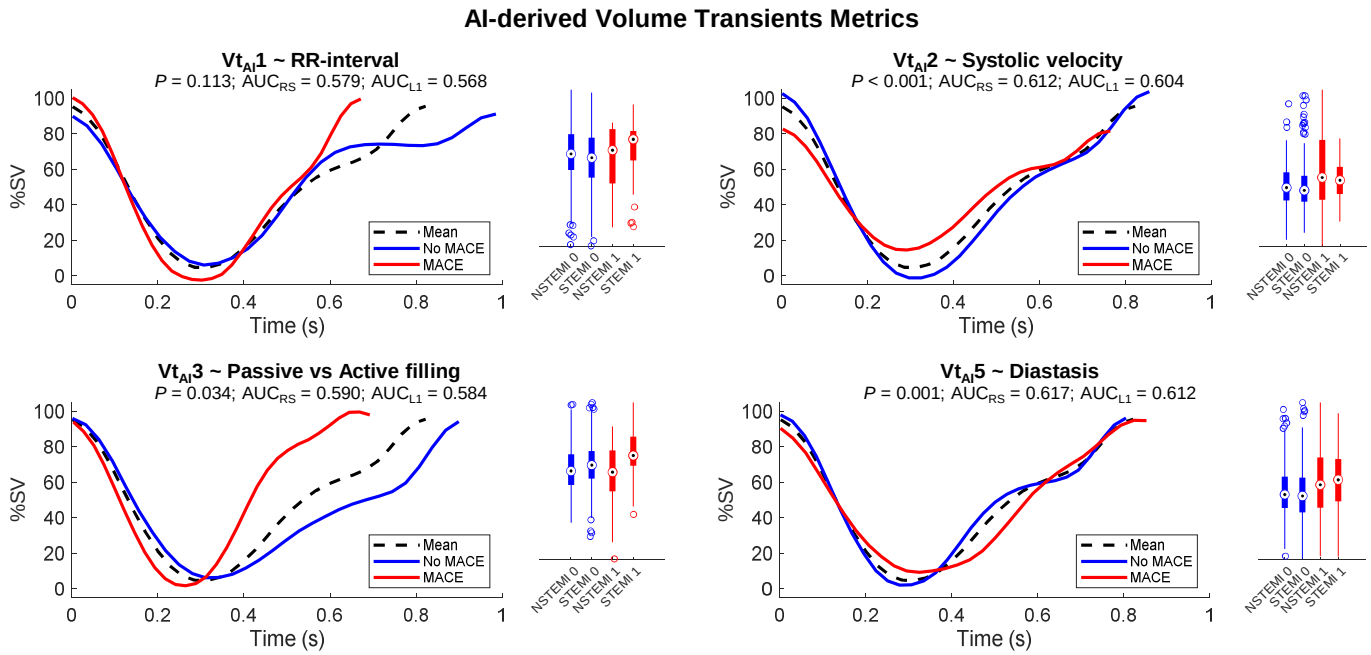


Fig. 2 - AI-derived volume transient features most relevant to MACE occurrence prediction, resulting from unsupervised learning ($Vt_{AI}1$, $Vt_{AI}2$, $Vt_{AI}3$ and $Vt_{AI}5$). The MACE (red, class 1) and No-MACE (blue, class 0) traces shown correspond to the 10th and 90th percentiles. This allows to visualize the pattern of change encoded by each of the unsupervised variables (RR-interval, passive vs active filling, etc.), as well as to describe how a representative MACE and No-MACE normalized volume transient would theoretically look like according to each of these four unsupervised variables. The *P*-value, re-substitution and leave-one-out AUCs are presented along each mode as MACE and No-MACE distributions, further stratified into infarct aetiology (STEMI and NSTEMI). This figure is available in video format, where the evolution from the MACE to the No MACE extremes is shown to better appreciate the pattern of change encoded in each of the VT_{AI} contraction modes (see Supplemental Videos).

Table 3 - Additional prognostic contribution of the proposed enhanced LV function characterization.

MODEL	LINEAR SELECTION	AUC_k	AUC_{RS}	COX SELECTION	C_{L1}	C_{RS}
LVEF	LVEF	0.683 (0.681 - 0.685)	0.687	LVEF	0.668	0.669
CMR	ESV, EDV	0.700 (0.698 - 0.702)	0.708	ESV, EDV	0.689	0.693
CMR + Vt	ESV, EDV, RR, $t_{diastolic}$	0.708 (0.706 - 0.710)	0.721	ESV, EDV	0.689	0.693
CMR + Vt_{AI}	ESV, EDV, $Vt_{AI}2$, $Vt_{AI}3$, $Vt_{AI}5$	0.726 (0.723 - 0.730)	0.742	ESV, EDV, $Vt_{AI}3$, $Vt_{AI}5$	0.705	0.716
CMR + Strain	ESV, EDV, GLS	0.723 (0.721 - 0.725)	0.733	ESV, EDV, GLS	0.710	0.717
CMR + LV_{3D}	ESV, EDV, C5, C16	0.738 (0.736 - 0.740)	0.750	ESV, EDV, C5, C16	0.729	0.736
ALL	ESV, EDV, Age, Killip	0.729 (0.727 - 0.733)	0.745	ESV, EDV, Age, Killip	0.715	0.726
ALL + Vt	ESV, Age, Killip, BSA, RR, $t_{diastolic}$	0.737 (0.734 - 0.740)	0.759	ESV, Age, Killip, RR, $v_{diastolic_avg}$	0.718	0.732
ALL + Vt_{AI}	ESV, EDV, Age, Killip, $Vt_{AI}3$, $Vt_{AI}5$	0.746 (0.743 - 0.749)	0.769	ESV, EDV, Age, Killip, $Vt_{AI}3$, $Vt_{AI}5$	0.729	0.747
ALL + Strain	ESV, Age, Killip, BSA, Weight, Vessels, GLS	0.752 (0.749 - 0.754)	0.776	ESV, Age, Killip, BSA, GLS	0.733	0.748
ALL + LV_{3D}	ESV, EDV, Age, Killip, C5, C16	0.747 (0.745 - 0.749)	0.766	ESV, EDV, Age, C5, C16	0.742	0.753
All + Vt_{AI} + Strain + LV_{3D}	ESV, EDV, Age, C16, GLS, $Vt_{AI}3$, $Vt_{AI}5$	**0.774** (0.771 - 0.777)	**0.796**	ESV, EDV, Age, C16, GLS, $Vt_{AI}5$	**0.761**	**0.773**

Backward stepwise LDA and Cox results of additional prognostic metrics for MACE risk assessment. Considered features include conventional (Vt) and AI-derived (Vt_{AI}) volume transient metrics, CMR-FT strains (strains), and LV 3D detail patterns (LV_{3D}) descriptors, each on top of considering only CMR biomarkers (CMR) or the cardiovascular risk factors and patient characteristics of the study (ALL). A final experiment combines the 3 proposed directions of enhancement. Performance of LVEF is likewise included. The resulting significant selection of variables is reported along with performance, expressed as AUC and C-index (C) re-substitution (rs) and 10-fold cross-validated (k), computed for a 100 random data splits and presented as median (interquartile range). All differences in AUC_k and C_k between models were statistically significant ($P<0.001$), except for the 'CMR' and 'CMR + Vt' resulting Cox models. BSA indicates body surface area (m^2); Killip, Killip class on admission; RR, RR-interval (ms); $t_{diastolic}$, diastolic time (ms); $v_{diastolic_avg}$, average diastolic velocity (mL/s); Vessel, number of diseased vessels.

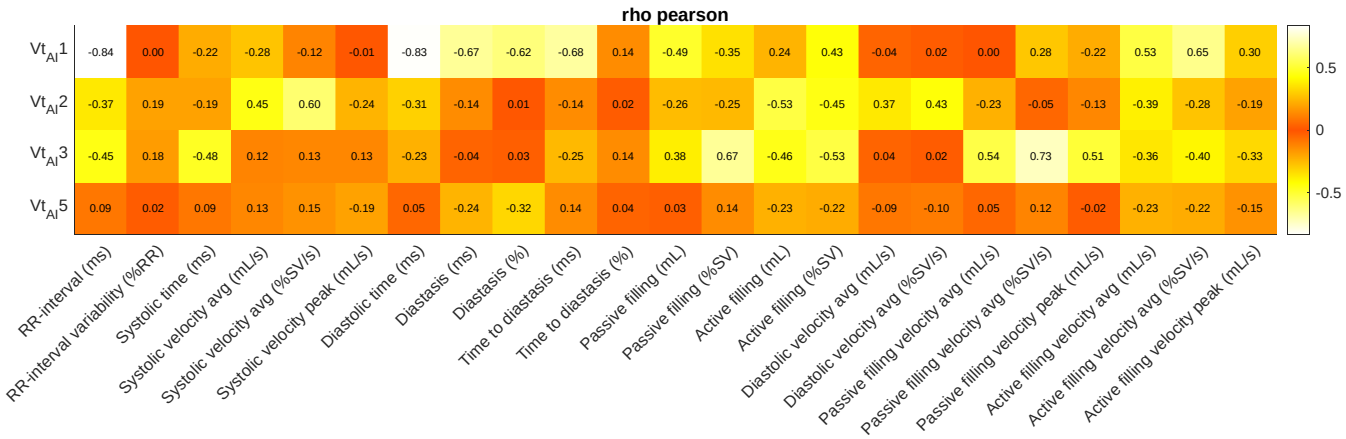


Fig. 3 - Heat map of Pearson correlation coefficients between AI-derived (columns) and conventional (rows) volume transient metrics.

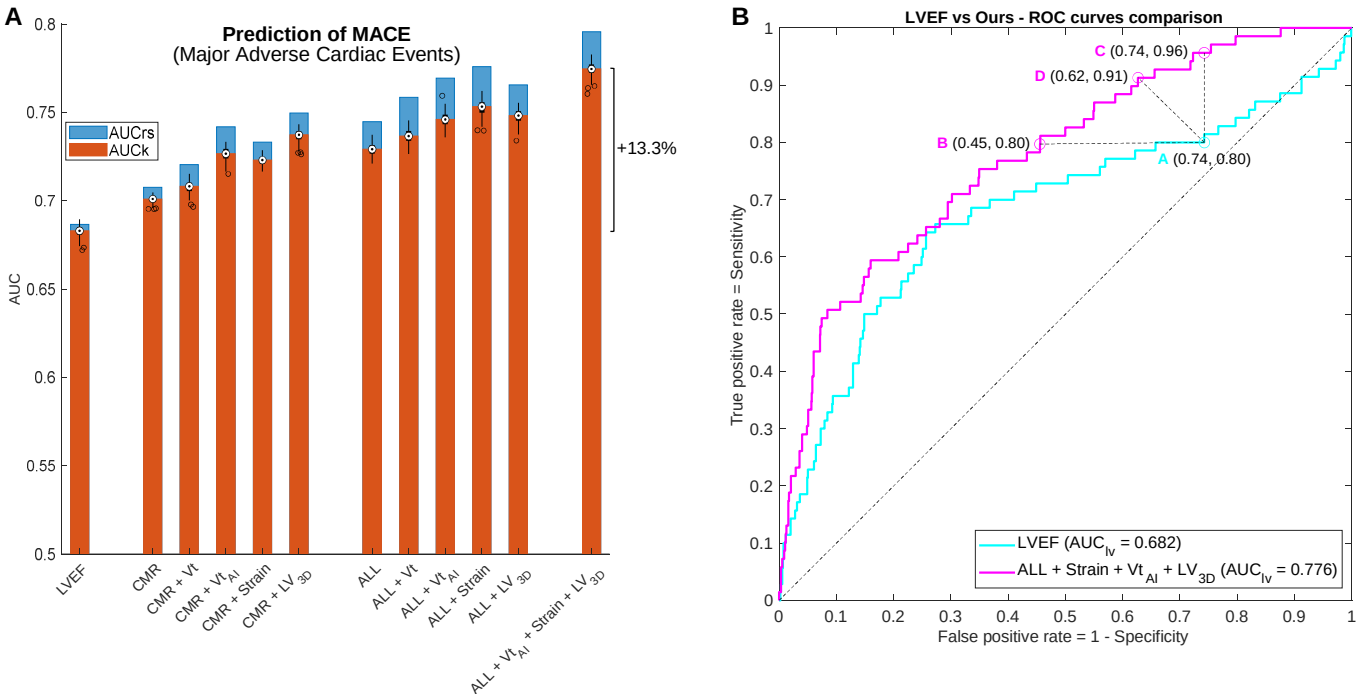


Fig. 4 - **Panel A:** MACE prediction performance results and prognostic contribution of the 3 proposed directions of enhancement. The performance of LVEF, the most established marker for AMI prognosis, is included as reference. Selection of significant variables is based on LDA multivariable stepwise analysis, and results are expressed in AUC resubstitution (rs, blue) and AUC 10-cross-fold validation (k, orange), repeated for a hundred random data splits (black distributions). All differences in performance were significant (P<0.001). Note the remarkable improvement of the final configuration, combining the 3 proposed directions of enhancement. **Panel B:** Interpretation and implications of the AUC differences in ROC prediction curves, illustrated by the comparison between LVEF and our multivariable model, combining the 3 directions of enhancement. Given a recommended operating point of high sensitivity (80%) in the LVEF curve (point A), our model is able to predict MACE at the same sensitivity but reducing the false positive rate a 40% (point B), that is, removing the burden of 379 false positive MACE predictions. Likewise, our proposed model could operate at the similar false positive rate than LVEF but improving the sensitivity a 20% (point C), that is, guessing the 96% of the MACE cases. Alternatively, the operating point D of our model could be chosen for MACE prediction, where both sensitivity and specificity are improved with respect to LVEF. It is therefore shown the potential for risk management improvement associated to superior AUC scores.

1 LV Fully Automated 3D Shape Analysis (LV_{3D})

2 The ES shape and 3D contraction descriptors were significant to MACE (P<0.001) and proved superior
3 to their corresponding stand-alone versions ESV and LVEF (17). Their statistics were reported in (17),
4 and are reproduced in Table SIV.1.

Prognostic Contributions

The results of the augmented prognostic contribution of proposed biomarkers to characterize LV contraction are shown in Table 3. Both LDA and Cox multivariable analysis converged to a nearly identical selection of variables for each of the models.

Prediction of MACE occurrence via conventional and AI automated volume transient metrics, strains and 3D LV markers proved significantly superior (P<0.001) to LVEF using either CMR-derived (CMR: 2.49%; CMR+Vt: 3.66%; CMR+Vt_{AI}: 6.3%; CMR+Strain: 5.86%; CMR+LV_{3D}: 8.05% AUC improvement) or the ALL clinical features (ALL: 6.73%; ALL+Vt: 7.91%; ALL+Vt_{AI}: 9.22%; ALL+Strain: 10.1%; ALL+LV_{3D}: 9.37% AUC improvement), as illustrated in Fig. 4A. Similar results were obtained in Cox analyses. The additional prognostic value was larger (P<0.001) when using AI-derived metrics vs conventional transient ones (Table 3 and Fig. 4A).

The combination of the 3 directions of enhancement led to a major improvement compared to ALL baseline, and to the selection of one variable from each of these directions (GLS, C16 and $Vt_{AI}5$ – and $Vt_{AI}3$, in LDA). In comparison to LVEF, the AUC improvement was superior to 13% in both LDA and Cox analyses (See Fig. 4A and Table 3).

At a univariate level (see Table 2), GLS was the most prognostic variable (AUC_k=0.710), followed by GCS, active filling contribution, and RR-interval. The AI-derived volume transient metrics were not among the most predictive variables, but ended up being the most complementary in multivariable settings.

DISCUSSION

This is the first large-sized multicentre CMR study to date that comprehensively characterises the LV cardiac pump function after AMI combining (1) CMR-FT strain, (2) LV volume temporal transient and (3) 3D shape analyses for risk assessment. We have successfully (a) developed a fully automated AI pipeline, available and easy to run, that segments the CMR SAx stacks and identifies the AMI acute contraction signatures associated to risk; (b) demonstrated the additional prognostic value of cardiac characterization via any of the 3 proposed directions of enhancement; and (c) combined these markers with the established CMR indexes, cardiovascular risk factors and patient characteristics towards a final CMR prediction model, that outdoes LVEF in a 13% (LDA, cross-validated AUC).

Impact and Clinical Translation

While multiple trials have shown the incremental prognostic information of CMR-based risk models in AMI management, these models have not found their role in clinical practice (6,9,14). The more common availability of echocardiography, the length of the CMR acquisitions, the requirement for manual interaction, and the increasing complexity of CMR multi-parameter models have been among the barriers for adoption (4,9). Nevertheless, CMR post-infarction protocols have been shortened and CMR availability has significantly increased in recent years. Moreover, complex CMR prognostic predictors can be now synthetized into simple risk scores models (9), and fully automated as shown in this study. Our work further facilitates the adoption of CMR post-infarction risk management by contributing to two of the pillars of healthcare systems: efficiency and efficacy (3). That is, it leads to an improved prognostic assessment while reducing the analysis time.

Our automated pipeline for volume temporal transients runs on a standard laptop: it receives a CMR scan, segments the LV cavity, calculates the volume transient and summarizes the patient contraction dynamics into 4 biomarkers (Vt_{AI}), that are used in the multivariable models to predict risk. The same applies to the automated shape analysis pipeline presented in (17). This is done in the order of

seconds, removing the burden of manual segmentations, and drastically reducing the analysis time. Moreover, the method is deterministic (i.e. given a scan, it always returns the same outputs) which, in turn, eliminates intra- and interobserver variability towards a more objective, standardized, and quantitative diagnosis. Conventional metrics extracted from volume transients often rely on rather qualitative detection landmarks (e.g. diastasis), which may be problematic especially in pathological cases. Importantly, our proposed Vt_{AI} markers are robust to outliers and abnormal cases. Furthermore, the strong correlation between the automated and manual volumes, along with the high values and small variance of the segmentations dice scores (see Results), supports the accuracy and precision of the automated pipeline, which additionally outdoes commercially available software in performance (35).

LV contraction dynamics are conventionally based on global function metrics evaluated at two single time points (ES and ED, or their ratio, LVEF). The tree proposed approaches to unveil further prognostic value complement each other in the final multi-variate regression model: first, the temporal resolution is increased by the volume transient metrics (from two points to the whole cycle); second, the bulk tissue contraction is assessed by the strain patterns (CMR-FT – from global function to quantification in longitudinal, radial and circumferential directions); and third, the 3D anatomical detail is studied by statistical shape models (17). The relevance of looking at AMI management from all these different angles is evidenced by the results of the final risk-assessment model that combines global function descriptors (ESV, EDV), 3D contraction patterns (C16), strains (GLS), contraction dynamics indexes ($Vt_{AI}3$, $Vt_{AI}5$) and patient characteristics (age).

Our model provides a remarkable boost in MACE prediction: 13% AUC improvement, when comparing against LVEF (Fig. 4A and Table 3), the reference marker according to AMI guidelines (4,6). The implications of these AUC improvements, in terms of specificity and sensitivity, are illustrated in Fig. 4B. The large size of the cohort of this study, the small gap between resubstitution and cross-validation metrics, and the narrow variance when using 100 random data splits (see Fig. 4A and Table 3) suggest the robustness of our method for MACE prediction.

As a result of the automated analysis, we have built a reference LV volume transient atlas from 1000+ AMI subjects that captures the average post-infarct contraction behavior along with the main variations (PCA modes). The atlas and the associated methods to calculate the herein proposed Vt_{AI} markers and the resulting risk models, made publicly available to the community (doi.org/10.6084/m9.figshare.16735300), will not only allow the validation of the presented findings but also further AMI studies, computer simulations, and synthetic patient data generation for training algorithms or educational use, among others. Moreover, while the AI decomposition methodology has been deployed using CRM volume transients, its application to echocardiography is straightforward. The segmentation algorithm, extendable to different protocols and scanners to the ones used in this study as well as to other cardiac diseases (36), is likewise available.

Interpretability: Unravelling the AMI Contraction

While infarct severity is associated with functional and morphological alterations, and certain acute changes have proven significant to AMI prognosis, their interplay in modulating risk remains unsolved (6,13,37,38). In this regard, we have identified contraction traits, in the form of strains and LV volume transient patterns, significantly relevant to MACE occurrence.

Our analysis of cardiac function characterized by conventional descriptors of the volume transients has identified the RR-interval length and the active and passive contributions to filling, in mL, as particularly related to MACE occurrence (see Table 2). Since patients at risk exhibit a decreased LVEF (4,6), their passive and active contributions are accordingly reduced and, in consequence, their RR-interval shortened to increase cardiac output (39). This, in turn, leads to a smaller preload and a less effective contraction, with the corresponding drop in LVEF (negative feedback loop).

RR variability was also found to be highly prognostic. It can be hypothesized that this beat-to-beat variation could be related to atrial fibrillation (AF), as evidenced in (40), especially given the mentioned fast heart rates, related to ventricular tachycardia and fibrillation risk according to the MADIT-ICD score (41), that characterize our MACE population. Nevertheless, and oppositely, the RR variability has been shown beneficial in other AF disease models (42).

One of the advantages of the AI deployed methodology is that the resulting Vt_{AI} variables are not obscure features for prediction improvement, but they rather encode explainable contraction patterns, as shown in Fig. 2. Thus, $Vt_{AI}1$ can be primarily interpreted as RR-interval length (r_p = -0.84); $Vt_{AI}2$, as average systolic velocity (r_p = 0.60); $Vt_{AI}3$, as balance between the active and passive contributions (passive filling: r_p = 0.67); and $Vt_{AI}5$, as diastasis once corrected for RR-interval length (diastasis: r_p = -0.32; r_s = -0.43. See Supplemental Material VI). Nevertheless, these contraction variations contain more information than just a single conventional metric, reason why they are also correlated to other variables (i.e. in a contraction with reduced RR-interval the diastolic and diastasis times will be likewise compromised; therefore, $Vt_{AI}1$, correlated to RR-interval, is also correlated to these indexes) and reason why they led to greater prognostic contributions.

In this context, the minmax normalization prior to the AI analysis, standardizing the transients by removing ventricle size and SV information, should be considered when visually interpreting the results (See Supplemental Material II). This normalization decreases the univariate predictive power of the Vt_{AI} markers to the benefit of removing cofounding factors and facilitating their contribution to multivariable settings. $Vt_{AI}3$ and $Vt_{AI}5$, both interpreted as related to diastolic function, are the AI-derived transient markers that contribute the most towards additional prognosis in the multivariable settings (Table 3). While mechanistic explanations are to be further explored, a plausible interpretation is that, in patients at risk, RR-interval shortening compromises ventricular relaxation. This may come at the expense of diastasis ($Vt_{AI}5$), and/or faster dynamics with a swift in the relative filling contributions (%SV) towards the passive phase ($Vt_{AI}3$). This in turn explains the consequent significant increase in normalized passive filling velocity (note that the actual passive filling velocities, in mL/s, without normalization, are smaller for the MACE population, due to a smaller LVEF – see Table 2). This is particularly relevant for the STEMI population (see Fig. 2).

Likewise, the 3D LV patterns, interpreted and digested in (17), demonstrated consistent with the three predominant LV remodeling descriptors according to the literature: ESV, LFEF and ES myocardial thickness (37,38). The interplay between these myocardial 3D shape changes and micro-damage, in the form of IS and MVO, was also evidenced.

All the CMR-FT strain metrics reported were highly significant to MACE occurrence (Table 2). The compromised GLS, GRS and GCS that the MACE population exhibits are in harmony with the LVEF decrease associated to risk (4,6). Besides, the observed relationship between risk and impaired contraction uniformity (CURE and RURE) is widely discussed in (9). The prognostic value of GLS, superior to LVEF and incremental to the baseline markers, is in accordance with (43), that argues for its adoption in clinical practice.

The proposed risk models have superior prognostic value for patients with decreased LV function (LVEF<35%), identified as high-risk group, in accordance with the literature (4,10) (Fig. SV.2). With respect to infarct aetiology, the risk models that include CMR-FT strain or 3D LV patterns are more predictive for the NSTEMI population, despite the study being biased to STEMI patients (see Fig. SV.3 and Results - Patients). The final model, that combines strains, AI transient metrics and 3D LV metrics, reports a similar performance for STEMI (AUC = 0.795) and NTEMI (AUC = 0.799), and, thus, clinical validity for both infarct types could be claimed. If analyzed separately, both systolic and diastolic components significantly contribute to the multivariable risk assessment models (see Supplemental Material VII). However, as expected, the combination of systolic and diastolic function, that is, the entire transient, is more prognostic than each of the components on its own, especially in the systolic

scenario (in line with the fact that the modes that provide more additional prognostic value, $Vt_{AI}3$ and $Vt_{AI}5$, are related to diastolic function, as mentioned above).

To sum up, patients at risk of MACE occurrence have a compromised cardiac function, which contributes to an impaired global contraction (decreased LVEF and GLS) and, therefore, to a reduced RR-interval, in order to maintain the cardiac output. This, in turn, leads to diastolic impairment, exhibited in the form of compromised diastasis and active filling.

Limitations

Patients were imaged within 10 days after infarction in the absence of optimal post-infarction CMR imaging time recommendations (9). The effect of this post-infarction imaging time on the proposed AI-derived metrics has not been assessed. It is however hypothesized that we are capturing acute injury, whose extent and phenotype can predict outcome, and therefore long-term LV adaptions. CMR long-axis views were excluded for simplicity, but ideally could be incorporated to correct for SAx misalignment and improve the cavity volume calculations. Likewise, while only CMR-FT global strains and uniformity ratio estimates have proved prognostic in the literature (7,9,13,18,19), future work could consider a more detailed strain analysis, in time and by regions, especially given the outstanding prognostic power of GLS. Similar methodologies to the PCA analysis formulated in this study could be considered for this. The resolution of the volume transients herein derived is limited to the resolution of the CMR scan. Given the limited number of MACE events (class imbalance), splitting the cohort into an independent testing dataset would be rather noisy. The combination of cross-fold validation, repeated for a hundred data-splits, is therefore preferred as it ensures fair performance and robust conclusions reporting, as argued in (32,34). Finally, the results are limited to AMI patients that can undergo a CMR scan.

CONCLUSION

This large multicenter study evidences the prognostic value of the enhanced cardiac characterization using CMR-FT strains, LV volume temporal transients and 3D shape analysis in post-infarction management. Besides, it further contributes towards the adoption of CMR multivariable models in clinical practice by demonstrating the feasibility of fully automated analysis and the relevance of these models in AMI stratification.

ACKNOWLEDGMENTS

This work was supported by the EU's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie (g.a. 764738), the DZHK (German Center for Cardiovascular Research), the British Heart Foundation (PG/16/75/32383, FS/17/22/32644), and the Wellcome Trust (209450/Z/17).

Conflict of interest: none declared.

REFERENCES

1. Virani SS, Alonso A, Benjamin EJ, et al. Heart Disease and Stroke Statistics - 2020 Update: A Report From the American Heart Association. Circulation. 2020;141(9):e139–596.

2. James SL, Abate D, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990-2017: A

systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1789–858.

3. Corral Acero J, Margara F, Marciniak M, et al. The 'Digital Twin' to enable the vision of precision cardiology. Eur Heart J 2020; 41:4556–64.

4. Ibanez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). Eur Heart J 2017; 39(2):119–77

5. Anderson JL, Morrow DA. Acute Myocardial Infarction. N Engl J Med 2017; 376(21):2053–64.

6. Eitel I, de Waha S, Wöhrle J, et al. Comprehensive prognosis assessment by CMR imaging after ST-segment elevation myocardial infarction. J Am Coll Cardiol 2014; 64(12):1217–26.

7. Amzulescu MS, de Craene M, Langet H, et al. Myocardial strain imaging: review of general principles, validation, and sources of discrepancies. Eur Heart J Cardiovasc Imaging 2019; 20(6):605–19

8. Smiseth OA, Torp H, Opdahl A, Haugaa KH, Urheim S. Myocardial strain imaging: how useful is it in clinical decision making? Eur Heart J 2015; 37(15):1196–207.

9. Stiermaier T, Backhaus SJ, Lange T, et al. Cardiac Magnetic Resonance Left Ventricular Mechanical Uniformity Alterations for Risk Assessment After Acute Myocardial Infarction. J Am Heart Assoc 2019; 8(16):e011576

10. Dagres N, Hindricks G. Risk stratification after myocardial infarction: is left ventricular ejection fraction enough to prevent sudden cardiac death? Eur Heart J 2013; 34(26):1964–71.

11. Bulluck H, Dharmakumar R, Arai AE, Berry C, Hausenloy DJ. Cardiovascular Magnetic Resonance in Acute ST-Segment–Elevation Myocardial Infarction. Circulation 2018; 137(18):1949–64

12. Dall'Armellina E. From Recognized to Novel Quantitative CMR Biomarkers of LV Recovery. JACC Cardiovasc Imaging 2017; 10(9):1000–2.

13. Reindl M, Eitel I, Reinstadler SJ. Role of Cardiac Magnetic Resonance to Improve Risk Prediction following Acute ST-elevation Myocardial Infarction. J Clin Med 2020; 9(4):1041.

14. Stiermaier T, Jobs A, de Waha S, et al. Optimized Prognosis Assessment in ST-Segment-Elevation Myocardial Infarction Using a Cardiac Magnetic Resonance Imaging Risk Score. Circ Cardiovasc Imaging 2017; 10(11):e006774.

15. Masci PG, Ganame J, Francone M, et al. Relationship between location and size of myocardial infarction and their reciprocal influences on post-infarction left ventricular remodelling. Eur Heart J 2011 Jul; 32(13):1640–8.

16. Ortiz-Pérez JT, Rodríguez J, Meyers SN, Lee DC, Davidson C, Wu E. Correspondence Between the 17-Segment Model and Coronary Arterial Anatomy Using Contrast-Enhanced Cardiac Magnetic Resonance Imaging. JACC Cardiovasc Imaging 2008; 1(3):282–93.

17. Corral Acero J, Schuster A, Zacur E, et al. Understanding and Improving Risk Assessment after Infarction using Automated Left Ventricular Shape Analysis. Under Review in *JACC Cardiovasc Imaging*. See: https://drive.google.com/drive/folders/1m78M0XGX3gqMkZNCNg7IYKT4BlsqF4-7?usp=sharing

18. Mangion K, McComb C, Auger DA, Epstein FH, Berry C. Magnetic resonance imaging of myocardial strain after acute st-segment-elevation myocardial infarction a systematic review. Circ Cardiovasc Imaging. 2017; 10(8):e006498

19. Eitel I, Stiermaier T, Lange T, et al. Cardiac Magnetic Resonance Myocardial Feature Tracking for Optimized Prediction of Cardiovascular Events Following Myocardial Infarction. JACC Cardiovasc Imaging 2018; 11(10):1433–44

20. Ruijsink B, Puyol-Antón E, Oksuz I, et al. Fully Automated, Quality-Controlled Cardiac Analysis From CMR: Validation and Large-Scale Application to Characterize Cardiac Function. JACC Cardiovasc Imaging 2020; 13(3):684–95.

21. Thiele H, de Waha S, Zeymer U, et al. Effect of aspiration thrombectomy on microvascular obstruction in NSTEMI patients: The TATORT-NSTEMI trial. J Am Coll Cardiol 2014; 64(11):1117–24.

22. Eitel I, Wöhrle J, Suenkel H, et al. Intracoronary compared with intravenous bolus abciximab application during primary percutaneous coronary intervention in ST-segment elevation myocardial infarction: Cardiac magnetic resonance substudy of the AIDA STEMI trial. J Am Coll Cardiol 2013; 61(13):1447–54.

23. Thiele H, Wöhrle J, Hambrecht R, et al. Intracoronary versus intravenous bolus abciximab during primary percutaneous coronary intervention in patients with acute ST-elevation myocardial infarction: A randomised trial. Lancet 2012; 379(9819):923–31.

24. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020; 368:m441.

25. Backhaus SJ, Metschies G, Zieschang V, et al. Head-to-head comparison of cardiovascular MR feature tracking cine versus acquisition-based deformation strain imaging using myocardial tagging and strain encoding. Magn Reson Med 2021; 85(1):357–68

26. Gertz RJ, Lange T, Kowallick JT, et al. Correction: Inter-vendor reproducibility of left and right ventricular cardiovascular magnetic resonance myocardial feature-tracking. PLoS One 2018; 13(6):e0199489

27. Kowallick JT, Morton G, Lamata P, et al. Quantitative assessment of left ventricular mechanical dyssynchrony using cine cardiovascular magnetic resonance imaging: Inter-study reproducibility. JRSM Cardiovasc Dis 2017; 6:1:8.

28. Corral Acero J, Zacur E, Xu H, et al. SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI. In: Coudière Y. et al. (eds). FIMH 2019. LNCS. Springer, Cham; 2019; 11504:361-69

29. Corral Acero J, Xu H, Zacur E, et al. Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation. In: Pop M. et al. (eds). STACOM 2019. LNCS. Springer, Cham. 2019; 12009:384–94.

30. Yang G, Hua T, Xue W, Shuo L. Left Ventricle Full Quantification Challenge MICCAI 2019. https://lvquan19.github.io/ (29 September 2021).

31. Belhumeur PN, Hespanha JP, Kriegman DJ. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 1997; 19(7):711–20.

32. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 2015; 162(1):W1-W73

33. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the Yield of Medical Tests. J Am Med Assoc 1982; 247(18):2543–6

34. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. Ann Intern Med 2019; 170:W1-W33.

35. Schuster A, Lange T, Backhaus SJ, et al. Fully Automated Cardiac Assessment for Diagnostic and Prognostic Stratification Following Myocardial Infarction. J Am Heart Assoc 2020; 9(18):e016612.

36. Corral Acero J, Sundaresan V, Dinsdale N, Grau V, Jenkinson M. A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation. In: Puyol E. et al. (eds) STACOM 2020. LNCS. Springer, Cham; 2021; 12592:196–207.

37. Cohn JN, Ferrari R, Sharpe N. Cardiac remodeling-concepts and clinical implications: A consensus paper from an International Forum on Cardiac Remodeling. J Am Coll Cardiol 2000; 35(3):569–82.

38. Hassell ME, Vlastra W, Robbers L, et al. Long-term left ventricular remodelling after revascularisation for ST-segment elevation myocardial infarction as assessed by cardiac magnetic resonance imaging. Open heart 2017; 4(1):e000569.

39. Jabre P, Roger VL, Weston SA, et al. Resting heart rate in first year survivors of myocardial infarction and long-term mortality: A community study. Mayo Clin Proc 2014; 89(12):1655–63.

40. Lyon A, van Mourik M, Cruts L, et al. Both beat-to-beat changes in RR-interval and left ventricular filling time determine ventricular function during atrial fibrillation. EP Europace 2021; 23:i21-i28

41. Younis A, Goldberger JJ, Kutyifa V, et al. Predicted benefit of an implantable cardioverter-defibrillator: the MADIT-ICD benefit score. Eur Heart J 2021; 42(17):1676-84.

42. Brateanu A. Heart rate variability after myocardial infarction: What we know and what we still need to find out. Curr Med Res Opin. 2015;31(10):1855-60.

43. Reindl M, Tiller C, Holzknecht M, et al. Prognostic Implications of Global Longitudinal Strain by Feature-Tracking Cardiac Magnetic Resonance in ST-Elevation Myocardial Infarction. Circ Cardiovasc Imaging 2019; 12(11):e009404.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Comprehensive Characterization of Cardiac Function for Improved Post-Infarction Risk Assessment
Publication Status	□ Published □ Accepted for Publication ☑ Submitted for Publication □ Unpublished and unsubmitted work written in a manuscript style
Publication Details	Jorge Corral Acero, Pablo Lamata, Ingo Eitel, Ernesto Zacur, Ruben Evertz, Torben Lange, Sören J. Backhaus, Thomas Stiermaier, Holger Thiele, Alfonso Bueno-Orovio, Andreas Schuster, Vicente Grau, "Comprehensive Characterization of Cardiac Function for Improved Post-Infarction Risk Assessment". Submitted to *Circulation* on 5th October 2021.

Student Confirmation

Student Name:	Jorge Corral Acero		
Contribution to the Paper	(1) Pre-processed the data; (2) developed and implemented methods; (3) designed and ran the experiments; (4) analysed the results; (5) wrote the manuscript; and (6) coordinated the collaboration.		
Signature		Date	15/11/2021

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. Vicente Grau		
Supervisor comments: I confirm that the above statement is accurate.		
Signature 	Date	15/11/2021

This completed form should be included in the thesis, at the end of the relevant chapter.

CHAPTER 7

Discussion and Conclusions

7. Discussion and Conclusions

7.1. Key Contributions and Main Findings

The aims of this thesis were to develop novel computational tools for the characterisation of cardiac structure and function, and to provide novel insights on how their interplay modulates adverse cardiac outcomes. The thesis largely focuses on acute myocardial infarction as the deadliest cardiovascular disease, while also presenting applications to other inherited cardiomyopathies, such as hypertrophic cardiomyopathy. The outcomes of this thesis are expected to ultimately reduce the socio-economic burden of cardiovascular disease, by means of a personalised and preventive medicine. In this regard, the thesis has successfully contributed to these directions as follows:

Firstly, it has reviewed the important role of computational modelling towards precision and preventive cardiovascular medicine, enabled by the recent big data expansion in cardiology and the increasing computational power. The vision of a synergistic combination between statistical and mechanistic modelling towards the future of precision cardiology is presented and discussed.

Secondly, a fully automated pipeline to accurately obtain the 3D shape of the heart from standard 2D CMR scans has been developed (Research Objective 1). The pipeline segments the 2D images, using state-of-the-art deep learning approaches, and generates the personalised heart meshes of the patient in less than 2 minutes on a standard laptop. To achieve this, novel methods were deployed to explicitly address the main challenges towards the automation of CMR segmentation: anatomical variability, imaging heterogeneity, labels imbalance, and limited clinical data availability for training. The resulting segmenter successfully removes the burden of manual annotation, which is costly, tedious, and often subjected to significant intra- and inter-variability as important confounders for diagnosis. Moreover, the proposed segmentation NNs were coupled to a meshing method that guarantees point-correspondence and robustness. This ensures 3D-spatial consistency, which leads to a more accurate quantification of volumes and, therefore, an improved prognosis compared to using standard clinical metrics (as shown in Chapter 5).

Thirdly, the Thesis has introduced an enhanced analysis of the heart anatomy (Research objective 2) and cardiac contraction (Research objective 3) to better understand the structural-functional interplay that modulates adverse outcomes. Three directions of enhanced characterisation have been proposed. Shape analysis provides anatomical 3D detail, while CMR-FT strains further characterise the properties of the heart at its tissue level, and the analysis of volume transients disentangle the temporal dynamics of LV contraction. Moreover, this comprehensive analysis of function and structure was coupled to ML techniques for the identification of interpretable patterns related to risk. All this was made possible by the automated pipeline developed before. The pipeline was accordingly expanded to not only outputting the 3D heart shape of the patient, but also to accurately estimate the LV volume transients, along with the cardiac cycle metrics.

Fourthly, further to the clinical impact of a full automation of CMR analysis, the pipeline developed in this Thesis, along with the enhanced analysis that it enables, has demonstrated its prognostic value on AMI. To strengthen the robustness of the clinical findings, a large cohort of AMI patients with one-year follow-up was considered. In this scenario, each of the three proposed directions of enhancement proved to significantly provide additional prognostic information towards a better post-infarction management. These three directions were further combined on top of established CMR indexes and cardiovascular risk factors towards a final prediction model. This synergistic combination reported a 13% AUC improvement with respect to the risk model established in the most recent AMI guidelines. While the implications derived from these results on saved lives and reduced healthcare costs can only be fully ascertained with further independent clinical trials, the margin is sufficiently solid to assume a substantial improvement of AMI prognosis and management.

Finally, the enhanced analysis in the AMI cohort has identified 3D shape and contraction patterns that are related to adverse outcomes. These patterns are interpretable and can be quantitatively described with a few continuous variables, thanks to the PCA approach that the enhanced analysis integrates. As a summary of main findings, patients at risk of MACE occurrence present a

compromised cardiac function, which contributes to an impaired global contraction. This is shown as a decrease in LVEF and global longitudinal strains, together with an increase in ESV, leading to ES thinning in consequence. As a result of the compromised function, patients at risk experience a shortening of the R-R interval to maintain the cardiac output. This, in turn, leads to diastolic impairment, exhibited as compromised diastasis and active filling. In addition, basal impairment and LV akinesia in the anterior wall showed to further increase the risk of MACE occurrence. A detailed digestion of 3D anatomical and contraction patterns can be found in in Chapters 5 and 6. Therefore, the proposed analysis proved capable of shedding new light on the structural-functional remodelling that modulates post-infarction outcomes. Hence, the overall goal of the thesis, linking cardiac structure and function as modulators of adverse cardiac events, is met.

To accelerate the vision proposed in this Thesis, the methods and computational atlases constructed will be made publicly available upon acceptance of the final papers in progress to be published.

7.2. Future Work

Future work could explore alternative methodological implementations to the ones presented in this Thesis, overcoming some of their limitations. Recently proposed DL approaches, such as Point-Net (see Chapter 2), could incorporate the long axis CMR acquisitions into proposed pipeline for LV shape regression. This could potentially improve the correction for SAx misalignment, and to remove noise in estimating the length of the ventricle. It could also potentially reduce the analysis time to less than a second per patient. However, this would come at the cost of mesh point-correspondence and robustness. In consequence, further methodological developments or alternative shape analysis pipelines would be required to overcome these limitations. Other DL approaches, in particular VAEs, could also be considered for the dimensionality reduction of the generated meshes. Likewise, more advanced ML classifiers could be deployed for improved outcome predictions. The adoption of such

approaches would most likely report an improvement in performance. However, it is still to be determined whether the improvement would be significant enough to justify a potential loss of interpretability. Moreover, an augmentation of the instances in appearance could also be explored to even further improve the CMR segmentation. Finally, SMOD could be implemented in 3D, addressing its current limitation to 2D planes.

Next steps regarding the AMI case study could include a more profound analysis of the CMR-FT strains. So far, the analysis presented in this thesis has been restricted to metrics previously shown of prognostic value, that is, global strains and uniformity ratio estimates. Similar methodologies to the PCA analysis for LV volume transients could be considered for this, especially given the outstanding prognostic power of global longitudinal strain. Furthermore, the additional prognostic information of later chronic remodelling (i.e., longitudinal scans at 3-6 months post-AMI), could be incorporated in the analysis. The current analysis is limited to the LV, whose remodelling has proven crucial in AMI prognosis. However, the RV could also be modelled and analysed similarly to the LV for complementary insights into the disease. While the present study identifies acute shape and contraction patterns related to risk and proves their impact on patient management, future work could relate these patterns to specific long-term remodelling and, thus, underpin personalised optimal treatment for patients at risk (e.g. reperfusion in cases of ischemic but viable myocardium due to an unresolve lack of perfusion or cardiac resynchronization therapy implants in cases of ventricular tachycardia or atrial fibrillation risk). The findings of the study were properly validated, via cross-fold validation (repeated for a large number of random data splits) and reproducibility analysis. In addition to this, they could be tested in an independent cohort to further validate the findings, and promote their adoption in clinical practice.

Alternative avenues of future work could use the herein deployed methods and enhanced analyses, proven useful for AMI management, to investigate other cardiomyopathies. Even further, they could be applied to some other diseases or alternative imagining modalities, upon certain modifications and

fine tuning. Inspired by the vision presented in Chapter 3, the methods developed in this Thesis ensure that the resulting 3D anatomical models can be directly used in mechanistic simulations (i.e., watertight and smooth LV meshes). Thus, the Thesis pipeline could be coupled to mechanics models of cardiac function, or emulators of them, allowing for personalised 3D computer simulations from standard clinical 2D CMR scans. This translates into the possibility of inferring mechanistical properties of the tissue, such us contractility or stiffness, from just CMR images. Finally, the 3D LV computational atlases can be useful to other AMI shape studies, and for the generation of synthetic patients for training algorithms or educational use, as mentioned before.

7.3. Conclusion

To conclude, this Thesis builds a refined understanding on the structural and functional interplay that modulates risk in AMI, by deploying of state-of-art tools, which automatise the CMR analysis. These ease the clinical workflow and pave the road towards the use of multivariate models and CMR imaging in clinical routine. Ultimately, the Thesis demonstrates the potential of 3D computational models of the heart for CVD management, towards preventive and personalised cardiovascular medicine.

REFERENCES

REFERENCES

[1] E. M. Antman and J. Loscalzo, “Precision medicine in cardiology,” *Nature Reviews Cardiology*, vol. 13, no. 10. Nature Publishing Group, pp. 591–602, Oct. 01, 2016, doi: 10.1038/nrcardio.2016.101.

[2] J. Corral-Acero *et al.*, “The ‘Digital Twin’ to enable the vision of precision cardiology,” *European Heart Journal*, Mar. 2020, doi: 10.1093/eurheartj/ehaa159.

[3] World Health Organization (WHO), “Cardiovascular diseases (CVDs),” Available: https://www.who.int/cardiovascular_diseases/en/ (accessed Jan. 25, 2019).

[4] Wilkins E. *et al.*, “European Cardiovascular Disease Statistics 2017 edition,” *European Heart Network, Brussels*, p. 192, 2017, doi: 978-2-9537898-1-2.

[5] N. Townsend, J. Williams, P. Bhatnagar, K. Wickramasinghe, and M. Rayner Nuffield, “Cardiovascular Disease Statistics 2014,” British Heart Foundation Centre on Population Approaches for Non-Communicable Disease Prevention, Report no. SC039426. 2014.

[6] S. S. Virani *et al.*, “Heart Disease and Stroke Statistics - 2020 Update: A Report From the American Heart Association,” *Circulation*, vol. 141, no. 9, pp. e139–e596, 2020, doi: 10.1161/CIR.0000000000000757.

[7] J. L. Anderson and D. A. Morrow, “Acute Myocardial Infarction,” *New England Journal of Medicine*, vol. 376, no. 21, pp. 2053–2064, May 2017, doi: 10.1056/NEJMra1606915.

[8] S. Zhang, L. Li, W. Guo, L. Wang, X. Yan, and C. Zhang, “Observation on left ventricular remodeling in acute myocardial infarction.,” *Zhonghua nei ke za zhi*, vol. 38, no. 2, pp. 107–9, Feb. 1999.

[9] E. Dall’Armellina, “From Recognized to Novel Quantitative CMR Biomarkers of LV Recovery,” *JACC: Cardiovascular Imaging*, vol. 10, no. 9, pp. 1000–1002, Sep. 2017, doi: 10.1016/j.jcmg.2016.07.007.

[10] M. E. Hassell *et al.*, “Long-term left ventricular remodelling after revascularisation for ST-segment elevation myocardial infarction as assessed by cardiac magnetic resonance imaging.,” *Open heart*, vol. 4, no. 1, p. e000569, 2017, doi: 10.1136/openhrt-2016-000569.

[11] B. Ibanez *et al.*, “2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Socie,” *European Heart Journal*, vol. 39, no. 2, pp. 119–177, Aug. 2017, doi: 10.1093/eurheartj/ehx393.

[12] S. L. James *et al.*, “Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017,” *The Lancet*, vol. 392, no. 10159, pp. 1789–1858, Nov. 2018, doi: 10.1016/S0140-6736(18)32279-7.

[13] J. S. Rumsfeld, K. E. Joynt, and T. M. Maddox, “Big data analytics to improve cardiovascular care: promise and challenges,” *Nature Reviews Cardiology*, vol. 13, no. 6, pp. 350–359, Jun. 2016, doi: 10.1038/nrcardio.2016.42.

[14] K. Shameer, K. W. Johnson, B. S. Glicksberg, J. T. Dudley, and P. P. Sengupta, “Machine learning in cardiovascular medicine: are we there yet?,” *Heart*, vol. 104, no. 14, pp. 1156–1164, Jul. 2018, doi: 10.1136/heartjnl-2017-311198.

[15] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/J.MEDIA.2017.07.005.

[16] E. Ranschaert, S. Morozov, and P. R. Algra, *Artificial Intelligence in Medical Imaging*, 1st ed. Cham: Springer International Publishing, 2019.

[17] X. Yi, E. Walia, and P. Babyn, "Generative Adversarial Network in Medical Imaging: A Review," *Medical Image Analysis*, vol. 58, Dec. 2019, doi:10.1016/j.media.2019.101552

[18] M. B. Stokes and R. Roberts-Thomson, "The role of cardiac imaging in clinical practice.," *Australian prescriber*, vol. 40, no. 4, pp. 151–155, Aug. 2017, doi: 10.18773/austprescr.2017.045.

[19] M. A. Peterzan, O. J. Rider, and L. J. Anderson, "The Role of Cardiovascular Magnetic Resonance Imaging in Heart Failure.," *Cardiac failure review*, vol. 2, no. 2, pp. 115–122, Nov. 2016, doi: 10.15420/cfr.2016.2.2.115.

[20] A. Suinesiaputra, A. D. McCulloch, M. P. Nash, B. Pontre, and A. A. Young, "Cardiac image modelling: Breadth and depth in heart disease," *Medical Image Analysis*, vol. 33, pp. 38–43, Oct. 2016, doi: 10.1016/J.MEDIA.2016.06.027.

[21] K. W. Johnson *et al.*, "Artificial Intelligence in Cardiology," *Journal of the American College of Cardiology*, vol. 71, no. 23. 2018, doi: 10.1016/j.jacc.2018.03.521.

[22] N. Forsch *et al.*, "Computational analysis of cardiac structure and function in congenital heart disease: Translating discoveries to clinical strategies," *Journal of Computational Science*, vol. 52, p. 101211, May 2021, doi: 10.1016/J.JOCS.2020.101211.

[23] K. W. Johnson *et al.*, "Enabling Precision Cardiology Through Multiscale Biology and Systems Medicine," *JACC: Basic to Translational Science*, vol. 2, no. 3. Elsevier Inc, pp. 311–327, Jun. 01, 2017, doi: 10.1016/j.jacbts.2016.11.010.

[24] S. A. Niederer, J. Lumens, and N. A. Trayanova, "Computational models in cardiology," *Nature Reviews Cardiology*, vol. 16, no. 2. 2019, doi: 10.1038/s41569-018-0104-y.

[25] D. Dey *et al.*, "Artificial Intelligence in Cardiovascular Imaging: JACC State-of-the-Art Review," *Journal of the American College of Cardiology*, vol. 73, no. 11. 2019, doi: 10.1016/j.jacc.2018.12.054.

[26] P. J. Slomka, D. Dey, A. Sitek, M. Motwani, D. S. Berman, and G. Germano, "Cardiac imaging: working towards fully-automated machine analysis & interpretation," *Expert Review of Medical Devices*, vol. 14, no. 3. Taylor and Francis Ltd, pp. 197–212, Mar. 04, 2017, doi: 10.1080/17434440.2017.1300057.

[27] C. Dai *et al.*, "Suggestive Annotation of Brain Tumour Images with Gradient-Guided Sampling," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 156–165, doi: 10.1007/978-3-030-59719-1_16.

[28] S. Wang *et al.*, "Deep Generative Model-Based Quality Control for Cardiac MRI Segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 88–97, doi: 10.1007/978-3-030-59719-1_9.

[29] C. Krittanawong, A. Tunhasiriwet, H. J. Zhang, Z. Wang, M. Aydar, and T. Kitai, "Deep Learning With Unsupervised Feature in Echocardiographic Imaging," *Journal of the American College of Cardiology*, vol. 69, no. 16. Elsevier USA, pp. 2100–2101, Apr. 25, 2017, doi: 10.1016/j.jacc.2016.12.047.

[30] W. Xue, G. Brahm, S. Pandey, S. Leung, and S. Li, "Full left ventricle quantification via deep multitask relationships learning," *Medical Image Analysis*, vol. 43, pp. 54–65, Jan. 2018, doi: 10.1016/J.MEDIA.2017.09.005.

[31] W. Xue, A. Islam, M. Bhaduri, and S. Li, "Direct Multitype Cardiac Indices Estimation via Joint Representation and Regression Learning," *IEEE Transactions on Medical Imaging*, vol. 36, no. 10, pp. 2057–2067, Oct. 2017, doi: 10.1109/TMI.2017.2709251.

[32] A. Raju *et al.*, "User-Guided Domain Adaptation for Rapid Annotation from User Interactions: A Study on Pathological Liver Segmentation," in *MICCAI 2020*, pp. 457–467, doi: 10.1007/978-3-030-59710-8_45.

[33] S. Ali *et al.*, "Deep learning for detection and segmentation of artefact and disease instances in gastrointestinal endoscopy," *Medical Image Analysis*, vol. 70, p. 102002, May 2021, doi: 10.1016/J.MEDIA.2021.102002.

[34] T. A. Nketia, H. Sailem, G. Rohde, R. Machiraju, and J. Rittscher, "Analysis of live cell images: Methods, tools and opportunities," *Methods*, vol. 115, pp. 65–79, Feb. 2017, doi: 10.1016/J.YMETH.2017.02.007.

[35] C. Martin-Isla *et al.*, "Image-Based Cardiac Diagnosis With Machine Learning: A Review," *Frontiers in Cardiovascular Medicine*, vol. 7. Frontiers Media S.A., Jan. 24, 2020, doi: 10.3389/fcvm.2020.00001.

[36] R. Arsanjani *et al.*, "Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population," *Journal of Nuclear Cardiology*, vol. 20, no. 4, pp. 553–562, Aug. 2013, doi: 10.1007/s12350-013-9706-2.

[37] D. Dey *et al.*, "Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: A multicentre study," *European Radiology*, vol. 28, no. 6, pp. 2655–2664, Jan. 2018, doi: 10.1007/s00330-017-5223-z.

[38] B. Bozorgtabar, D. Mahapatra, G. Vray, and J.-P. Thiran, "SALAD: Self-supervised Aggregation Learning for Anomaly Detection on X-Rays," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 468–478, doi: 10.1007/978-3-030-59710-8_46 .

[39] A. Mayr, H. Binder, O. Gefeller, and M. Schmid, "The evolution of boosting algorithms: From machine learning to statistical modelling," *Methods of Information in Medicine*, vol. 53, no. 6, pp. 419–427, Jan. 2014, doi: 10.3414/ME13-01-0122.

[40] O. Ronneberger, P. Fischer, T. Brox, and P. and B. T. Ronneberger Olaf and Fischer, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2015*, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[41] W. Bai *et al.*, "Automated cardiovascular magnetic resonance image analysis with fully convolutional networks," *Journal of Cardiovascular Magnetic Resonance*, vol. 20, no. 1, p. 65, Dec. 2018, doi: 10.1186/s12968-018-0471-x.

[42] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features," *IEEE Transactions on Medical Imaging*, vol. 27, no. 11, pp. 1668–1681, Nov. 2008, doi: 10.1109/TMI.2008.2004421.

[43] S. Ali *et al.*, "An objective comparison of detection and segmentation algorithms for artefacts in clinical endoscopy," *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–15, Feb. 2020, doi: 10.1038/s41598-020-59413-5.

[44] K. Sirinukunwattana, N. K. Alham, C. Verrill, and J. Rittscher, "Improving Whole Slide Segmentation Through Visual Context - A Systematic Study," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 2018, pp. 192–200, doi: 10.1007/978-3-030-00934-2_22.

[45] W. Xue, I. Ben Nachum, S. Pandey, J. Warrington, S. Leung, and S. Li, "Direct Estimation of Regional Wall Thicknesses via Residual Recurrent Neural Network," in *Information Processing in Medical Imaging - IPMI 2017*, 2017, pp. 505–516, doi: 10.1007/978-3-319-59050-9_40.

[46] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, vol. 2, pp. 1097–1105.

[47] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, "Fast and accurate view classification of echocardiograms using deep learning," *npj Digital Medicine*, vol. 1, no. 1, p. 6, Dec. 2018, doi: 10.1038/s41746-017-0013-1.

[48] S. Anker, F. W. Asselbergs, G. Brobert, P. Vardas, D. E. Grobbee, and M. Cronin, "Big Data in Cardiovascular Disease," *European Heart Journal*, vol. 38, no. 24, pp. 1863–1865, Jun. 2017, doi: 10.1093/eurheartj/ehx283.

[49] T. Zhou *et al.*, "M^2Net: Multi-modal Multi-channel Network for Overall Survival Time Prediction of Brain Tumor Patients," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 221–231, doi: 10.1007/978-3-030-59713-9_22.

[50] J. Yang, J. Chen, K. Kuang, T. Lin, J. He, and B. Ni, "MIA-Prognosis: A Deep Learning Framework to Predict Therapy Response," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 211–220, doi: 10.1007/978-3-030-59713-9_21.

[51] Y. H. Qu, H. T. Zhu, K. Cao, X. T. Li, M. Ye, and Y. S. Sun, "Prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer using a deep learning (DL) method," *Thoracic Cancer*, vol. 11, no. 3, pp. 651–658, Mar. 2020, doi: 10.1111/1759-7714.13309.

[52] G. Farrar *et al.*, "Atlas-based ventricular shape analysis for understanding congenital heart disease," *Progress in Pediatric Cardiology*, vol. 43, pp. 61–69, Dec. 2016, doi: 10.1016/J.PPEDCARD.2016.07.010.

[53] A. Lyon *et al.*, "Electrocardiogram phenotypes in hypertrophic cardiomyopathy caused by distinct mechanisms: Apico-basal repolarization gradients vs. Purkinje-myocardial coupling abnormalities," *Europace*, vol. 20, no. suppl_3, pp. III102–III112, Nov. 2018, doi: 10.1093/europace/euy226.

[54] S. J. Shah *et al.*, "Phenomapping for novel classification of heart failure with preserved ejection fraction," *Circ.*, 131, 269–79, 2015, doi: 10.1161/CIRCULATIONAHA.114.010637.

[55] M. Varela *et al.*, "Novel computational analysis of left atrial anatomy improves prediction of atrial fibrillation recurrence after ablation," *Frontiers in Physiology*, vol. 8, no. FEB, Feb. 2017, doi: 10.3389/fphys.2017.00068.

[56] A. Lyon *et al.*, "Distinct ECG phenotypes identified in hypertrophic cardiomyopathy using machine learning associate with arrhythmic risk markers," *Frontiers in Physiology*, vol. 9, no. MAR, 2018, doi: 10.3389/fphys.2018.00213.

[57] T. G. Dietterich, "Ensemble methods in machine learning," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2000, vol. 1857 LNCS, pp. 1–15, doi: 10.1007/3-540-45014-9_1.

[58] M. A. Ganaie, M. Hu, M. Tanveer*, and P. N. Suganthan*, "Ensemble deep learning: A review," *arXiv*, Apr. 2021, Available: http://arxiv.org/abs/2104.02395.

[59] J. S. Domingos, R. v. Stebbing, P. Leeson, and J. Alison Noble, "Structured random forests for myocardium delineation in 3d echocardiography," *Lecture Notes in Computer Science*, vol. 8679, pp. 215–222, 2014, doi: 10.1007/978-3-319-10581-9_27.

[60] K. Chykeyuk, D. A. Clifton, and J. A. Noble, “Feature extraction and wall motion classification of 2D stress echocardiography with relevance vector machines,” in *Proceedings - ISBI*, 2011, pp. 677–680, doi: 10.1109/ISBI.2011.5872497.

[61] Ian Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, vol. 13, no. 1. MIT Press, 2017.

[62] F. Rosenblatt, *The Perceptron, a Perceiving and Recognizing Automaton Project Para*. Cornell Aeronautical Laboratory, 1957.

[63] [C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation Functions: Comparison of trends in Practice and Research for Deep Learning,” *arXiv*, Nov. 2018, Available: http://arxiv.org/abs/1811.03378.

[64] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *arXiv*, Dec. 2015, Available: https://arxiv.org/abs/1412.6980v9.

[65] Y. Lecun, “A theoretical framework for back-propagation,” *Proceedings of the 1988 Connectionist Models Summer School, Pittsburg, PA*. Morgan Kaufmann, pp. 21–28, 1988.

[66] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020, doi: 10.1007/s10462-020-09825-6.

[67] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.

[68] A. J. Lewandowski *et al.*, “Preterm heart in adult life: Cardiovascular magnetic resonance reveals distinct differences in left ventricular mass, geometry, and function,” *Circulation*, vol. 127, no. 2, pp. 197–206, Jan. 2013, doi: 10.1161/CIRCULATIONAHA.112.126920.

[69] D. R. Warriner *et al., An Asymmetric Wall-Thickening Pattern Predicts Response to Cardiac Resynchronization Therapy*, vol. 11, no. 10. 2018, pp. 1545–1546.

[70] K. P. Vincent *et al.*, “Atlas-based methods for efficient characterization of patient-specific ventricular activation patterns,” *EP Europace*, vol. 23, no. Supplement_1, pp. i88–i95, Mar. 2021, doi: 10.1093/EUROPACE/EUAA397.

[71] K. Gilbert *et al.*, “Atlas-Based Computational Analysis of Heart Shape and Function in Congenital Heart Disease,” *Journal of Cardiovascular Translational Research* , vol. 11, no. 2, pp. 123–132, Jan. 2018, doi: 10.1007/S12265-017-9778-5.

[72] UK Biobank, “UK Biobank.” Available: https://www.ukbiobank.ac.uk/ (Jan. 30, 2020).

[73] Z Z. Hussain, F. Gimenez, D. Yi, and D. Rubin, “Differential Data Augmentation Techniques for Medical Imaging Classification Tasks.,” in *AMIA Symposium 2107*, vol. 2017, pp. 979–984, 2017, Available: www.ncbi.nlm.nih.gov/pubmed/29854165.

[74] J. Lemley, S. Bazrafkan, and P. Corcoran, “Smart Augmentation Learning an Optimal Data Augmentation Strategy,” *IEEE Access*, vol. 5, pp. 5858–5869, 2017, doi: 10.1109/ACCESS.2017.2696121.

[75] A. M. Molinaro, R. Simon, and R. M. Pfeiffer, “Prediction error estimation: A comparison of resampling methods,” *Bioinformatics*, vol. 21, no. 15, pp. 3301–3307, Aug. 2005, doi: 10.1093/bioinformatics/bti499.

[76] A. Santamaria-Pang, J. Kubricht, A. Chowdhury, C. Bhushan, and P. Tu, “Towards Emergent Language Symbolic Semantic Segmentation and Model Interpretability,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 326–334, doi: 10.1007/978-3-030-59710-8_32.

[77] W. Silva, A. Poellinger, J. S. Cardoso, and M. Reyes, “Interpretability-Guided Content-Based Medical Image Retrieval,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 305–314, doi: 10.1007/978-3-030-59710-8_30.

[78] D. Lenis, D. Major, M. Wimmer, A. Berg, G. Sluiter, and K. Bühler, “Domain Aware Medical Image Classifier Interpretation by Counterfactual Impact Analysis,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 315–325, doi: 10.1007/978-3-030-59710-8_31.

[79] R. LaLonde, D. Torigian, and U. Bagci, “Encoding Visual Attributes in Capsules for Explainable Medical Diagnoses,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 294–304, doi: 10.1007/978-3-030-59710-8_29.

[80] J. Sun, F. Darbehani, M. Zaidi, and B. Wang, “SAUNet: Shape Attentive U-Net for Interpretable Medical Image Segmentation,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 797–806, doi: 10.1007/978-3-030-59719-1_77.

[81] R. A. Gray and P. Pathmanathan, “Patient-Specific Cardiovascular Computational Modeling: Diversity of Personalization and Challenges,” *Journal of Cardiovascular Translational Research*, vol. 11, no. 2, pp. 80–88, Apr. 2018, doi: 10.1007/s12265-018-9792-2.

[82] M. R. Davies *et al.*, “Recent developments in using mechanistic cardiac modelling for drug safety evaluation,” *Drug Discovery Today*, vol. 21, no. 6, pp. 924–938, Jun. 2016, doi: 10.1016/J.DRUDIS.2016.02.003.

[83] N. A. Trayanova and J. J. Rice, “Cardiac Electromechanical Models: From Cell to Organ,” *Frontiers in Physiology*, vol. 2, p. 43, Aug. 2011, doi: 10.3389/fphys.2011.00043.

[84] X. Zhou, A. Bueno-Orovio, and B. Rodriguez, “In silico evaluation of arrhythmia,” *Current Opinion in Physiology*, vol. 1, pp. 95–103, Feb. 2018, doi: 10.1016/J.COPHYS.2017.11.003.

[85] C. E. Clancy and Y. Rudy, “Linking a genetic defect to its cellular phenotype in a cardiac arrhythmia,” *Nature*, vol. 400, no. 6744, pp. 566–569, Aug. 1999, doi: 10.1038/23034.

[86] K. Gilbert *et al.*, “Creating shape templates for patient specific biventricular modeling in congenital heart disease,” *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2015-November, pp. 679–682, Nov. 2015, doi: 10.1109/EMBC.2015.7318453.

[87] J. Walmsley, W. van Everdingen, M. J. Cramer, F. W. Prinzen, T. Delhaas, and J. Lumens, “Combining computer modelling and cardiac imaging to understand right ventricular pump function,” *Cardiovascular Research*, vol. 113, no. 12, pp. 1486–1498, Oct. 2017, doi: 10.1093/cvr/cvx154.

[88] A. Lopez-Perez, R. Sebastian, and J. M. Ferrero, “Three-dimensional cardiac computational modelling: methods, features and applications,” *BioMedical Engineering OnLine*, vol. 14, no. 1, p. 35, Dec. 2015, doi: 10.1186/s12938-015-0033-5.

[89] M. Reumann, V. Gurev, and J. J. Rice, “Computational modeling of cardiac disease: potential for personalized medicine,” *Personalized Medicine*, vol. 6, no. 1, pp. 45–66, Jan. 2009, doi: 10.2217/17410541.6.1.45.

[90] H. J. Arevalo *et al.*, “Arrhythmia risk stratification of patients after myocardial infarction using personalized heart models,” *Nature Communications*, vol. 7, no. 1, p. 11437, Dec. 2016, doi: 10.1038/ncomms11437.

[91] Z. J. Wang, V. Y. Wang, C. P. Bradley, M. P. Nash, A. A. Young, and J. J. Cao, “Left Ventricular Diastolic Myocardial Stiffness and End-Diastolic Myofibre Stress in Human Heart Failure Using Personalised Biomechanical Analysis,” *J. Cardiovasc. Transl. Res*, vol. 11, no. 4, pp. 346–356, Aug. 2018, doi: 10.1007/s12265-018-9816-y.

[92] F. Donati, C. A. Figueroa, N. P. Smith, P. Lamata, and D. A. Nordsletten, "Non-invasive pressure difference estimation from PC-MRI using the work-energy equation.," *Medical image analysis*, vol. 26, no. 1, pp. 159–72, Dec. 2015, doi: 10.1016/j.media.2015.08.012.

[93] B. L. Nørgaard *et al.*, "Diagnostic Performance of Noninvasive Fractional Flow Reserve Derived From Coronary Computed Tomography Angiography in Suspected Coronary Artery Disease: The NXT Trial (Analysis of Coronary Blood Flow Using CT Angiography: Next Steps)," *Journal of the American College of Cardiology*, vol. 63, no. 12, pp. 1145–1155, Apr. 2014, doi: 10.1016/J.JACC.2013.11.043.

[94] C. T. Villongco, D. E. Krummen, J. H. Omens, and A. D. McCulloch, "Non-invasive, model-based measures of ventricular electrical dyssynchrony for predicting CRT outcomes," *EP Europace*, vol. 18, no. suppl_4, pp. 104–112, Dec. 2016, doi: 10.1093/EUROPACE/EUW356.

[95] E. Kayvanpour *et al.*, "Towards Personalized Cardiology: Multi-Scale Modeling of the Failing Heart," *PLOS ONE*, vol. 10, no. 7, p. e0134869, 2015, doi: 10.1371/journal.pone.0134869.

[96] A. Krishnamurthy *et al.*, "Patient-specific models of cardiac biomechanics," *Journal of Computational Physics*, vol. 244, pp. 4–21, Jul. 2013, doi: 10.1016/J.JCP.2012.09.015.

[97] E. Passini *et al.*, "Mechanisms of pro-arrhythmic abnormalities in ventricular repolarisation and anti-arrhythmic therapies in human hypertrophic cardiomyopathy.," *Journal of molecular and cellular cardiology*, vol. 96, pp. 72–81, 2016, doi: 10.1016/j.yjmcc.2015.09.003.

[98] P. M. Boyle, S. Zahid, and N. A. Trayanova, "Using personalized computer models to custom-tailor ablation procedures for atrial fibrillation patients: are we there yet?," *Expert Review of Cardiovascular Therapy*, vol. 15, no. 5, pp. 339–341, May 2017, doi: 10.1080/14779072.2017.1317593.

[99] I. Fernández-Ruiz, "Computer modelling to personalize bioengineered heart valves," *Nature Reviews Cardiology*, vol. 15, 8, pp. 440–441, Aug. 2018, doi: 10.1038/s41569-018-0040-x.

[100] H. J. Arevalo *et al.*, "Arrhythmia risk stratification of patients after myocardial infarction using personalized heart models," *Nature Communications 2016 7:1*, vol. 7, no. 1, pp. 1–8, May 2016, doi: 10.1038/ncomms11437.

[101] H. Martinez-Navarro, A. Mincholé, A. Bueno-Orovio, and B. Rodriguez, "High arrhythmic risk in antero-septal acute myocardial ischemia is explained by increased transmural reentry occurrence," *Scientific Reports 2019 9:1*, vol. 9, no. 1, pp. 1–12, Nov. 2019, doi: 10.1038/s41598-019-53221-2.

[102] J. K. Min *et al.*, "Noninvasive Fractional Flow Reserve Derived From Coronary CT Angiography: Clinical Data and Scientific Principles," *JACC: Cardiovascular Imaging*, vol. 8, no. 10, pp. 1209–1222, Oct. 2015, doi: 10.1016/J.JCMG.2015.08.006.

[103] M. Haissaguerre *et al.*, "Noninvasive Panoramic Mapping of Human Atrial Fibrillation Mechanisms: A Feasibility Report," *Journal of Cardiovascular Electrophysiology*, vol. 24, no. 6, pp. 711–717, Jun. 2013, doi: 10.1111/jce.12075.

[104] T. M. Morrison, P. Pathmanathan, M. Adwan, and E. Margerrison, "Advancing Regulatory Science With Computational Modeling for Medical Devices at the FDA's Office of Science and Engineering Laboratories," *Frontiers in Medicine*, vol. 5, p. 241, Sep. 2018, doi: 10.3389/fmed.2018.00241.

[105] N. Zemzemi *et al.*, "Computational assessment of drug-induced effects on the electrocardiogram: from ion channel to body surface potentials.," *British journal of pharmacology*, vol. 168, no. 3, pp. 718–33, Feb. 2013, doi: 10.1111/j.1476-5381.2012.02200.x.

[106] I. Cavero and H. Holzgrefe, “CiPA: Ongoing testing, future qualification procedures, and pending issues,” *Journal of Pharmacological and Toxicological Methods*, vol. 76, pp. 27–37, Nov. 2015, doi: 10.1016/j.vascn.2015.06.004.

[107] E. Passini *et al.*, “Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity,” *Frontiers in Physiology*, vol. 8, p. 668, Sep. 2017, doi: 10.3389/fphys.2017.00668.

[108] C. H. Roney *et al.*, “Variability in pulmonary vein electrophysiology and fibrosis determines arrhythmia susceptibility and dynamics.,” *PLoS computational biology*, vol. 14, no. 5, p. e1006166, 2018, doi: 10.1371/journal.pcbi.1006166.

[109] Y. R. Hill *et al.*, “Investigating a Novel Activation-Repolarisation Time Metric to Predict Localised Vulnerability to Reentry Using Computational Modelling,” *PLOS ONE*, vol. 11, no. 3, p. e0149342, Mar. 2016, doi: 10.1371/journal.pone.0149342.

[110] P. D. Morris *et al.*, “Computational fluid dynamics modelling in cardiovascular medicine.,” *Heart (British Cardiac Society)*, vol. 102, no. 1, pp. 18–28, Jan. 2016, doi: 10.1136/heartjnl-2015-308044.

[111] Y. Feng *et al.*, “An efficient cardiac mapping strategy for radiofrequency catheter ablation with active learning,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 7, pp. 1199–1207, Jul. 2017, doi: 10.1007/s11548-017-1587-4.

[112] J. George *et al.*, “SLAC: Statistical total lesion metabolic activity computation by fuzzy unsupervised learning of PET images,” *Machine Vision and Applications*, vol. 24, no. 7, pp. 1341–1358, Oct. 2013, doi: 10.1007/s00138-012-0454-0.

[113] E. T. Y. Chang, M. Strong, and R. H. Clayton, “Bayesian Sensitivity Analysis of a Cardiac Cell Model Using a Gaussian Process Emulator,” *PLOS ONE*, vol. 10, no. 6, p. e0130252, Jun. 2015, doi: 10.1371/journal.pone.0130252.

[114] B. A. Lawson, K. Burrage, P. Burrage, C. C. Drovandi, and A. Bueno-Orovio, “Slow Recovery of Excitability Increases Ventricular Fibrillation Risk as Identified by Emulation,” *Frontiers in Physiology*, vol. 9, p. 1114, Aug. 2018, doi: 10.3389/fphys.2018.01114.

[115] R. H. Johnstone *et al.*, “Uncertainty and variability in models of the cardiac action potential: Can we build trustworthy models?,” *Journal of Molecular and Cellular Cardiology*, vol. 96, pp. 49–62, Jul. 2016, doi: 10.1016/J.YJMCC.2015.11.018.

[116] B. Lawson, C. Drovandi, P. Burrage, B. Rodriguez, and K. Burrage, “Dimension Reduction for the Emulation of Cardiac Electrophysiology Models for Single Cells and Tissue,” Sep. 2017, doi: 10.22489/CinC.2017.309-340.

[117] E. A. Sobie, “Parameter Sensitivity Analysis in Electrophysiological Models Using Multivariable Regression,” *Biophysical Journal*, vol. 96, no. 4, pp. 1264–1274, Feb. 2009, doi: 10.1016/j.bpj.2008.10.056.

[118] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, “Design and Analysis of Computer Experiments,” *Statistical Science*, vol. 4, no. 4, pp. 409–423, Nov. 1989, doi: 10.1214/ss/1177012413.

[119] C. Edward. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. MIT Press, 2006.

[120] W. E. Sorteberg, S. Garasto, A. S. Pouplin, C. D. Cantwell, and A. Bharath, “Approximating the solution to wave propagation using deep neural networks,” *arXiv*, Dec. 2018, Available: https://arxiv.org/abs/1812.01609.

[121] J. Verrelst *et al.*, "SCOPE-Based Emulators for Fast Generation of Synthetic Canopy Reflectance and Sun-Induced Fluorescence Spectra," *Remote Sensing*, vol. 9, no. 9, p. 927, Sep. 2017, doi: 10.3390/rs9090927.

[122] F. Meister *et al.*, "Graph Convolutional Regression of Cardiac Depolarization from Sparse Endocardial Maps," in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, 2021, pp. 23–34, doi: 10.1007/978-3-030-68107-4_3.

[123] X. Jiang, S. Ghimire, J. Dhamala, Z. Li, P. K. Gyawali, and L. Wang, "Learning Geometry-Dependent and Physics-Based Inverse Image Reconstruction," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 487–496, doi: 10.1007/978-3-030-59725-2_47.

[124] P. D. Dueben and P. Bauer, "Challenges and design choices for global weather and climate models based on machine learning," *Geoscientific Model Development*, vol. 11, no. 10, pp. 3999–4009, Oct. 2018, doi: 10.5194/gmd-11-3999-2018.

[125] R. Prudden, A. Arribas, J. Tomlinson, and N. Robinson, "Using Machine Learning as a fast emulator of physical processes within the Met Office's Unified Model," *AGU Fall Meeting Abstracts*, Dec. 2017.

[126] J. Rivera *et al.*, "An Emulator Toolbox to Approximate Radiative Transfer Models with Statistical Learning," *Remote Sensing*, vol. 7, no. 7, pp. 9347–9370, Jul. 2015, doi: 10.3390/rs70709347.

[127] A. Toni Albert, A. Rhoades, S. Ganguly, D. Feldman, A. D. Jones, and Prabhat, "Towards Generative Deep Learning Emulators for Fast Hydroclimate Simulation," *U.S. Department of Energy*, 2018. Available: https://climatemodeling.science.energy.gov/presentations/towards-generative-deep-learning-emulators-fast-hydroclimate-simulations.

[128] Z. J. Wang *et al.*, "Efficient estimation of load-free left ventricular geometry and passive myocardial properties using principal component analysis," *Int J Numer Method Biomed Eng*. 2020 Mar;36(3):e3313. doi: 10.1002/cnm.3313.

[129] J. R. Levick, *An introduction to cardiovascular physiology*. London; New York: Arnold ; Distributed in the United States of America by Oxford University Press, 2003.

[130] Wikipedia.org, "Heart." Available: en.wikipedia.org/wiki/Heart (accessed Jul. 08, 2021).

[131] R. E. Klabunde, *Cardiovascular Physiology Concepts*, 3rd ed. Wolters Kluwer, 2021.

[132] Khurana, *Essentials of Medical Physiology*. Elsevier India, 2008.

[133] Wikipedia.org, "Electrocardiography" Available: en.wikipedia.org/wiki/Electrocardiography (accessed Jul. 18, 2021)

[134] L. S. Lilly and Harvard Medical School., *Pathophysiology of heart disease: a collaborative project of medical students and faculty*. 5th Edition. Lippincott Williams & Wilkins, 2012.

[135] H. Watkins, H. Ashrafian, and C. Redwood, "Inherited Cardiomyopathies," *New England Journal of Medicine*, vol. 364, no. 17, pp. 1643–1656, Apr. 2011, doi: 10.1056/NEJMra0902923.

[136] B. J. Maron and M. S. Maron, "Hypertrophic cardiomyopathy," *The Lancet*, vol. 381, no. 9862, pp. 242–255, Jan. 2013, doi: 10.1016/S0140-6736(12)60397-3.

[137] R. Coppini *et al.*, "Late Sodium Current Inhibition Reverses Electromechanical Dysfunction in Human Hypertrophic Cardiomyopathy," *Circulation*, vol. 127, no. 5, pp. 575–584, Feb. 2013, doi: 10.1161/CIRCULATIONAHA.112.134932.

[138] B. A. Houston and G. R. Stevens, “Hypertrophic cardiomyopathy: a review.,” *Clinical Medicine Insights. Cardiology*, vol. 8, no. Suppl 1, pp. 53–65, 2014, doi: 10.4137/CMC.S15717.

[139] C. Semsarian, J. Ingles, M. S. Maron, and B. J. Maron, “New Perspectives on the Prevalence of Hypertrophic Cardiomyopathy,” *Journal of the American College of Cardiology*, vol. 65, no. 12, pp. 1249–1254, Mar. 2015, doi: 10.1016/j.jacc.2015.01.019.

[140] Mayo Foundation for Medical Education and Research, “Mayo Foundation for Medical Education and Research.” Available: https://www.mayo.edu/ (accessed Jul. 08, 2021).

[141] Wikipedia, “Myocardial infarction.” Available: en.wikipedia.org/wiki/Myocardial_infarction (accessed Jul. 08, 2021).

[142] R. A. Noureldin *et al.*, “The diagnosis of hypertrophic cardiomyopathy by cardiovascular magnetic resonance.,” *Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance*, vol. 14, no. 1, p. 17, Feb. 2012, doi: 10.1186/1532-429X-14-17.

[143] C. O’Mahony, P. Elliott, and W. McKenna, “Sudden Cardiac Death in Hypertrophic Cardiomyopathy,” *Circulation: Arrhythmia and Electrophysiology*, vol. 6, no. 2, pp. 443–451, Apr. 2013, doi: 10.1161/CIRCEP.111.962043.

[144] QxMD, “HCM Risk-SCD calculator.” Available: qxmd.com/calculate/calculator_303/hcm-risk-scd (accessed Jul. 09, 2021).

[145] Authors/Task Force members *et al.*, “2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy,” *European Heart Journal*, vol. 35, no. 39, pp. 2733–2779, Oct. 2014, doi: 10.1093/eurheartj/ehu284.

[146] B. J. Maron, S. A. Casey, R. H. Chan, R. F. Garberich, E. J. Rowin, and M. S. Maron, “Independent Assessment of the European Society of Cardiology Sudden Death Risk Model for Hypertrophic Cardiomyopathy,” *The American Journal of Cardiology*, vol. 116, no. 5, pp. 757–764, Sep. 2015, doi: 10.1016/j.amjcard.2015.05.047.

[147] K. Thygesen *et al.*, “Third universal definition of myocardial infarction,” *Nature Reviews Cardiology*, vol. 9, no. 11, pp. 620–633, Nov. 2012, doi: 10.1038/nrcardio.2012.122.

[148] Ph. G. Steg *et al.*, “ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation,” *European Heart Journal*, vol. 33, no. 20, pp. 2569–2619, Oct. 2012, doi: 10.1093/eurheartj/ehs215.

[149] M. G. John Sutton and N Sharpe, “Left Ventricular Remodeling After Myocardial Infarction,” *Circulation*, vol. 101, no. 25, pp. 2981–2988, Jun. 2000, doi: 10.1161/01.CIR.101.25.2981.

[150] J. W. Holmes, T. K. Borg, and J. W. Covell, “Structure and Mechanics of Healing Myocardial Infarcts,” *Annual review of biomedical engineering*, vol. 7, pp. 223–253, Jul. 2005, doi: 10.1146/ANNUREV.BIOENG.7.060804.100453.

[151] C. Bauters *et al.*, “Long-term prognostic impact of left ventricular remodeling after a first myocardial infarction in modern clinical practice,” *PLOS ONE*, vol. 12, no. 11, p. e0188884, Nov. 2017, doi: 10.1371/journal.pone.0188884.

[152] S. Desch *et al.*, “Cardiac magnetic resonance imaging parameters as surrogate endpoints in clinical trials of acute myocardial infarction.,” *Trials*, vol. 12, p. 204, Sep. 2011, doi: 10.1186/1745-6215-12-204.

[153] R. F. Rushmer, D. L. Franklin, and R. M. Ellis, “Left Ventricular Dimensions Recorded by Sonocardiometry,” *Circulation Research*, vol. 4, no. 6, pp. 684–688, Nov. 1956, doi: 10.1161/01.RES.4.6.684.

[154] H. N. Sabbah, M. Marzilli, and P. D. Stein, "The relative role of subendocardium and subepicardium in left ventricular mechanics," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 240, no. 6, pp. H920–H926, Jun. 1981, doi: 10.1152/ajpheart.1981.240.6.H920.

[155] G. M. Pohost, "The history of cardiovascular magnetic resonance.," *JACC. Cardiovascular imaging*, vol. 1, no. 5, pp. 672–8, Sep. 2008, doi: 10.1016/j.jcmg.2008.07.009.

[156] P. Peng, K. Lekadir, A. Gooya, L. Shao, S. E. Petersen, and A. F. Frangi, "A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging.," *Magma (New York, N.Y.)*, vol. 29, no. 2, pp. 155–95, Apr. 2016, doi: 10.1007/s10334-015-0521-4.

[157] T. Leiner *et al.*, "SCMR Position Paper (2020) on clinical indications for cardiovascular magnetic resonance," *Journal of Cardiovascular Magnetic Resonance*, vol. 22, no. 1. BioMed Central Ltd, Dec. 01, 2020, doi: 10.1186/s12968-020-00682-4.

[158] E. Dall'Armellina, T. D. Karamitsos, S. Neubauer, and R. P. Choudhury, "CMR for characterization of the myocardium in acute coronary syndromes," *Nature Reviews Cardiology*, vol. 7, no. 11, pp. 624–636, Nov. 2010, doi: 10.1038/nrcardio.2010.140.

[159] T. Podlesnikar, V. Delgado, and J. J. Bax, "Cardiovascular magnetic resonance imaging to assess myocardial fibrosis in valvular heart disease," *The International Journal of Cardiovascular Imaging*, vol. 34, no. 1, pp. 97–112, Jan. 2018, doi: 10.1007/s10554-017-1195-y.

[160] M.-T. Wu *et al.*, "Sequential Changes of Myocardial Microstructure in Patients Postmyocardial Infarction by Diffusion-Tensor Cardiac MR," *Circulation: Cardiovascular Imaging*, vol. 2, no. 1, pp. 32–40, Jan. 2009, doi: 10.1161/CIRCIMAGING.108.778902.

[161] A. J. Taylor, M. Salerno, R. Dharmakumar, and M. Jerosch-Herold, "T1 Mapping: Basic Techniques and Clinical Applications," *JACC: Cardiovascular Imaging*, vol. 9, no. 1, pp. 67–81, Jan. 2016, doi: 10.1016/J.JCMG.2015.11.005.

[162] E.-S. H. Ibrahim, "Myocardial tagging by cardiovascular magnetic resonance: evolution of techniques--pulse sequences, analysis algorithms, and applications.," *Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance*, vol. 13, no. 1, p. 36, Jul. 2011, doi: 10.1186/1532-429X-13-36.

[163] A. Scatteia, A. Baritussio, and C. Bucciarelli-Ducci, "Strain imaging using cardiac magnetic resonance.," *Heart failure reviews*, vol. 22, no. 4, pp. 465–476, 2017, doi: 10.1007/s10741-017-9621-8.

[164] M. D. Cerqueira *et al.*, "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association.," *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology*, vol. 9, no. 2, pp. 240–5, 2002, doi: 10.1067/MNC.2002.123122.

[165] P. Bucius *et al.*, "Comparison of feature tracking, fast-SENC, and myocardial tagging for global and segmental left ventricular strain," *ESC Heart Failure*, vol. 7, no. 2, pp. 523–532, Apr. 2020, doi: 10.1002/ehf2.12576.

[166] A. Scatteia, A. Baritussio, and C. Bucciarelli-Ducci, "Strain imaging using cardiac magnetic resonance," *Heart Failure Reviews*, vol. 22, no. 4. Springer New York LLC, pp. 465–476, Jul. 01, 2017, doi: 10.1007/s10741-017-9621-8.

[167] National Institute of Biomedical Imaging and Bioengineering, “Magnetic Resonance Imaging (MRI).” https://www.nibib.nih.gov/science-education/science-topics/magnetic-resonance-imaging-mri (accessed Jun. 02, 2021).

[168] A. D. Scott, J. Keegan, and D. N. Firmin, “Motion in cardiovascular MR imaging,” *Radiology*, vol. 250, no. 2, pp. 331–351, Feb. 2009, doi: 10.1148/radiol.2502071998.

[169] J. Marshall, T. Martin, J. Downie, and K. Malisza, “A comprehensive analysis of MRI research risks: In support of full disclosure,” *Canadian Journal of Neurological Sciences*, vol. 34, no. 1. Canadian Journal of Neurological Sciences, pp. 11–17, 2007, doi: 10.1017/S0317167100005734.

[170] T. Force *et al.*, “Research priorities in hypertrophic cardiomyopathy: report of a Working Group of the National Heart, Lung, and Blood Institute.,” *Circulation*, vol. 122, no. 11, pp. 1130–3, Sep. 2010, doi: 10.1161/CIRCULATIONAHA.110.950089.

[171] D. P. Ripley, T. A. Musa, L. E. Dobson, S. Plein, and J. P. Greenwood, “Cardiovascular magnetic resonance imaging: what the general cardiologist should know.,” *Heart (British Cardiac Society)*, vol. 102, no. 19, pp. 1589–603, 2016, doi: 10.1136/heartjnl-2015-307896.

[172] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, “Magnetic resonance imaging : physical principles and sequence design.” Wiley-Blackwell, 2014.

[173] J. P. Ridgway, “Cardiovascular magnetic resonance physics for clinicians: Part i,” *Journal of Cardiovascular Magnetic Resonance*, vol. 12, no. 1. BioMed Central, pp. 1–28, Nov. 30, 2010, doi: 10.1186/1532-429X-12-71.

[174] M. Saeed, T. A. Van, R. Krug, S. W. Hetts, and M. W. Wilson, “Cardiac MR imaging: current status and future direction.,” *Cardiovascular diagnosis and therapy*, vol. 5, no. 4, pp. 290–310, Aug. 2015, doi: 10.3978/j.issn.2223-3652.2015.06.07.

[175] A. A. Goshtasby and D. A. Turner, “Fusion of short-axis and long-axis cardiac MR images,” *Computerized Medical Imaging and Graphics*, vol. 20, no. 2, pp. 77–87, 1996, doi: 10.1016/0895-6111(96)00035-3.

[176] A. G. Chandler *et al.*, “Correction of misaligned slices in multi-slice cardiovascular magnetic resonance using slice-to-volume registration,” *Journal of Cardiovascular Magnetic Resonance*, vol. 10, no. 1, pp. 1–9, Feb. 2008, doi: 10.1186/1532-429X-10-13.

[177] Y. Su *et al.*, “Automatic correction of motion artifacts in 4D left ventricle model reconstructed from MRI,” in *Computing in Cardiology 2014*, 2014, pp. 705–708.

[178] E. Ferrante and N. Paragios, “Slice-to-volume medical image registration: A survey,” *Medical Image Analysis*, vol. 39, pp. 101–123, Jul. 2017, doi: 10.1016/j.media.2017.04.010.

[179] G. Shechter, C. Ozturk, J. R. Resar, and E. R. McVeigh, “Respiratory motion of the heart from free breathing coronary angiograms,” in *IEEE Transactions on Medical Imaging*, Aug. 2004, vol. 23, no. 8, pp. 1046–1056, doi: 10.1109/TMI.2004.828676.

[180] K. McLeish, D. L. G. Hill, D. Atkinson, J. M. Blackall, and R. Razavi, “A study of the motion and deformation of the heart due to respiration,” *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1142–1150, Sep. 2002, doi: 10.1109/TMI.2002.804427.

[181] C. Chen, Y. Liu, O. P. Simonetti, and R. Ahmad, “Automatic Extraction and Sign Determination of Respiratory Signal in Real-Time Cardiac Magnetic Resonance Imaging,” in *Proceedings - International Symposium on Biomedical Imaging*, Apr. 2020, vol. 2020-April, pp. 830–833, doi: 10.1109/ISBI45749.2020.9098315.

[182] S. E. Petersen *et al.*, "Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort," *Journal of Cardiovascular Magnetic Resonance*, vol. 19, no. 1, p. 18, Dec. 2017, doi: 10.1186/s12968-017-0327-9.

[183] M. S. Nacif *et al.*, "Left atrial volume quantification using cardiac MRI in atrial fibrillation: comparison of the Simpson's method with biplane area-length, ellipse, and three-dimensional methods," *Diagnostic and Interventional Radiology*, Dec. 2012, doi: 10.5152/dir.2012.002.

[184] R. Lang *et al.*, "Recommendations for chamber quantification☆," *European Journal of Echocardiography*, vol. 7, no. 2, pp. 79–108, Mar. 2006, doi: 10.1016/j.euje.2005.12.014.

[185] A. Suinesiaputra *et al.*, "Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours," *Journal of Cardiovascular Magnetic Resonance*, vol. 17, no. 1, p. 63, Dec. 2015, doi: 10.1186/s12968-015-0170-9.

[186] A. F. Frangi, W. J. Niessen, and M. A. Viergever, "Three-dimensional modeling for functional analysis of cardiac images, a review," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 2–5, 2001, doi: 10.1109/42.906421.

[187] M. L. Chuang *et al.*, "CMR reference values for left ventricular volumes, mass, and ejection fraction using computer-aided analysis: the Framingham Heart Study.," *Journal of magnetic resonance imaging : JMRI*, vol. 39, no. 4, pp. 895–900, Apr. 2014, doi: 10.1002/jmri.24239.

[188] J. Bogaert, S. Dymarkowski, A. M. Taylor, and V. Muthurangu, *Clinical Cardiac MRI*. Springer Berlin Heidelberg, 2012.

[189] Z. B. Popović and J. D. Thomas, "Assessing observer variability: a user's guide.," *Cardiovascular diagnosis and therapy*, vol. 7, no. 3, pp. 317–324, Jun. 2017, doi: 10.21037/cdt.2017.03.12.

[190] T. Leiner *et al.*, "Machine learning in cardiovascular magnetic resonance: Basic concepts and applications," *Journal of Cardiovascular Magnetic Resonance*, vol. 21, no. 1. BioMed Central Ltd., Oct. 07, 2019, doi: 10.1186/s12968-019-0575-y.

[191] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543–563, Aug. 2009, doi: 10.1016/j.media.2009.05.004.

[192] S. Ardekani *et al.*, "Shape analysis of hypertrophic and hypertensive heart disease using MRI-based 3D surface models of left ventricular geometry," *Medical Image Analysis*, vol. 29, pp. 12–23, Apr. 2016, doi: 10.1016/j.media.2015.11.004.

[193] C. Mauger *et al.*, "Right ventricular shape and function: Cardiovascular magnetic resonance reference morphology and biventricular risk factor morphometrics in UK Biobank," *Journal of Cardiovascular Magnetic Resonance*, vol. 21, no. 1, Jul. 2019, doi: 10.1186/s12968-019-0551-6.

[194] I. Eitel *et al.*, "Comprehensive prognosis assessment by CMR imaging after ST-segment elevation myocardial infarction," *Journal of the American College of Cardiology*, vol. 64, no. 12, pp. 1217–1226, Sep. 2014, doi: 10.1016/j.jacc.2014.06.1194.

[195] N. Dagres and G. Hindricks, "Risk stratification after myocardial infarction: is left ventricular ejection fraction enough to prevent sudden cardiac death?," *European Heart Journal*, vol. 34, no. 26, pp. 1964–1971, May 2013, doi: 10.1093/eurheartj/eht109.

[196] J. L. Rouleau *et al.*, "Myocardial infarction patients in the 1990s - Their risk factors, stratification and survival in Canada: The Canadian Assessment of Myocardial Infarction (CAMI) study," *Journal of the American College of Cardiology*, vol. 27, no. 5, pp. 1119–1127, 1996, doi: 10.1016/0735-1097(95)00599-4.

[197] H. D. White, R. M. Norris, M. A. Brown, P. W. Brandt, R. M. Whitlock, and C. J. Wild, "Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction.," *Circulation*, vol. 76, no. 1, pp. 44–51, Jul. 1987, doi: 10.1161/01.CIR.76.1.44.

[198] J. N. Cohn, R. Ferrari, and N. Sharpe, "Cardiac remodeling-concepts and clinical implications: A consensus paper from an International Forum on Cardiac Remodeling," *Journal of the American College of Cardiology*, vol. 35, no. 3, pp. 569–582, Mar. 2000, doi: 10.1016/S0735-1097(99)00630-0.

[199] H. Bulluck, R. Dharmakumar, A. E. Arai, C. Berry, and D. J. Hausenloy, "Cardiovascular Magnetic Resonance in Acute ST-Segment–Elevation Myocardial Infarction," *Circulation*, vol. 137, no. 18, pp. 1949–1964, May 2018, doi: 10.1161/CIRCULATIONAHA.117.030693.

[200] M. Reindl, I. Eitel, and S. J. Reinstadler, "Role of Cardiac Magnetic Resonance to Improve Risk Prediction following Acute ST-elevation Myocardial Infarction," *Journal of Clinical Medicine*, vol. 9, no. 4, p. 1041, Apr. 2020, doi: 10.3390/jcm9041041.

[201] P. G. Masci *et al.*, "Relationship between location and size of myocardial infarction and their reciprocal influences on post-infarction left ventricular remodelling," *European Heart Journal*, vol. 32, no. 13, pp. 1640–1648, Jul. 2011, doi: 10.1093/eurheartj/ehr064.

[202] J. T. Ortiz-Pérez, J. Rodríguez, S. N. Meyers, D. C. Lee, C. Davidson, and E. Wu, "Correspondence Between the 17-Segment Model and Coronary Arterial Anatomy Using Contrast-Enhanced Cardiac Magnetic Resonance Imaging," *JACC: Cardiovascular Imaging*, vol. 1, no. 3, pp. 282–293, May 2008, doi: 10.1016/j.jcmg.2008.01.014.

[203] M. S. Amzulescu *et al.*, "Myocardial strain imaging: review of general principles, validation, and sources of discrepancies," *European Heart Journal - Cardiovascular Imaging*, vol. 20, no. 6, pp. 605–619, Mar. 2019, doi: 10.1093/ehjci/jez041.

[204] K. Mangion, C. McComb, D. A. Auger, F. H. Epstein, and C. Berry, "Magnetic resonance imaging of myocardial strain after acute st-segment-elevation myocardial infarction a systematic review," *Circulation: Cardiovascular Imaging*, vol. 10, no. 8. Lippincott Williams and Wilkins, Aug. 2017, doi: 10.1161/CIRCIMAGING.117.006498.

[205] T. Stiermaier *et al.*, "Cardiac Magnetic Resonance Left Ventricular Mechanical Uniformity Alterations for Risk Assessment After Acute Myocardial Infarction," *Journal of the American Heart Association*, vol. 8, no. 16, Aug. 2019, doi: 10.1161/JAHA.118.011576.

[206] I. Eitel *et al.*, "Cardiac Magnetic Resonance Myocardial Feature Tracking for Optimized Prediction of Cardiovascular Events Following Myocardial Infarction," *JACC: Cardiovascular Imaging*, vol. 11, no. 10, pp. 1433–1444, Oct. 2018, doi: 10.1016/j.jcmg.2017.11.034.

[207] T. Stiermaier *et al.*, "Optimized Prognosis Assessment in ST-Segment-Elevation Myocardial Infarction Using a Cardiac Magnetic Resonance Imaging Risk Score," *Circulation. Cardiovasc. imaging*, vol. 10, no. 11, Nov. 2017, doi: 10.1161/CIRCIMAGING.117.006774.

[208] A. Schuster *et al.*, "Fully Automated Cardiac Assessment for Diagnostic and Prognostic Stratification Following Myocardial Infarction," *Journal of the American Heart Association*, vol. 9, no. 18, p. e016612, Sep. 2020, doi: 10.1161/JAHA.120.016612.

[209] S. M. Cho *et al.*, "Machine Learning Compared With Conventional Statistical Models for Predicting Myocardial Infarction Readmission and Mortality: A Systematic Review," *Canadian Journal of Cardiology*, Mar. 2021, doi: 10.1016/J.CJCA.2021.02.020.

[210] B. Patel and P. Sengupta, “Machine learning for predicting cardiac events: what does the future hold?,” *Expert review of cardiovascular therapy*, vol. 18, no. 2, p. 77, Feb. 2020, doi: 10.1080/14779072.2020.1732208.

[211] D. Mandair, P. Tiwari, S. Simon, K. L. Colborn, and M. A. Rosenberg, “Prediction of incident myocardial infarction using machine learning applied to harmonized electronic health record data,” *BMC Medical Informatics and Decision Making 2020 20:1*, vol. 20, no. 1, pp. 1–10, Oct. 2020, doi: 10.1186/S12911-020-01268-X.

[212] J. Kwon *et al.*, “Deep-learning-based risk stratification for mortality of patients with acute myocardial infarction,” *PLOS ONE*, vol. 14, no. 10, p. e0224502, Oct. 2019, doi: 10.1371/JOURNAL.PONE.0224502.

[213] Y. Li, L. Jiang, J. He, K. Jia, Y. Peng, and M. Chen, “Machine Learning to Predict the 1-Year Mortality Rate After Acute Anterior Myocardial Infarction in Chinese Patients,” *Therapeutics and Clinical Risk Management*, vol. 16, p. 1, 2020, doi: 10.2147/TCRM.S236498.

[214] R. Khera *et al.*, “Use of Machine Learning Models to Predict Death After Acute Myocardial Infarction,” *JAMA Cardiology*, vol. 6, no. 6, pp. 633–641, Jun. 2021, doi: 10.1001/JAMACARDIO.2021.0122.

[215] M. P. Than *et al.*, “Machine Learning to Predict the Likelihood of Acute Myocardial Infarction,” *Circulation*, vol. 140, no. 11, pp. 899–909, Sep. 2019, doi: 10.1161/CIRCULATIONAHA.119.041980.

[216] C. C. Wu *et al.*, “An artificial intelligence approach to early predict non-ST-elevation myocardial infarction patients with chest pain,” *Computer Methods and Programs in Biomedicine*, vol. 173, pp. 109–117, May 2019, doi: 10.1016/J.CMPB.2019.01.013.

[217] Y. Cho *et al.*, “Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography,” *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–10, Nov. 2020, doi: 10.1038/s41598-020-77599-6.

[218] Hae-Yeoun Lee, N. C. F. Codella, M. D. Cham, J. W. Weinsaft, and Yi Wang, “Automatic Left Ventricle Segmentation Using Iterative Thresholding and an Active Contour Model With Adaptation on Short-Axis Cardiac MRI,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 905–913, Apr. 2010, doi: 10.1109/TBME.2009.2014545.

[219] Y.-L. Lu, K. A. Connelly, A. J. Dick, G. A. Wright, and P. E. Radau, “Automatic functional analysis of left ventricle in cardiac cine MRI.,” *Quantitative imaging in medicine and surgery*, vol. 3, no. 4, pp. 200–9, Aug. 2013, doi: 10.3978/j.issn.2223-4292.2013.08.02.

[220] J. Ringenberg *et al.*, “Automated segmentation and reconstruction of patient-specific cardiac anatomy and pathology from *in vivo* MRI*,” *Measurement Science and Technology*, vol. 23, no. 12, p. 125405, Dec. 2012, doi: 10.1088/0957-0233/23/12/125405.

[221] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, Jan. 1988, doi: 10.1007/BF00133570.

[222] L. M. Koch *et al.*, “Multi-Atlas Segmentation Using Partially Annotated Data: Methods and Annotation Strategies,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 7, pp. 1683–1696, Jul. 2018, doi: 10.1109/TPAMI.2017.2711020.

[223] X. Alba *et al.*, “An Algorithm for the Segmentation of Highly Abnormal Hearts Using a Generic Statistical Shape Model,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 3, pp. 845–859, Mar. 2016, doi: 10.1109/TMI.2015.2497906.

[224] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, “Active Shape Models-Their Training and Application,” *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995, doi: 10.1006/CVIU.1995.1004.

[225] Y. Zhu, X. Papademetris, A. J. Sinusas, and J. S. Duncan, "Bidirectional segmentation of three-dimensional cardiac MR images using a subject-specific dynamical model.," *Medical image computing and computer-assisted intervention - MICCAI 2008*, vol. 11, no. Pt 2, pp. 450–7, 2008, doi: 10.1007/978-3-540-85990-1_54.

[226] L. E. Juarez-Orozco, O. Martinez-Manzanera, S. V. Nesterov, S. Kajander, and J. Knuuti, "The machine learning horizon in cardiac hybrid imaging," *European Journal of Hybrid Imaging*, vol. 2, no. 1, p. 15, Dec. 2018, doi: 10.1186/s41824-018-0033-3.

[227] G. Stalidis, N. Maglaveras, S. N. Efstratiadis, A. S. Dimitriadis, and C. Pappas, "Model-based processing scheme for quantitative 4-D cardiac MRI analysis.," *IEEE transactions on information technology in biomedicine*, vol. 6, no. 1, pp. 59–72, Mar. 2002, doi: 10.1109/4233.992164.

[228] W. Bai, W. Shi, C. Ledig, and D. Rueckert, "Multi-atlas segmentation with augmented features for cardiac MR images," *Medical Image Analysis*, vol. 19, no. 1, pp. 98–109, Jan. 2015, doi: 10.1016/J.MEDIA.2014.09.005.

[229] D. Mahapatra, "Analyzing Training Information From Random Forests for Improved Image Segmentation," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1504–1512, Apr. 2014, doi: 10.1109/TIP.2014.2305073.

[230] H. Hu, H. Liu, Z. Gao, and L. Huang, "Hybrid segmentation of left ventricle in cardiac MRI using gaussian-mixture model and region restricted dynamic programming," *Magnetic Resonance Imaging*, vol. 31, no. 4, pp. 575–584, May 2013, doi: 10.1016/J.MRI.2012.10.004.

[231] X. Zhen, Z. Wang, A. Islam, M. Bhaduri, I. Chan, and S. Li, "Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation," *Medical Image Analysis*, vol. 30, pp. 120–129, May 2016, doi: 10.1016/J.MEDIA.2015.07.003.

[232] Z. Zhijie Wang, M. Ben Salah, B. Bin Gu, A. Islam, A. Goela, and S. Shuo Li, "Direct Estimation of Cardiac Biventricular Volumes With an Adapted Bayesian Formulation," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1251–1260, Apr. 2014, doi: 10.1109/TBME.2014.2299433.

[233] C. Chen *et al.*, "Deep Learning for Cardiac Image Segmentation: A Review," *Frontiers in Cardiovascular Medicine*, vol. 7, p. 25, Mar. 2020, doi: 10.3389/fcvm.2020.00025.

[234] P. V. Tran, "A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis MRI," *arXiv*, Apr. 2016. Available: https://arxiv.org/abs/1604.00494.

[235] D. M. Vigneault, W. Xie, C. Y. Ho, D. A. Bluemke, and J. A. Noble, "Ω-Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks," *Medical Image Analysis*, vol. 48, pp. 95–106, Nov. 2018, doi: 10.1016/j.media.2018.05.008.

[236] O. Oktay *et al.*, "Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 384–395, Feb. 2018, doi: 10.1109/TMI.2017.2743464.

[237] T. A. Ngo, Z. Lu, and G. Carneiro, "Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance," *Medical Image Analysis*, vol. 35, pp. 159–171, Jan. 2017, doi: 10.1016/J.MEDIA.2016.05.009.

[238] A. Newell, K. Yang, and J. Deng, "Stacked Hourglass Networks for Human Pose Estimation," *Computer Vision – ECCV 2016*, vol 9912, Springer, Cham, 2016, pp. 483–499, doi: 10.1007/978-3-319-46484-8_29.

[239] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," *Deep Learning in Medical Image Analysis*

and Multimodal Learning for Clinical Decision Support. DLMIA 2018, vol. 11045 LNCS, pp. 3–11, Jul. 2018, doi: 10.1007/978-3-030-00889-5_1.

[240] J. M. J. Valanarasu, V. A. Sindagi, I. Hacihaliloglu, and V. M. Patel, "KiU-Net: Towards Accurate Segmentation of Biomedical Images Using Over-Complete Representations," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 363–373, doi: 10.1007/978-3-030-59719-1_36.

[241] L. Jiao, L. Huo, C. Hu, and P. Tang, "Refined UNet: UNet-Based Refinement Network for Cloud and Shadow Precise Segmentation," *Remote Sensing*, vol. 12, no. 12, 2020, doi: 10.3390/rs12122001.

[242] C. F. Baumgartner *et al.*, "PHiSeg: Capturing Uncertainty in Medical Image Segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, 2019, pp. 119–127, doi: 10.1007/978-3-030-32245-8_14.

[243] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas." a*rXiv*, 2018. Available: https://arxiv.org/abs/1804.03999.

[244] F. Kong and S. C. Shadden, "A Generalizable Deep-Learning Approach for Cardiac Magnetic Resonance Image Segmentation Using Image Augmentation and Attention U-Net," in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, 2020, pp. 287–296, doi: 10.1007/978-3-030-68107-4_29.

[245] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," a*rXiv*, Sep. 2014. Available: https://arxiv.org/abs/1409.1556.

[246] A. J. Larrazabal, C. Martinez, and E. Ferrante, "Anatomical Priors for Image Segmentation via Post-processing with Denoising Autoencoders," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, pp. 585–593, doi: 10.1007/978-3-030-32226-7_65.

[247] Q. Zeng *et al.*, "Liver Segmentation in Magnetic Resonance Imaging via Mean Shape Fitting with Fully Convolutional Neural Networks," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, pp. 246–254, doi: 10.1007/978-3-030-32245-8_28.

[248] D. H. Ballard, "Modular Learning in Neural Networks," in *Proceedings of the Sixth National Conference on Artificial Intelligence - Vol 1*, 1987, pp. 279–284, ISBN: 978-0-934613-42-2.

[249] N. Painchaud, Y. Skandarani, T. Judge, O. Bernard, A. Lalande, and P.-M. Jodoin, "Cardiac MRI Segmentation with Strong Anatomical Guarantees," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, 2019, pp. 632–640, doi: 10.1007/978-3-030-32245-8_70.

[250] Q. Yue, X. Luo, Q. Ye, L. Xu, and X. Zhuang, "Cardiac Segmentation from LGE MRI Using Deep Neural Network Incorporating Shape and Spatial Priors," *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, Oct. 2019, vol. 11765 LNCS, pp. 559–567, doi: 10.1007/978-3-030-32245-8_62.

[251] C. Chen, C. Biffi, G. Tarroni, S. Petersen, W. Bai, and D. Rueckert, "Learning Shape Priors for Robust Cardiac MR Segmentation from Multi-view Images," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, 2019, pp. 523–531, doi: 10.1007/978-3-030-32245-8_58.

[252] A. Chartsias *et al.*, "Disentangled representation learning in cardiac image analysis," *Medical Image Analysis*, vol. 58, p. 101535, Dec. 2019, doi: 10.1016/j.media.2019.101535.

[253] J. Duan *et al.*, "Automatic 3D Bi-Ventricular Segmentation of Cardiac Images by a Shape-Refined Multi- Task Deep Learning Approach," *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2151–2164, 2019, doi: 10.1109/TMI.2019.2894322.

[254] J. Chu *et al.*, “Pay More Attention to Discontinuity for Medical Image Segmentation,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 166–175, doi: 10.1007/978-3-030-59719-1_17.

[255] P. Nardelli and R. San José Estépar, “Targeting Precision with Data Augmented Samples in Deep Learning,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, 2019, pp. 284–292, doi: 10.1007/978-3-030-32226-7_32.

[256] S S. Li, C. Zhang, and X. He, “Shape-Aware Semi-supervised 3D Semantic Segmentation for Medical Images,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 552–561, doi: 10.1007/978-3-030-59710-8_54.

[257] M. N. N. To *et al.*, “Improving Dense Pixelwise Prediction of Epithelial Density Using Unsupervised Data Augmentation for Consistency Regularization,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 572–581, doi: 10.1007/978-3-030-59710-8_56.

[258] P. K. Gyawali, S. Ghimire, P. Bajracharya, Z. Li, and L. Wang, “Semi-supervised Medical Image Classification with Global Latent Mixing,” in *MICCAI 2020*, pp. 604–613, doi: 10.1007/978-3-030-59710-8_59.

[259] Q. Liu, Q. Dou, and P.-A. Heng, “Shape-Aware Meta-learning for Generalizing Prostate MRI Segmentation to Unseen Domains,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 475–485, doi: 10.1007/978-3-030-59713-9_46.

[260] N. Byrne, J. R. Clough, G. Montana, and A. P. King, “A Persistent Homology-Based Topological Loss Function for Multi-class CNN Segmentation of Cardiac MRI,” in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges. STACOM 2020*, vol. 12592 LNCS, pp. 3–13, doi: 10.1007/978-3-030-68107-4_1.

[261] T. Liu, Y. Tian, S. Zhao, and X. Huang, “Graph Reasoning and Shape Constraints for Cardiac Segmentation in Congenital Heart Defect,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 607–616, doi: 10.1007/978-3-030-59719-1_59.

[262] D. Sun *et al.*, “Automatic Tooth Segmentation and Dense Correspondence of 3D Dental Model,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 703–712, doi: 10.1007/978-3-030-59719-1_68.

[263] S. Valipour, M. Siam, M. Jagersand, and N. Ray, “Recurrent Fully Convolutional Networks for Video Segmentation,” in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Mar. 2017, pp. 29–36, doi: 10.1109/WACV.2017.11.

[264] T. Young, D. Hazarika, S. Poria, and E. Cambria, “Recent Trends in Deep Learning Based Natural Language Processing,” *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, Aug. 2018, doi: 10.1109/MCI.2018.2840738.

[265] R. P. K. Poudel, P. Lamata, and G. Montana, “Recurrent Fully Convolutional Neural Networks for Multi-slice MRI Cardiac Segmentation,” in *Reconstruction, Segmentation, and Analysis of Medical Images. RAMBO 2016, HVSMR 2016*, Springer, Cham, 2017, pp. 83–94, doi: 10.1007/978-3-319-52280-7_8.

[266] D. Zhang *et al.*, “A multi-level convolutional LSTM model for the segmentation of left ventricle myocardium in infarcted porcine cine MR images,” in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Apr. 2018, pp. 470–473, doi: 10.1109/ISBI.2018.8363618.

[267] J.-Y. Lin, Y.-C. Chang, and W. H. Hsu, “Efficient and Phase-Aware Video Super-Resolution for Cardiac MRI,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 66–76, doi: 10.1007/978-3-030-59719-1_7.

[268] S. Dong *et al.*, “DeU-Net: Deformable U-Net for 3D Cardiac MRI Video Segmentation,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 98–107, doi: 10.1007/978-3-030-59719-1_10.

[269] H. Wei *et al.*, “Temporal-Consistent Segmentation of Echocardiography with Co-learning from Appearance and Shape,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 623–632, doi: 10.1007/978-3-030-59713-9_60.

[270] Q. Huang *et al.*, “Measure Anatomical Thickness from Cardiac MRI with Deep Neural Networks,” in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, 2021, pp. 44–55, doi: 10.1007/978-3-030-68107-4_5.

[271] G. Yang, T. Hua, W. Xue, and L. Shuo, “Left Ventricle Full Quantification Challenge MICCAI 2019,” Available: https://lvquan19.github.io/ (Accessed Jul. 01, 2020).

[272] V. M. Campello *et al.*, “Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge,” *IEEE Transactions on Medical Imaging*, pp. 1–1, Jun. 2021, doi: 10.1109/tmi.2021.3090082.

[273] M. Regehr, A. Volk, M. Noga, and K. Punithakumar, “Machine Learning and Graph Based Approach to Automatic Right Atrial Segmentation from Magnetic Resonance Imaging,” in *Proceedings - International Symposium on Biomedical Imaging*, Apr. 2020, vol. 2020-April, pp. 826–829, doi: 10.1109/ISBI45749.2020.9098437.

[274] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A Comprehensive Survey on Graph Neural Networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, Jan. 2021, doi: 10.1109/tnnls.2020.2978386.

[275] Y. Meng *et al.*, “CNN-GCN Aggregation Enabled Boundary Regression for Biomedical Image Segmentation,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 352–362, doi: 10.1007/978-3-030-59719-1_35.

[276] A. Asperti and C. Mastronardo, “The Effectiveness of Data Augmentation for Detection of Gastrointestinal Diseases from Endoscopical Images,” *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*, pp. 199–205, 2017, doi: 10.5220/0006730901990205.

[277] L. Perez and J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” *arXiv*, 2017. Available: http://arxiv.org/abs/1712.04621.

[278] J. Corral Acero *et al.*, “SMOD - Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI,” in *Functional Imaging and Modeling of the Heart. FIMH 2019*, Jun. 2019, vol. 11504 LNCS, pp. 361–369, doi: 10.1007/978-3-030-21949-9_39.

[279] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, “Artificial intelligence in radiology,” *Nature Reviews Cancer*, vol. 18, no. 8, pp. 500–510, Aug. 2018, doi: 10.1038/s41568-018-0016-5.

[280] G. Litjens *et al.*, “State-of-the-Art Deep Learning in Cardiovascular Image Analysis,” *JACC: Cardiovascular Imaging*, vol. 12, no. 8P1. 2019, doi: 10.1016/j.jcmg.2019.06.009.

[281] W. Bai *et al.*, “Self-Supervised Learning for Cardiac MR Image Segmentation by Anatomical Position Prediction,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, 2019, pp. 541–549, doi: 10.1007/978-3-030-32245-8_60.

[282] X. Tao, Y. Li, W. Zhou, K. Ma, and Y. Zheng, “Revisiting Rubik’s Cube: Self-supervised Learning with Volume-Wise Transformation for 3D Medical Image Segmentation,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 238–248, doi: 10.1007/978-3-030-59719-1_24.

[283] S. Kayal, S. Chen, and M. de Bruijne, "Region-of-Interest Guided Supervoxel Inpainting for Self-supervision," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 500–509, doi: 10.1007/978-3-030-59710-8_49.

[284] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10. pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.

[285] J. Ma, "Histogram Matching Augmentation for Domain Adaptation with Application to Multi-centre, Multi-vendor and Multi-disease Cardiac Image Segmentation," in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, 2021, pp. 177–186, doi: 10.1007/978-3-030-68107-4_18.

[286] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," *Machine Learning*, vol. 79, no. 1–2, pp. 151–175, Oct. 2010, doi: 10.1007/s10994-009-5152-4.

[287] M. Parreño, R. Paredes, and A. Albiol, "Deidentifying MRI Data Domain by Iterative Backpropagation," in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, 2021, pp. 277–286, doi: 10.1007/978-3-030-68107-4_28.

[288] Y. Zhang *et al.*, "Semi-supervised Cardiac Image Segmentation via Label Propagation and Style Transfer," in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, 2021, pp. 219–227, doi: 10.1007/978-3-030-68107-4_22.

[289] Y. Xue, S. Feng, Y. Zhang, X. Zhang, and Y. Wang, "Dual-Task Self-supervision for Cross-modality Domain Adaptation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 408–417, doi: 10.1007/978-3-030-59710-8_40.

[290] Y. He, A. Carass, L. Zuo, B. E. Dewey, and J. L. Prince, "Self Domain Adapted Network," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 437–446, doi: 10.1007/978-3-030-59710-8_43.

[291] J. Corral Acero, V. Sundaresan, N. Dinsdale, V. Grau, and M. Jenkinson, "A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation," in *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges. STACOM 2020*, LNCS, vol. 12592, Springer, Cham, 2021, pp. 196–207, doi: 10.1007/978-3-030-68107-4_20.

[292] B. Murugesan, K. Sarveswaran, V. S. Raghavan, S. M. Shankaranarayana, K. Ram, and M. Sivaprakasam, "A Context Based Deep Learning Approach for Unbalanced Medical Image Segmentation," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 1949–1953, doi: 10.1109/ISBI45749.2020.9098597.

[293] J. Corral Acero *et al.*, "Left Ventricle Quantification with Cardiac MRI: Deep Learning Meets Statistical Models of Deformation," in *Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. STACOM 2019*, LNCS, Oct. 2020, vol. 12009, pp. 384–394, doi: 10.1007/978-3-030-39074-7_40.

[294] B. Audelan, D. Hamzaoui, S. Montagne, R. Renard-Penna, and H. Delingette, "Robust Fusion of Probability Maps," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 259–268, doi: 10.1007/978-3-030-59719-1_26.

[295] L. Zhang *et al.*, "Learning to Segment When Experts Disagree," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 179–190, doi: 10.1007/978-3-030-59710-8_18.

[296] T. Moroseos *et al.*, "Comparison of Simpson's method and three-dimensional reconstruction for measurement of right ventricular volume in patients with complete or corrected

transposition of the great arteries.," *The American journal of cardiology*, vol. 105, no. 11, pp. 1603–9, Jun. 2010, doi: 10.1016/j.amjcard.2010.01.025.

[297] M. S. Nacif *et al.*, "Three-Dimensional Volumetric Assessment of Diastolic Function by Cardiac Magnetic Resonance Imaging: The Multi-Ethnic Study of Atherosclerosis (MESA).," *Arquivos brasileiros de cardiologia*, vol. 108, no. 6, pp. 552–563, 2017, doi: 10.5935/abc.20170063.

[298] P. Medrano-Gracia *et al.*, "An Atlas for Cardiac MRI Regional Wall Motion and Infarct Scoring," *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges. STACOM 2012*, Springer, Berlin, Heidelberg, 2013, pp. 188–197, doi: 10.1007/978-3-642-36961-2_22.

[299] B. Villard, V. Grau, E. Zacur, B. Villard, V. Grau, and E. Zacur, "Surface Mesh Reconstruction from Cardiac MRI Contours," *Journal of Imaging*, vol. 4, no. 1, p. 16, Jan. 2018, doi: 10.3390/jimaging4010016.

[300] P. Lamata *et al.*, "An automatic service for the personalization of ventricular cardiac meshes," *Journal of The Royal Society Interface*, vol. 11, no. 91, pp. 20131023–20131023, Dec. 2013, doi: 10.1098/rsif.2013.1023.

[301] A. A. Young and A. F. Frangi, "Computational cardiac atlases: from patient to population and back," *Experimental Physiology*, vol. 94, no. 5, pp. 578–596, May 2009, doi: 10.1113/expphysiol.2008.044081.

[302] H. Lehmann *et al.*, "Integrating Viability Information into a Cardiac Model for Interventional Guidance," *Functional Imaging and Modeling of the Heart. FIMH 2009*, Springer, Berlin, Heidelberg, 2009, pp. 312–320, doi: 10.1007/978-3-642-01932-6_34.

[303] M. Vukicevic, B. Mosadegh, J. K. Min, and S. H. Little, "Cardiac 3D Printing and its Future Directions," *JACC: Cardiovascular Imaging*, vol. 10, no. 2, pp. 171–184, Feb. 2017, doi: 10.1016/j.jcmg.2016.12.001.

[304] W. Bai *et al.*, "A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion," *Medical Image Analysis*, vol. 26, no. 1, pp. 133–145, Dec. 2015, doi: 10.1016/J.MEDIA.2015.08.009.

[305] N. Smith *et al.*, "euHeart: personalized and integrated cardiac care using patient-specific cardiovascular modelling," *Interface Focus*, vol. 1, no. 3, pp. 349–364, Jun. 2011, doi: 10.1098/rsfs.2010.0048.

[306] M. Berger *et al.*, "A Survey of Surface Reconstruction from Point Clouds," *Computer Graphics Forum*, vol. 36, no. 1, pp. 301–329, Jan. 2017, doi: 10.1111/cgf.12802.

[307] B. Villard, E. Zacur, E. Dall'Armellina, and V. Grau, "Correction of Slice Misalignment in Multi-breath-hold Cardiac MRI Scans," *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges. STACOM 2016*, Springer, Cham, 2017, pp. 30–38, doi: 10.1007/978-3-319-52718-5_4.

[308] Min Wan *et al.*, "Reconstructing patient-specific cardiac models from contours via Delaunay triangulation and graph-cuts," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2013, pp. 2976–2979, doi: 10.1109/EMBC.2013.6610165.

[309] C. W. Lim *et al.*, "Automatic 4D reconstruction of patient-specific cardiac mesh with 1-to-1 vertex correspondence from segmented contours lines.," *PloS one*, vol. 9, no. 4, p. e93747, 2014, doi: 10.1371/journal.pone.0093747.

[310] Z. Wang, N. Geng, and Z. Zhang, “Surface Mesh Reconstruction Based on Contours,” in *2009 International Conference on Computational Intelligence and Software Engineering*, Dec. 2009, pp. 1–4, doi: 10.1109/CISE.2009.5365813.

[311] C. Mauger *et al.*, “An Iterative Diffeomorphic Algorithm for Registration of Subdivision Surfaces: Application to Congenital Heart Disease,” *IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2018, p. 596, Oct. 2018, doi: 10.1109/EMBC.2018.8512394.

[312] H. Hoppe *et al.*, “Piecewise smooth surface reconstruction,” in *Proceedings of the 21st annual conference on Computer graphics and interactive techniques - SIGGRAPH ’94*, 1994, pp. 295–302, doi: 10.1145/192161.192233.

[313] C. M. Lee, F. H. Sheehan, B. Bouzas, S. S. M. Chen, M. A. Gatzoulis, and P. J. Kilner, “The shape and function of the right ventricle in Ebstein’s anomaly,” *International Journal of Cardiology*, vol. 167, no. 3, pp. 704–710, Aug. 2013, doi: 10.1016/J.IJCARD.2012.03.062.

[314] A. A. Young, B. R. Cowan, S. F. Thrupp, W. J. Hedley, and L. J. Dell’Italia, “Left Ventricular Mass and Volume: Fast Calculation with Guide-Point Modeling on MR Images,” *Radiology*, vol. 216, no. 2, pp. 597–602, Aug. 2000, doi: 10.1148/radiology.216.2.r00au14597.

[315] D. A. Auger, X. Zhong, F. H. Epstein, E. M. Meintjes, and B. S. Spottiswoode, “Semi-automated left ventricular segmentation based on a guide point model approach for 3D cine DENSE cardiovascular magnetic resonance.,” *Journal of cardiovascular magnetic resonance*, vol. 16, no. 1, p. 8, Jan. 2014, doi: 10.1186/1532-429X-16-8.

[316] X.-F. Han, H. Laga, and M. Bennamoun, “Image-Based 3D Object Reconstruction: State-of-the-Art and Trends in the Deep Learning Era,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, n5, pp. 1578–1604, 2021, doi: 10.1109/TPAMI.2019.2954885.

[317] H. Xu, E. Zacur, J. E. Schneider, and V. Grau, “Ventricle Surface Reconstruction from Cardiac MR Slices Using Deep Learning.”, *Functional Imaging and Modeling of the Heart. FIMH 2019*, LNCS, vol 11504. Springer, Cham, doi: 10.1007/978-3-030-21949-9_37.

[318] W. E. Lorensen and H. E. Cline, “Marching cubes: a high resolution 3d surface construction algorithm.,” *Computer Graphics (ACM)*, vol. 21, no. 4, pp. 163–169, Aug. 1987, doi: 10.1145/37402.37422.

[319] D. Yang, P. Wu, C. Tan, K. M. Pohl, L. Axel, and D. Metaxas, “3D motion modeling and reconstruction of left ventricle wall in cardiac MRI,” in *Functional Imaging and Modelling of the Heart. FIMH 2017*, vol. 10263 LNCS, pp. 481–492, doi: 10.1007/978-3-319-59448-4_46.

[320] Z. Wu *et al.*, “3D ShapeNets: A deep representation for volumetric shapes,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Oct. 2015, vol. 07-12-June-2015, pp. 1912–1920, doi: 10.1109/CVPR.2015.7298801.

[321] I. J. Goodfellow *et al.*, “Generative Adversarial Nets,” in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, 2014, pp. 2672–2680.

[322] Y. Xue, T. Xu, H. Zhang, R. Long, and X. Huang, “SegAN: Adversarial Network with Multi-scale L_1 Loss for Medical Image Segmentation,” *Neuroinform* 16, 383–392 (2018), doi: 10.1007/s12021-018-9377-x.

[323] D. Yang *et al.*, “Automatic Liver Segmentation Using an Adversarial Image-to-Image Network,” in *Medical Image Computing and Computer-Assisted Intervention − MICCAI 2017*, 2017, pp. 507–515, doi: 10.1007/978-3-319-66179-7_58.

[324] V. K. Singh *et al.*, “Breast Tumor Segmentation and Shape Classification in Mammograms using Generative Adversarial and Convolutional Neural Network,” *arXiv*, Sep. 2018. Available: https://arxiv.org/abs/1809.01687.

[325] D. P. Kingma and M. Welling, “Auto-Encoding Variational Bayes,” *arXiv*, Dec. 2013. Available: https://arxiv.org/abs/1312.6114.

[326] D. J. Im, S. Ahn, R. Memisevic, and Y. Bengio, “Denoising criterion for variational auto-encoding framework,” in *31st AAAI Conference on Artificial Intelligence, AAAI 2017*, pp. 2059–2065. Available: https://ojs.aaai.org/index.php/AAAI/article/view/10777.

[327] K. Sohn, H. Lee, and X. Yan, “Learning Structured Output Representation using Deep Conditional Generative Models,” in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 3483–3491.

[328] J. J. Cerrolaza *et al.*, “3D Fetal Skull Reconstruction from 2DUS via Deep Conditional Generative Networks,” in *Medical Image Computing and Computer Assisted Intervention -- MICCAI 2018*, 2018, pp. 383–391, doi: 10.1007/978-3-030-00928-1_44.

[329] M. Mirza and S. Osindero, “Conditional Generative Adversarial Nets,” *arXiv*, Nov. 2014. Available: https://arxiv.org/abs/1411.1784.

[330] M. Rezaei, H. Yang, and C. Meinel, “Whole Heart and Great Vessel Segmentation with Context-aware of Generative Adversarial Networks,” *Bildverarbeitung für die Medizin 2018. Informatik aktuell*, Springer, Berlin, 2018, pp. 353–358, doi: 10.1007/978-3-662-56537-7_89.

[331] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[332] A. Bhattacharyya, B. Schiele, and M. Fritz, “Accurate and Diverse Sampling of Sequences Based on a ‘Best of Many’ Sample Objective,” *arXiv*, Jun. 2018. Available: https://arxiv.org/abs/1806.07772.

[333] U. Wickramasinghe, E. Remelli, G. Knott, and P. Fua, “Voxel2Mesh: 3D Mesh Model Generation from Volumetric Data,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 299–308, doi: 10.1007/978-3-030-59719-1_30.

[334] K. Tóthová *et al.*, “Probabilistic 3D Surface Reconstruction from Sparse MRI Information,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 813–823, doi: 10.1007/978-3-030-59710-8_79.

[335] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation,” *arXiv*, Dec. 2016. Available: arxiv.org/abs/1612.00593.

[336] F. Madesta, R. Schmitz, T. Rösch, and R. Werner, “Widening the Focus: Biomedical Image Segmentation Challenges and the Underestimated Role of Patch Sampling and Inference Strategies,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 289–298, doi: 10.1007/978-3-030-59719-1_29.

[337] Y. Song *et al.*, “Learning 3D Features with 2D CNNs via Surface Projection for CT Volume Segmentation,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020, pp. 176–186, doi: 10.1007/978-3-030-59719-1_18.

[338] B. Gutierrez-Becker and C. Wachinger, “Deep Multi-Structural Shape Analysis: Application to Neuroanatomy,” *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. LNCS, vol 11072. Springer, Cham, doi: 10.1007/978-3-030-00931-1_60.

[339] F. G. Zanjani *et al.*, “Mask-MCNet: Instance Segmentation in 3D Point Cloud of Intra-oral Scans,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, 2019, pp. 128–136, doi: 10.1007/978-3-030-32254-0_15.

[340] M. Beetz, A. Banerjee, and V. Grau, “Biventricular Surface Reconstruction From Cine Mri Contours Using Point Completion Networks,” in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, Apr. 2021, pp. 105–9, doi: 10.1109/ISBI48211.2021.9434040.

[341] C. Hane, S. Tulsiani, and J. Malik, “Hierarchical surface prediction for 3D object reconstruction,” in *Proceedings - 2017 International Conference on 3D Vision, 3DV 2017*, May 2018, pp. 412–420, doi: 10.1109/3DV.2017.00054.

[342] M. Simonovsky and N. Komodakis, “Dynamic edge-conditioned filters in convolutional neural networks on graphs,” in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Nov. 2017, vol. 2017, pp. 29–38, doi: 10.1109/CVPR.2017.11.

[343] A. Suinesiaputra *et al.*, “Statistical Shape Modeling of the Left Ventricle: Myocardial Infarct Classification Challenge,” *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 2, pp. 503–515, Mar. 2018, doi: 10.1109/JBHI.2017.2652449.

[344] B. Ambale-Venkatesh *et al.*, “Left ventricular shape predicts different types of cardiovascular events in the general population.,” *Heart (British Cardiac Society)*, vol. 103, no. 7, pp. 499–507, Apr. 2017, doi: 10.1136/heartjnl-2016-310052.

[345] F. Vadakkumpadan, N. Trayanova, and K. C. Wu, “Image-based left ventricular shape analysis for sudden cardiac death risk stratification,” *Heart Rhythm*, vol. 11, no. 10, pp. 1693–1700, Oct. 2014, doi: 10.1016/J.HRTHM.2014.05.018.

[346] B. Amberg, S. Romdhani, and T. Vetter, “Optimal Step Nonrigid ICP Algorithms for Surface Registration,” in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2007, pp. 1–8, doi: 10.1109/CVPR.2007.383165.

[347] T. F. Cootes and C. Taylor, *Statistical Models of Appearance for computer vision*, vol. 1. Technical report, University of Manchester, 2004.

[348] K. McLeod, T. Mansi, M. Sermesant, G. Pongiglione, and X. Pennec, “Statistical Shape Analysis of Surfaces in Medical Images Applied to the Tetralogy of Fallot Heart.” *Modeling in Computational Biology and Biomedicine*. 2012, Springer, Berlin, Heidelbergm, doi: 10.1007/978-3-642-31208-3_5.

[349] X. Zhang *et al.*, “Atlas-Based Quantification of Cardiac Remodeling Due to Myocardial Infarction,” *PLoS ONE*, vol. 9, no. 10, p. e110243, Oct. 2014, doi: 10.1371/journal.pone.0110243.

[350] G. Farrar *et al.*, “Atlas-based ventricular shape analysis for understanding congenital heart disease,” *Progress in Pediatric Cardiology*, vol. 43, pp. 61–69, Dec. 2016, doi: 10.1016/J.PPEDCARD.2016.07.010.

[351] D. R. Warriner *et al.*, “An Asymmetric Wall-Thickening Pattern Predicts Response to Cardiac Resynchronization Therapy,” *JACC: Cardiovascular Imaging*, vol. 11, no. 10, pp. 1545–1546, Oct. 2018, doi: 10.1016/J.JCMG.2018.01.022.

[352] P. Medrano-Gracia *et al.*, “Left ventricular shape variation in asymptomatic populations: the multi-ethnic study of atherosclerosis,” *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 1, p. 56, Dec. 2014, doi: 10.1186/s12968-014-0056-2.

[353] H. K. Narayan *et al.*, “Atlas-based measures of left ventricular shape may improve characterization of adverse remodeling in anthracycline-exposed childhood cancer survivors:

a cross-sectional imaging study," *Cardio-Oncology*, vol. 6, no. 1, Dec. 2020, doi: 10.1186/S40959-020-00069-5.

[354] X. Zhang *et al.*, "Information maximizing component analysis of left ventricular remodeling due to myocardial infarction.," *Journal of translational medicine*, vol. 13, p. 343, Nov. 2015, doi: 10.1186/s12967-015-0709-4.

[355] Juwei Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face recognition using LDA-based algorithms," *IEEE Transactions on Neural Networks*, vol. 14, no. 1, pp. 195–200, Jan. 2003, doi: 10.1109/TNN.2002.806647.

[356] X. Zhang, P. Medrano-Gracia, B. Ambale-Venkatesh, *et al.*, "Orthogonal decomposition of left ventricular remodeling in myocardial infarction," *GigaScience*, vol. 6, no. 3, pp. 1–15, Mar. 2017, doi: 10.1093/GIGASCIENCE/GIX005.

[357] D. Yang, Q. Huang, K. Mikael, S. al Al'aref, L. Axel, and D. Metaxas, "Mri-Based Characterization of Left Ventricle Dyssynchrony with Correlation to Crt Outcomes," in *Proceedings - International Symposium on Biomedical Imaging*, Apr. 2020, vol. 2020-April, pp. 822–825, doi: 10.1109/ISBI45749.2020.9098519.

[358] D. Yang, Q. Huang, K. Mikael, S. al Al'aref, L. Axel, and D. Metaxas, "Mri-Based Characterization of Left Ventricle Dyssynchrony with Correlation to Crt Outcomes," in *Proceedings - International Symposium on Biomedical Imaging*, Apr. 2020, vol. 2020-April, pp. 822–825, doi: 10.1109/ISBI45749.2020.9098519.

[359] Q. Tan, L. Gao, Y.-K. Lai, and S. Xia, "Variational Autoencoders for Deforming 3D Mesh Models," *arXiv*, Sep. 2017. Available: https://arxiv.org/abs/1709.04307.

[360] L. Hoang, S. H. Lee, O. H. Kwon, and K. R. Kwon, "A deep learning method for 3D object classification using the wave kernel signature and a center point of the 3D-triangle mesh," *Electronics (Switzerland)*, vol. 8, no. 10, p. 1196, Oct. 2019, doi: 10.3390/electronics8101196.

[361] R. Hanocka, A. Hertz, N. Fish, R. Giryes, S. Fleishman, and D. Cohen-Or, "MeshCNN: A Network with an Edge," *ACM Transactions on Graphics*, vol. 38, no. 4, Sep. 2018, doi: 10.1145/3306346.3322959.

[362] C. Biffi *et al.*, "Explainable Anatomical Shape Analysis through Deep Hierarchical Generative Models," *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 2088–2099, Jun. 2020, doi: 10.1109/TMI.2020.2964499.

[363] M. Reindl, I. Eitel, and S. J. Reinstadler, "Role of Cardiac Magnetic Resonance to Improve Risk Prediction following Acute ST-elevation Myocardial Infarction," *Journal of Clinical Medicine*, vol. 9, no. 4, p. 1041, Apr. 2020, doi: 10.3390/jcm9041041.

[364] V. M. Campello, "Multi-Centre, Multi-Vendor & Multi-Disease Cardiac Image Segmentation Challenge (M&Ms) MICCAI 2020." https://www.ub.edu/mnms/ (accessed Jul. 01, 2020).

[365] P. Lamata *et al.*, "An accurate, fast and robust method to generate patient-specific cubic Hermite meshes," *Medical Image Analysis*, vol. 15, no. 6, pp. 801–813, Dec. 2011, doi: 10.1016/j.media.2011.06.010.

[366] P. Lamata *et al.*, "An automatic service for the personalization of ventricular cardiac meshes," *Journal of the Royal Society Interface*, vol. 11, no. 91, pp. 20131023–20131023, Feb. 2014, doi: 10.1098/rsif.2013.1023.

[367] M. J. Gonzales *et al.*, "A three-dimensional finite element model of human atrial anatomy: New methods for cubic Hermite meshes with extraordinary vertices," *Medical Image Analysis*, vol. 17, no. 5, pp. 525–537, Jul. 2013, doi: 10.1016/J.MEDIA.2013.03.005.

[368] A. Krishnamurthy, M. J. Gonzales, G. Sturgeon, W. P. Segars, and A. D. McCulloch, "Biomechanics simulations using cubic Hermite meshes with extraordinary nodes for isogeometric cardiac modeling," *Computer Aided Geometric Design*, vol. 43, pp. 27–38, Mar. 2016, doi: 10.1016/J.CAGD.2016.02.016.

[369] B. Ruijsink *et al.*, "Fully Automated, Quality-Controlled Cardiac Analysis From CMR: Validation and Large-Scale Application to Characterize Cardiac Function" *JACC Cardiovasc. Imaging*, vol. 13, no. 3, pp. 684–695, Mar. 2020, doi: 10.1016/j.jcmg.2019.05.030.

APPENDICES

Appendix A:

Supplemental Material Paper V

SUPPLEMENTAL MATERIAL

Understanding and Improving Risk Assessment after Myocardial Infarction using Automated Left Ventricular Shape Analysis

Jorge Corral Acero, Andreas Schuster, Ernesto Zacur, Torben Lange, Thomas Stiermaier, Sören J. Backhaus, Holger Thiele, Alfonso Bueno-Orovio, Pablo Lamata, Ingo Eitel, Vicente Grau.

Table of Contents

SUPPLEMENTAL MATERIAL I – LV Segmentation 2
SUPPLEMENTAL MATERIAL II – Dimensionality Reduction 3
SUPPLEMENTAL MATERIAL III – Endpoint Prediction Additional Results 5
SUPPLEMENTAL MATERIAL IV – Independent Testing 10
SUPPLEMENTAL MATERIAL V – Subgroup Analysis 12
SUPPLEMENTAL MATERIAL VI – Reproducibility Analysis 13
SUPPLEMENTAL MATERIAL VII – Modes Correlation Analysis 16
SUPPLEMENTAL MATERIAL VIII – TRIPOD Reporting 18
SUPPLEMENTAL MATERIAL IX – ES Shape Modes 20
SUPPLEMENTAL MATERIAL X – 3D Contraction Modes 22

SUPPLEMENTAL MATERIAL I – LV Segmentation

A 2-step state-of-art approach was applied to segment the LV endo- and epicardium, from which the LV myocardium can be estimated (14,15):

- In the first step, pre-processing, the images are re-oriented, cropped and normalized. Considering the mid-ventricular SAx slice as reference, the first neural network (NN) detects the position of the heart and defines a region of interest (ROI) of 139.7x139.7 mm centered in the LV. Based on these LV centroid and ROI, the LV is aligned to a canonical position by a rigid registration of the same SAx reference slice to the atlas built in (14). The SAx slices are then cropped accordingly, and the intensities are normalized.
- The second step, fine segmentation, applies another NN to the pre-processed images to regress the enhanced LV segmentation. A final postprocessing is applied to binarize the segmentation predictions and improve the segmentation quality.

A cohort of 100 patients of the study, manually segmented with random resolutions and endpoints, was used to train the 2 NNs, following a 5-fold cross-validation strategy and using the same training-validation-testing split ratio as in (14). Translation, rotation and flipping were used for augmentation. Architectures and implementation are detailed in (14,15). Segmentation performance assessment is based on endocardium and epicardium gold standard Dice scores. This metric accounts for the overlap between manual segmentation and automated prediction and varies between 0 and 1, with 1 corresponding to a perfect match. Fig. SI.1 provides a visual sample of the segmentation results.

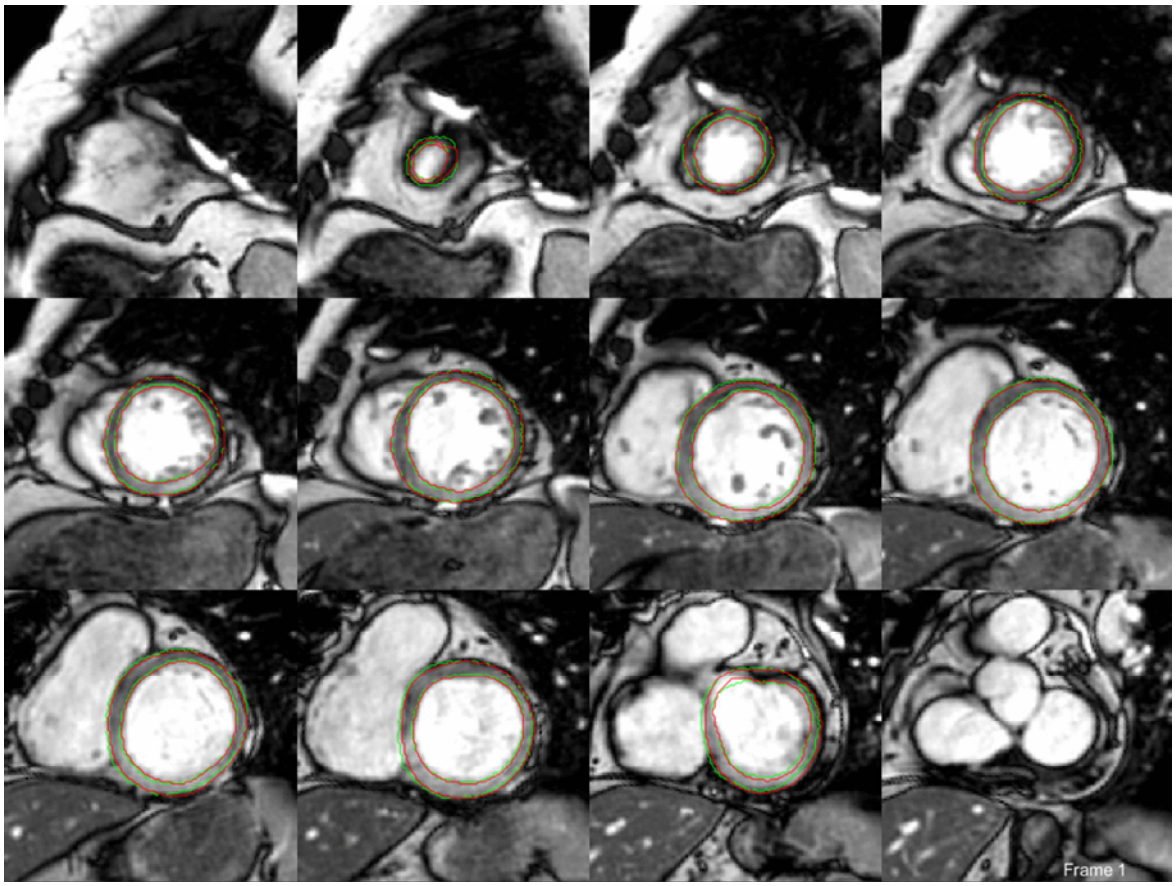


Fig. SI.1 – Segmentation results of a representative patient (median Dice) at ED. The green contours correspond to the LV reference segmentation (manual segmentation); and the red contours, to the prediction results based on our proposed 2-step deep learning approach.

SUPPLEMENTAL MATERIAL II – Dimensionality Reduction

As explained in the manuscript, the personalized LV shape of each patient is obtained by fitting (deforming) a template mesh to the myocardial segmentation of the patient. The LV shape of the patient is therefore described by the resulting 3D position (x, y and z cartesian coordinates) of the 2,450 nodes that conform the mesh template, that is, 7,350 variables (nodes × coordinates), which take a certain different value per patient. PCA, the dimensionality reduction technique applied in this work, is able to encode the information contained in these 7,350 variables (shape variability) into a few variables, the anatomical PCA modes, that describe the main shape variations. Mathematically, this is done by finding the vector space whose basis are the orthogonal directions that maximize the variance of the data, and subsequently projecting the data in this new space (See Fig. SII.1). These directions that maximize the variance are, precisely, the PCA modes of variation that represent the way in which the shape varies in the population (i.e. thickening, scaling, lengthening, etc.). As illustrated in Fig. SII.2, if we move in the direction of a mode, we can see how the mean shape deforms in its specific way. This allows describing each 3D mesh as a mean shape plus the amount of shape change (PCA coefficients) encoded by each linear anatomical mode (See Fig. SII.3):

$$Shape = \varphi_0 + \sum_i \omega_i \varphi_i,$$

where φ_0 represents the mean, φ_i the anatomical PCA modes, and ω_i their respective PCA coefficients. Each of the modes, i, is therefore a continuous variable, that accounts for a particular shape variation, which has a certain value for each patient, ω_i, and whose MACE predictive power can be analyzed. In other words, we can investigate which shape features (modes) are related to AMI prognosis.

The modes are sorted in descending order of importance according to the amount of variability of the population that they explain (i.e. in our results 42.6% variability is described by ES mode 1; 19.2%, by ES2; etc.). As we progressively incorporate modes, the shape reconstruction improves to the point that with only a few modes we can accurately approximate any shape (i.e. 95% of ES shape variance described by the first 12 modes), and hence the dimensionality reduction is achieved (See Fig. SII.4).

The same principles, herein illustrated for ES shape, apply to contraction.

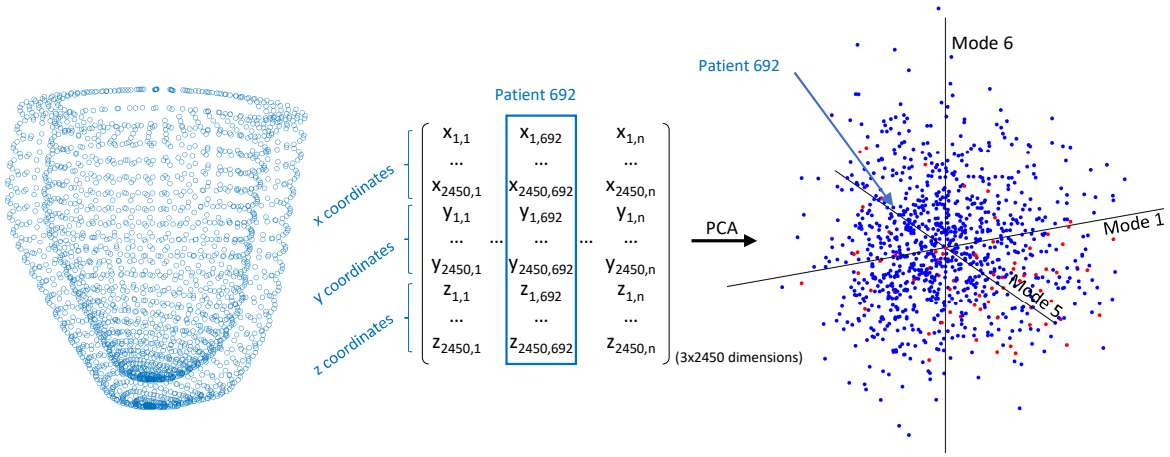


Fig. SII.1 – **LEFT**: 3D spatial representation of the 2,450 mesh nodes for patient 692. **MIDDLE**: The 7,350 variables that describe the patient mesh are sorted in columns, one per patient, to apply PCA. **RIGHT**: Projection of the patients in the PCA directions that maximize MACE (red) vs No MACE (blue) differences, i.e. modes 1, 5 and 6. We reduced the number of variables from 7,350 to 3 per patient.

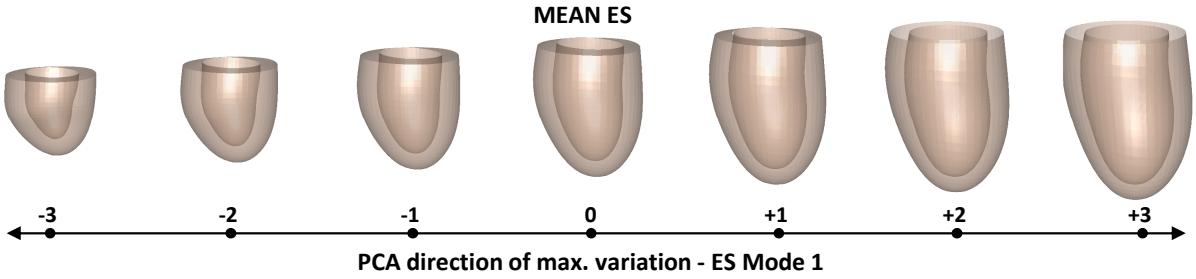


Fig. SII.2 – The PCA direction of maximum variability corresponds to ES mode 1. In our results, the shape variation that it encodes is interpreted as mainly scaling (related to ESV, as described in the main manuscript), although some other changes in thickness and morphology can be appreciated. As we move along the ES mode 1 direction, the mean ES shape is reduced (negative ω_1) or enlarged (positive ω_1), proportionally to the value of its PCA coefficient ω_1.

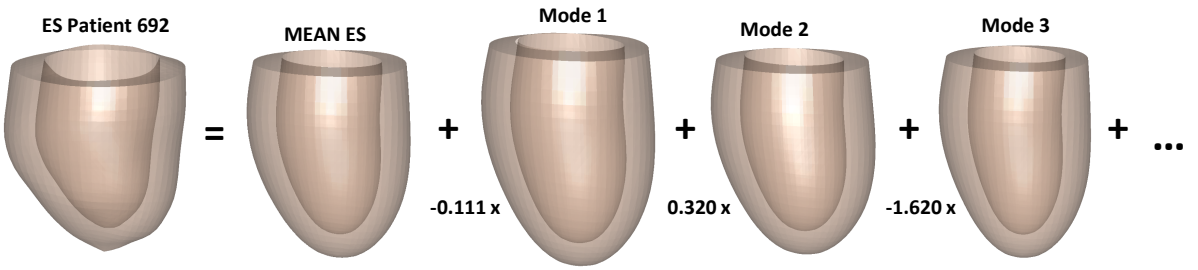


Fig. SII.3 – Following PCA application, each LV is decomposed into the mean shape (average LV) plus the anatomical modes (shape variations, illustrated here as mean plus a mode representation) times the corresponding PCA coefficient. The figure illustrates the decomposition of patient 692.

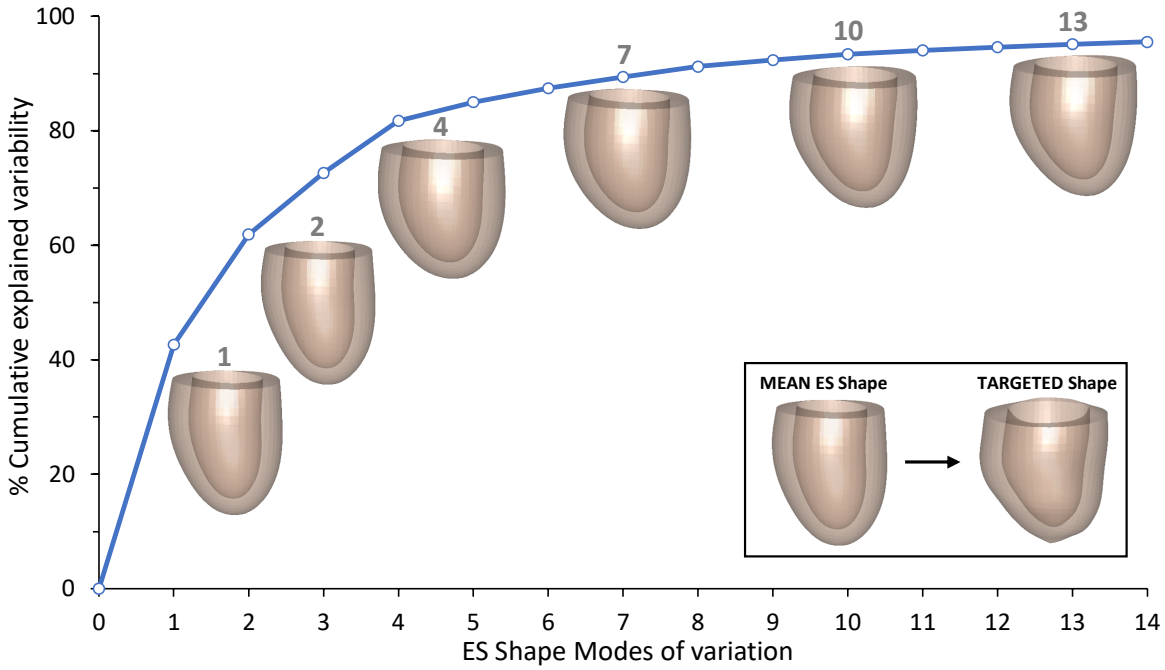


Fig. SII.4 – Cumulative % of the population variance explained by the modes. As we incorporate more modes, we account for more variance and we accurately approximate the target shape (Patient 692). The first few modes account for the majority of the variance, and as we move to latter modes the improvement in reconstruction is smaller (i.e. mode 10 vs mode 13).

SUPPLEMENTAL MATERIAL III – Endpoint Prediction Additional Results

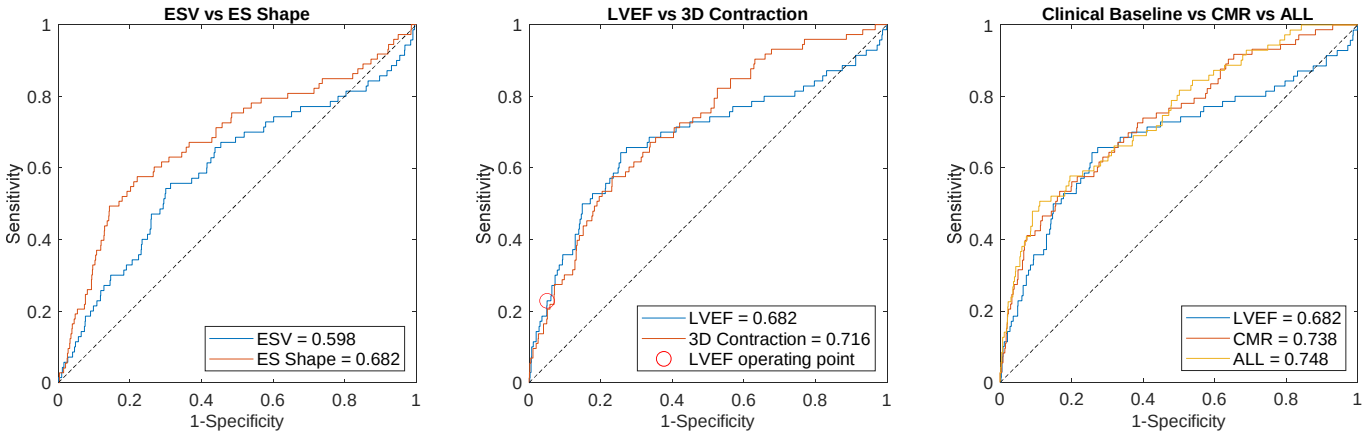


Fig. SIII.1 – Receiver-Operator Characteristic (ROC) curves illustrating the AUC differences in MACE prediction. The curves are obtained applying leave-one-out cross-validation, and the resulting AUCs are reported in the figure legends. The interpretation of the ROC in terms of specificity and sensitivity is explained in detail in Fig. SIII.4. Preventive treatment for patients at risk usually involves costly and invasive procedures (e.g. implantation of CRT devices). In such scenario, classifiers are typically operated at high sensitivity (i.e. ensuring that the patients predicted as at risk are indeed likely to suffer from MACE and, thus, minimizing false positives). This is the region where our proposed multivariable models most outperform the clinical baselines. **LEFT**: ESV vs ES shape. **MIDDLE**: LVEF vs 3D contraction. The operation point of the LVEF curve based on the threshold indicated in the clinical guidelines (0.35 LVEF) is illustrated (red circle). **RIGHT**: Best performing clinical baseline, i.e., LVEF vs the proposed multivariable models, combining the novel 3D descriptors with only CMR biomarkers (CMR) and with all the variables of the study (ALL).

Table SIII.1 – Multivariable Models: Endpoint prediction.

MODEL	LINEAR SELECTION	AUC_K	AUC_{RS}	COX SELECTION	$C\text{-index}_k$	$C\text{-index}_{RS}$
CMR Baseline	ESV, EDV	0.701 (0.699 - 0.702)	0.708	ESV, EDV	0.688 (0.687 - 0.690)	0.693
CMR Substitution	ES1, ES5, ES6, C3, C5, C16	0.729 (0.727 - 0.732)	0.748	ES1, ES5, ES6, C3, C5, C16	0.715 (0.713 - 0.718)	0.730
CMR	ESV, EDV, C5, C16	0.738 (0.736 - 0.740)	0.750	ESV, EDV, C5, C16	0.728 (0.727 - 0.730)	0.736
ALL Baseline	ESV, EDV, Age, Killip	0.728 (0.725 - 0.731)	0.745	ESV, EDV, Age, Killip	0.715 (0.713 - 0.718)	0.726
ALL Substitution	Age, Killip, BSA, ES1, C5, C16	0.757 (0.755 - 0.760)	0.776	Age, Killip, BSA, ES1, C5, C16, Vessels	0.746 (0.744 - 0.749)	0.764
ALL	ESV, EDV, Age, Killip, C5, C16	0.747 (0.745 - 0.749)	0.766	ESV, EDV, Age, C5, C16	0.741 (0.739 - 0.744)	0.753

Backward stepwise LDA and Cox results of the multivariable models including only CMR biomarkers (CMR) and all the variables of the study (ALL). While the baseline multivariable models exclude the novel 3D descriptors from the analysis, the substitution multivariable models do not consider the conventional imaging markers LVEF and ES. In the "CMR" and "ALL" models, both conventional and novel 3D markers are included as predictor candidates, as described in Methods. This demonstrates that ES shape and 3D contraction produce a significant overall improvement in a multivariable setting both in substitution and in combination with the conventional imaging markers. The ESV, EDV and LVEF calculated from automated volumes were considered for the analysis. The resulting significant selection of variables is reported along with the predictive performance, expressed as AUC and C-index re-substitution (RS) and 10-fold cross-validated (K), respectively. AUC_k and $C\text{-index}_k$ are presented as median (interquartile range). All differences were statistically significant (P<0.001). Killip indicates Killip class on admission; and vessel, number of diseased vessels.

Table SIII.2 – LDA models coefficients

MODEL	VARIABLES	LDA COEFFICIENTS	CONSTANT
ES Shape	ES1, ES5, ES6	0.42, -0.35, -0.46	-2.77
3D Contraction	C3, C5, C16	0.56, -0.53, -0.47	-2.89
CMR	ESV, EDV, C5, C16	1.76, -1.34, -0.33, -0.47	-3.07
ALL	ESV, EDV, Age, Killip, C5, C16	1.40, -0.98, 0.48, -0.34, -0.43, 0.51	-3.28

Coefficients and constant of the resulting ES shape, 3D contraction, CMR and ALL multivariable LDA models following the backward stepwise variable selection (see Methods). All the variables were normalized to zero mean and unit variance prior to LDA fitting. The performance of these models is summarized in Table 2 of the main manuscript.

Table SIII.3 – Cox models hazard ratios

TYPE	MODEL	VARIABLES	HR	*P* value
Clinical Baselines	ESV	ESV	1.43 (1.18 - 1.73)	< 0.001
	LVEF	LVEF	0.80 (0.74 - 0.87)	< 0.001
3D analysis	ES Shape	ES1	1.46 (1.17 - 1.83)	< 0.001
		ES5	0.73 (0.58 - 0.91)	0.005
		ES6	0.61 (0.48 - 0.78)	< 0.001
	3D Contraction	C3	1.66 (1.34 - 2.06)	< 0.001
		C5	0.63 (0.51 - 0.78)	< 0.001
		C16	0.66 (0.53 - 0.83)	< 0.001
Multivariable	CMR	ESV	4.23 (2.72 - 6.60)	< 0.001
		EDV	0.29 (0.18 - 0.48)	< 0.001
		C16	0.64 (0.51 - 0.80)	< 0.001
		C5	0.76 (0.61 - 0.95)	0.018
	ALL	ESV	3.41 (2.15 - 5.40)	< 0.001
		C16	0.65 (0.51 - 0.81)	< 0.001
		EDV	0.39 (0.23 - 0.65)	< 0.001
		Age	1.55 (1.19 - 2.02)	0.001
		C5	0.76 (0.60 - 0.95)	0.019

Hazard ratios, HR (95% confidence interval), and predictor significance, *P* value, of the selected variables in the resulting Cox models, presented in Table 2 of the main manuscript, along with their performance. All the variables were normalized to zero mean and unit variance prior to Cox analysis.

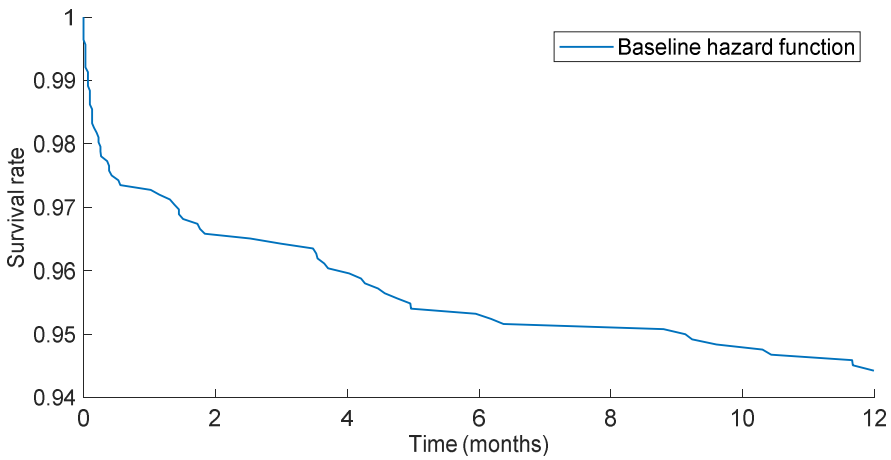


Fig. SIII.2 – Underlying baseline hazard function

Table SIII.4 – Endpoint prediction using volumes from manual segmentations

MODEL	LINEAR SELECTION	AUC_K	AUC_{RS}	COX SELECTION	$C\text{-index}_k$	$C\text{-index}_{RS}$
CMR	ESV, EDV, C5, C16, ES1	0.754 (0.752 - 0.756)	0.767	ESV, EDV, C5, C16, ES1	0.737 (0.735 - 0.739)	0.748
ALL	ESV, EDV, Age, C5, C16, ES1	0.762 (0.760 - 0.765)	0.779	ESV, EDV, Age, C5, C16, ES1	0.746 (0.743 - 0.749)	0.760

Backward stepwise LDA and Cox results of the multivariable models including only CMR biomarkers (CMR) and all the variables of the study (ALL). The ESV, EDV and LVEF calculated from manual segmentations were selected for the analysis. The resulting significant selection of variables (in red, differences with respect to the selection results using automated volumes) is reported along with the predictive performance, expressed as AUC and C-index, under cross-validation (k=10) and re-substitution (RS). AUC_k and $C\text{-index}_k$ are presented as median (interquartile range). All differences were statistically significant (P<0.001). Note the improvement in performance achieved by these models in comparison to the results reported in Table 2 of the main manuscript, related to the same analysis considering automated volumes rather than manual ones. This improvement could be explained by a double observer effect, where ESV and EDV are manually obtained, and a second observer (a machine) gets the contours again to find the meshes and modes.

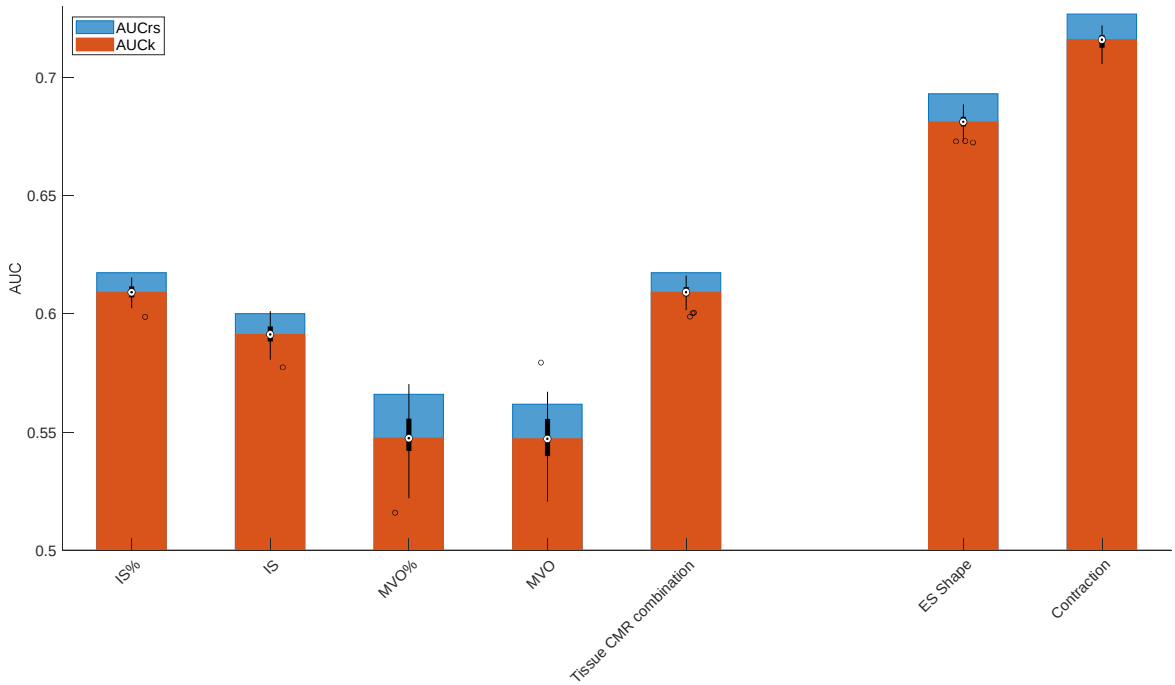


Fig. SIII.3 – MACE prediction comparison between the proposed ES shape and 3D contraction novel markers versus the tissue CMR biomarkers. Assessment based on AUC re-substitution (AUCrs, blue) and AUC 10-cross-fold validated (AUCk, orange), repeated for 100 random data splits (dark distributions). IS indicates infarct size; IS%, infarct size to LV mass ratio; MVO, microvascular obstruction; and MVO% microvascular obstruction to LV mass ratio. Tissue CMR combination is the resulting model from the backward stepwise LDA including all the tissue CMR biomarkers. Note that this model is identical to IS%, i.e. the rest of the tissue CMR biomarkers do not add significant prognostic value. All the differences in performance between the proposed 3D biomarkers and the tissue CMR ones were significant (P<0.001).

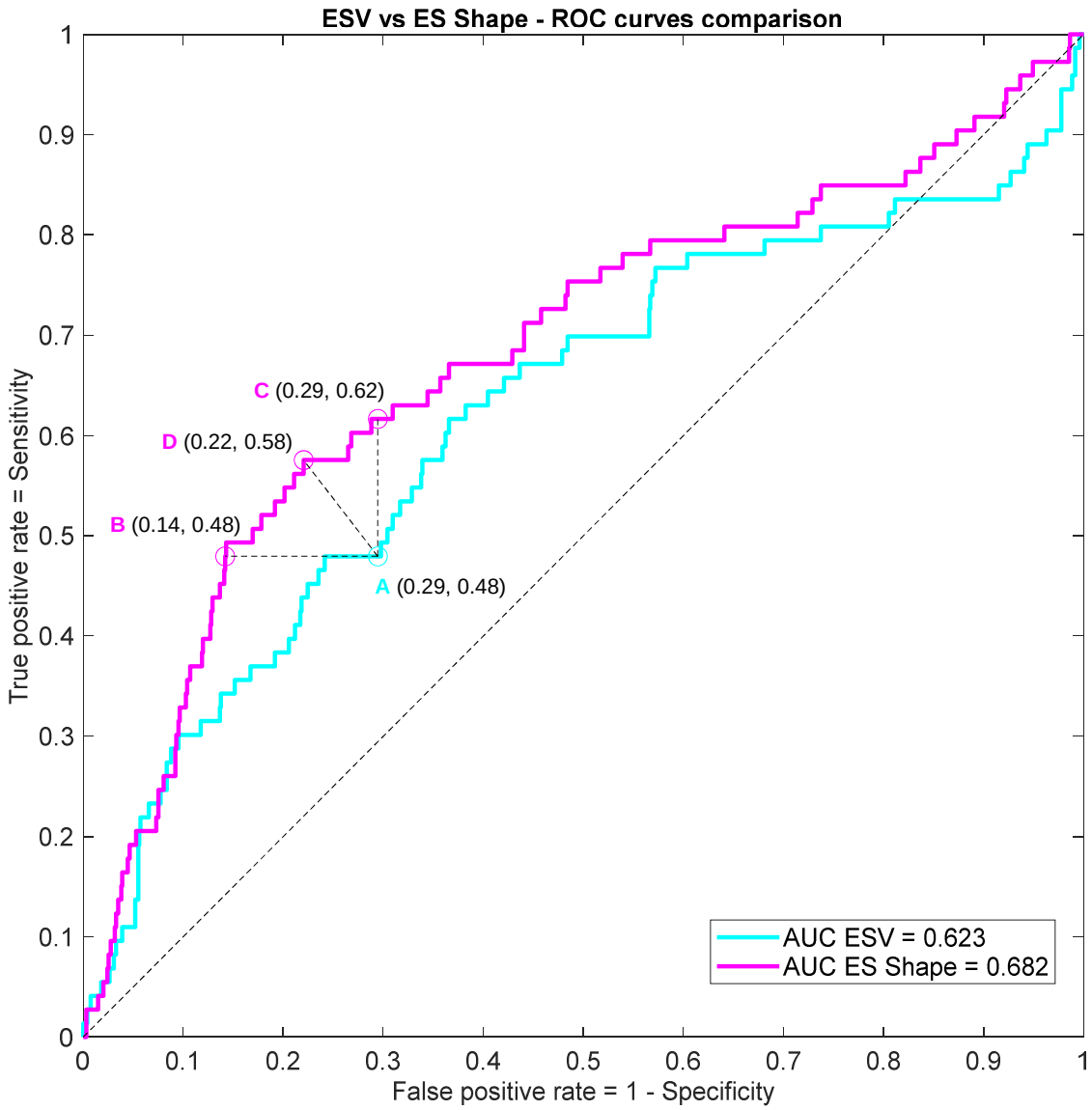


Fig. SIII.4 – Interpretation and implications of the AUC differences in ROC prediction curves, illustrated by the ESV vs ES shape comparison. Given a certain sensitivity and specificity operating point of the ESV curve, i.e. point A, the ES shape model is able to predict MACE at the same sensitivity but reducing the false positive rate a 50% (point B), i.e., from 275 false positive MACE predictions to 133. Likewise, the ES shape model could operate at a similar false positive rate to the ESV model while improving the sensitivity by 29% (point C). Alternatively, the ES shape operating point D could be chosen for MACE prediction, with both sensitivity and specificity improved with respect to the ESV model. This demonstrates the potential for risk management improvement associated to superior AUC scores.

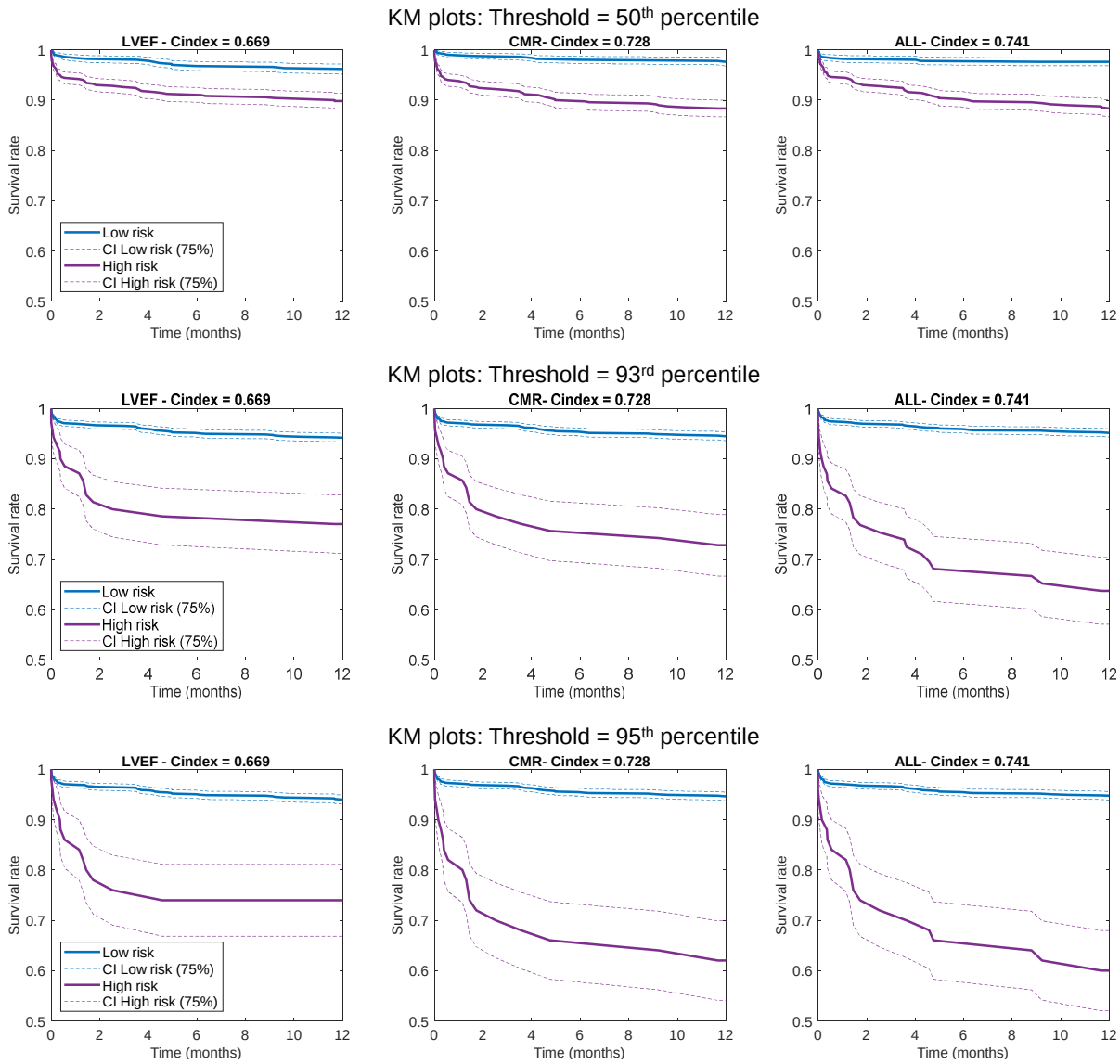


Fig. SIII.5 – Kaplan–Meier (KM) curve estimates for the high (purple) and low (blue) risk groups according to the best performing clinical baseline model, LVEF (**left**), and our proposed multivariable models, CMR (**middle**) and ALL (**right**). The KM plots are calculated as follows: (1) once trained, the Cox model is used to predict the patient scores; (2) the patients are then sorted by score, and the high and low risk groups are separated based on a threshold (a percentile); and, finally, (3) the KM curves, along with the confidence intervals, are estimated for each group of patients based on their initial survival data. This means that the KM curves illustrate survival at a single operating point, determined by the threshold selection. To illustrate this, 3 operating thresholds are considered: median percentile (**top**), 93rd percentile (**middle** – the operating point of a theoretical classifier that matches the 7% MACE incidence) and 95th percentile (**bottom** – operating point to increase sensitivity). Given our MACE incidence, the high and low risk KM curves are necessarily similar when operating at the median percentile, since survival is the main trend in both groups. As we move towards higher percentiles, the differences between high/low groups increase. The CMR model performs only slightly better than LVEF at 93rd percentile, but it is as good as the ALL model at 95th percentile. The assessment is biased to the threshold selection. This is why we prefer to establish comparisons based on C-indices (herein reported) and the ROC curves (Fig. SIII.1), which illustrate performance at any of the sensitivity-specificity operating points.

SUPPLEMENTAL MATERIAL IV – Independent Testing

Given the limited number of MACE events (class imbalance), the split of the cohort into fixed training and testing datasets would affect the statistical strength of the results. Cross-fold validation is therefore preferred and argued in the TRIPOD and PROBAST guidelines to be more robust than the traditional training-testing split, ensuring fair performance and conclusions reporting (17,18). In the *k*-fold cross-validation, the method deployed in this work, the data is randomly partitioned into *k* roughly equal subsets (i.e. folds). Iteratively, each of the *k* subsets is retained (completely unseen) for testing the model, and the remaining *k*-1 subsets are used as training data. The *k* results are then averaged to produce a performance estimation, which reduces noise variability. In addition, the *k*-fold cross-validation process can be repeated for a number of random data splits (e.g. a hundred) to estimate the variance of the model performance.

For completeness, the experiments are repeated, and the performance is evaluated considering an independent testing cohort. The population is split into training-testing based on the two independent randomized trials that compose the study (see Methods). Thus, the developed models are trained from scratch with patients from the AIDA-STEMI trial (n=723) and are evaluated on the TATORT-NSTEMI independent cohort (n=298). Note that the predictor selection is not repeated since this experiment is meant to assess generality and robustness of deployed models and not to develop new ones. The resubstitution and cross-fold validated AUC scores resulting from the training set, along with the final performance on the independent testing set, are illustrated in Fig. SIV.1. The performance trends reported in the main manuscript (ES shape>ESV; 3D contraction>LVEF; Multivariable models>>Clinical baselines) hold.

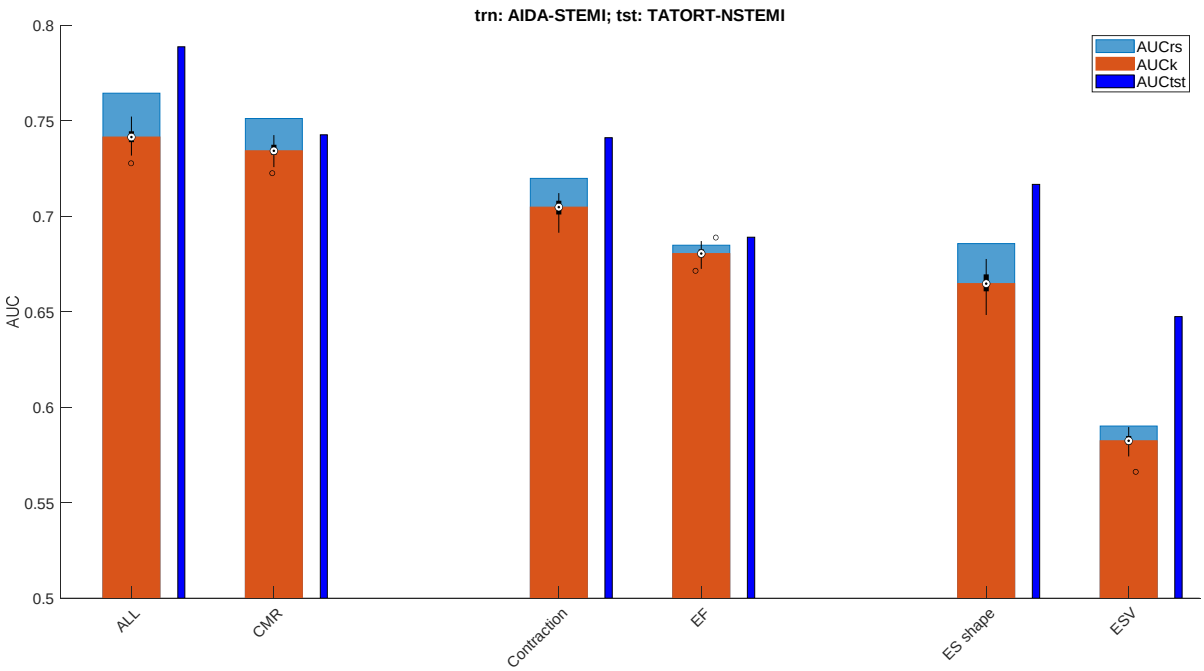


Fig. SIV.1 –MACE prediction results of the clinical baselines, 3D descriptors and multivariable models, including only CMR biomarkers (CMR) and all the variables of the study (ALL), on the independent testing validation. Patients are split into training (trn, n=723) and testing (tst, n=298) according to the two randomized trials, AIDA-STEMI and TATORT-NSTEMI. The AUC re-substitution (AUCrs, blue) and AUC 10-cross-fold validated (AUCk, orange), repeated for a hundred random data splits (dark distributions) resulting from the training set are reported. Likewise, the final testing AUC (AUCtst, purple), is illustrated.

While these results strengthen the generality of the findings, it can be argued that this experiment is not strictly equivalent to considering a new additional cohort since the TATORT-NSTEMI patients were indirectly used in the predictor selection. The size of the study, the intrinsic design of the methods and the reproducibility analysis further argue for the robustness of the findings (see Discussion).

SUPPLEMENTAL MATERIAL V – Subgroup Analysis

The proposed methodology has superior prognostic value for patients with decreased LV function (LVEF<35%), identified as high-risk group, in accordance to the literature (3) (see Fig. SV.1). With respect to infarct aetiology, the method is more predictive for the NSTEMI population (AUC of 0.79 vs 0.76, respectively, applying 'ALL' multivariable model – see Methods) despite the study being biased to STEMI patients (see Results). While mechanistic explanations are to be further explored, clinical validity for both types can be argued. This is also evidenced in the independent testing experiment, where models trained on only STEMI patients are suitable to NSTEMI ones (see Supplemental Material IV). Additional investigations show that ES5 and ES6 are significantly higher for the STEMI population, i.e. thinner ES shapes (as result of larger infarct size, P<0.001, in agreement with (5)) and higher anterior curvatures (LAD infarct (9)), respectively. Finally, while C16 is not relevant to infarct type, it is significantly higher for the NSTEMI MACE subgroups, which in turn may explain its predictive power.

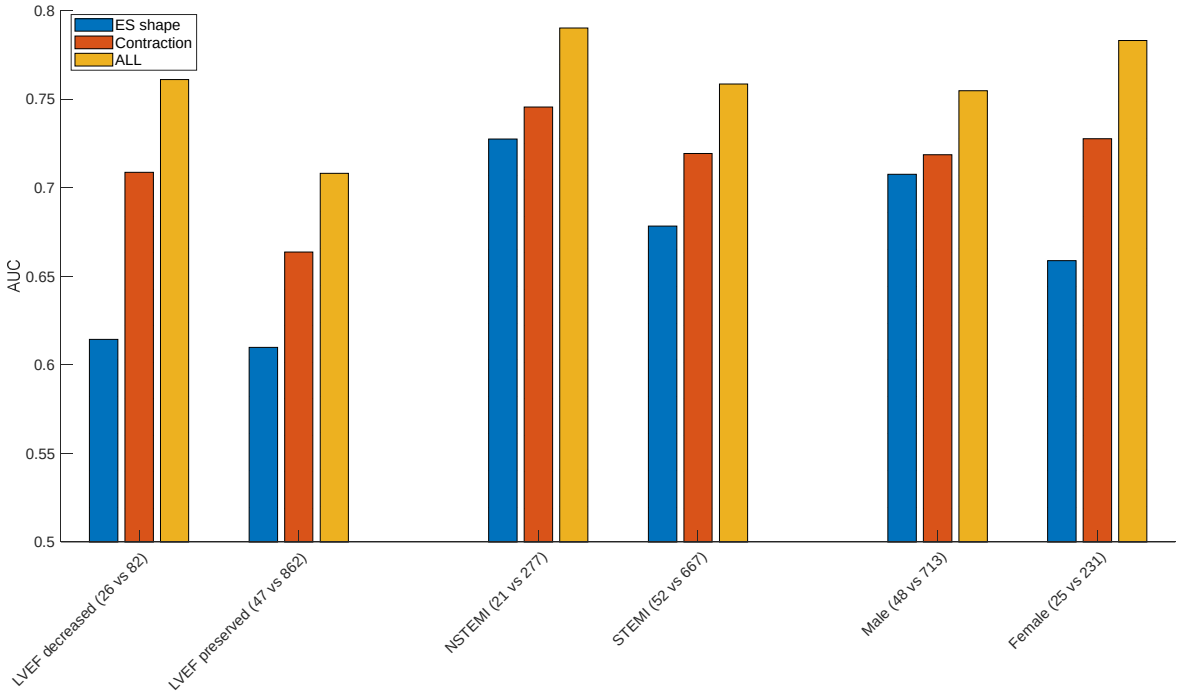


Fig. SV.1 – AUC endpoints prediction results stratifying by gender, infarct aetiology and LVEF (threshold: 0.35); and applying the 'ES shape', '3D contraction' and 'ALL' LDA models resulting from the backward stepwise analysis (See Table 2 and Fig. SIII.2). The labels are presented as subgroup (MACE vs No MACE cases).

SUPPLEMENTAL MATERIAL VI – Reproducibility Analysis

In the absence of a patient rescanning protocol in the MRI study, we assess the reproducibility of the method by removing the ED and ES frames of the patients' CMR scans. We thus find an alternative pair of ED and ES frames, those with the second largest and smallest volumes in the original sequence, respectively. The different frames selected emulate a new acquisition, with the same orientation and patient status, but with the bias caused by the removal of the true ES and ED frames.

The volumes and meshes at ES and the 3D contractions (i.e. ES shape – ED shape) resulting from this reproducibility analysis are projected using the original statistical PCA and LDA models, finding the second observation of the proposed biomarkers (ES1, ES3, ES6, C3, C5, C16). Both conventional (EDV, ESV and EF) and proposed biomarkers are regressed to the original observation (See Fig. SVI.1), and the predictive power of the second observation is assessed applying the models previously deployed based on the originals (See Fig. SVI.2) without any further training.

The results show a good robustness of the method in the above described drastic scenario. ESV, EDV and ES shape are calculated from a single frame; as such, their reproducibility is superior to LVEF and 3D contraction, which relay on two segmentations. While the robustness of ES shape ($R = 0.980$) is comparable to ESV and EDV ($R = 0.995$ and $R = 0.994$, respectively), the reproducibility of 3D contraction ($R = 0.864$) decreases with respect to LVEF ($R = 0.967$). This could be explained since 3D contraction encodes mode 16, and reproducibility decreases with the higher order modes as one would expect. That is, the higher level of detail we capture, the larger the impact of noise.

Even though the proposed variables that encode contraction variations are more sensitive to noise than LVEF or ESV, their MACE predictive performance is barely compromised. The novel 3D descriptors remain significantly superior, both considering stand-alone or multivariable models (See Fig. SVI.2), and by considering the combined score with all or with the variables that have been reproduced (i.e. partial in Fig. SVI.1). Finally, it is worth highlighting the extraordinary reproducibility of the automated volumes ($R > 0.99$).

The intraclass correlation coefficients (one-way random, single measures) and the Spearman correlation coefficients results, shown in Table SVI.1, further support the above discussed. Coefficients of variation (within-subject standard deviation method) are also reported for ESV, EDV and LVEF.

Table SVI.1 – Reproducibility analysis results

DESCRIPTION	DIFFERENCE	R	r_s	CoV (%)	ICC
ESV, mL	0.227 (3.324)	0.995	0.993	2.824	0.995
EDV, mL	0.168 (4.654)	0.994	0.992	2.075	0.994
EF	0.000 (0.027)	0.967	0.967	3.949	0.967
C3	-	0.986	0.985	-	0.986
C5	-	0.942	0.943	-	0.941
C16	-	0.973	0.969	-	0.973
ES1	-	0.892	0.869	-	0.891
ES5	-	0.840	0.829	-	0.837
ES6	-	0.832	0.830	-	0.829
ES shape	-	0.980	0.977	-	0.980
Contraction	-	0.864	0.849	-	0.862
Multivariable ALL	-	0.952	0.936	-	0.952
Multivariable ALL (partial)	-	0.910	0.895	-	0.910

Reproducibility analysis results reported as: difference, expressed as mean error (SD); linear correlation coefficient, R; Spearman correlation coefficient, r_s; coefficient of variation, CoV; and intraclass correlation coefficient, ICC. Differences and CoV are not reported for the LDA variables (normalized) and the modes of variation (centered at zero by definition). Multivariable ALL indicates scores of the resulting LDA model including all the variables of the study on top of the proposed 3D descriptors; Multivariable ALL (partial), partial scores of the Multivariable ALL model excluding the invariant markers in the reproducibility analysis (i.e. Age, etc.).

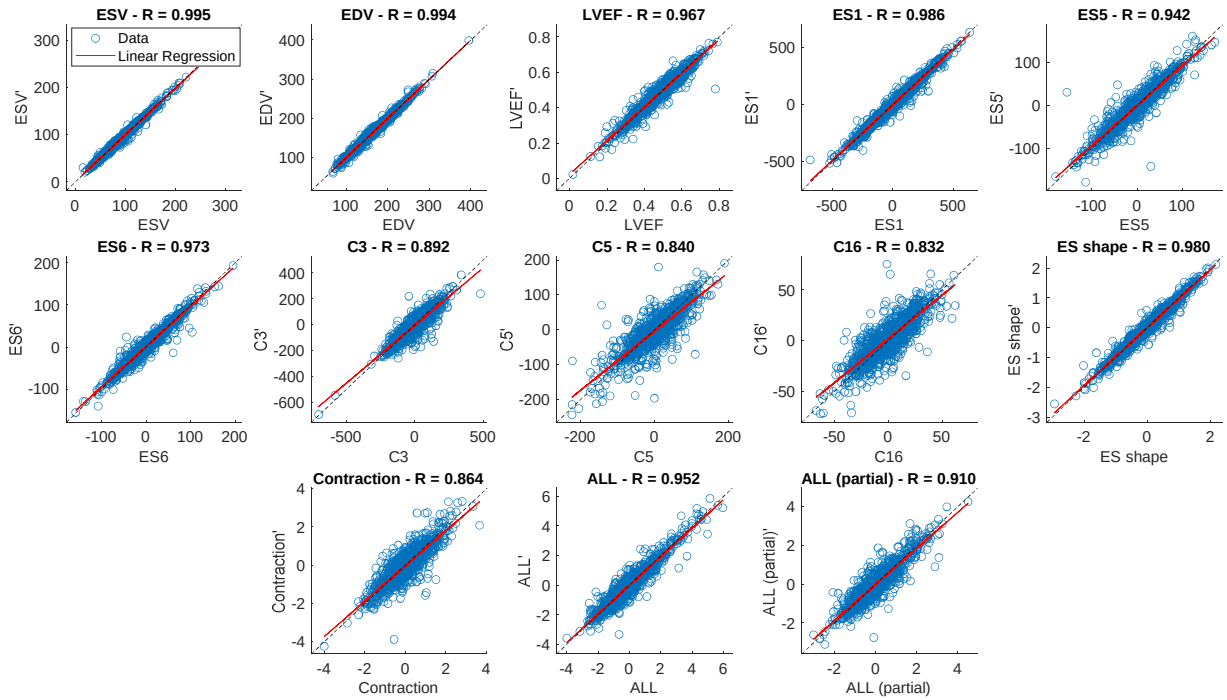


Fig. SVI.1 – Linear correlation results of the reproducibility analysis, expressed as R score. 'x' axis corresponds to originals; 'y', reproduced. Legend reads the same as in Table SV.1.

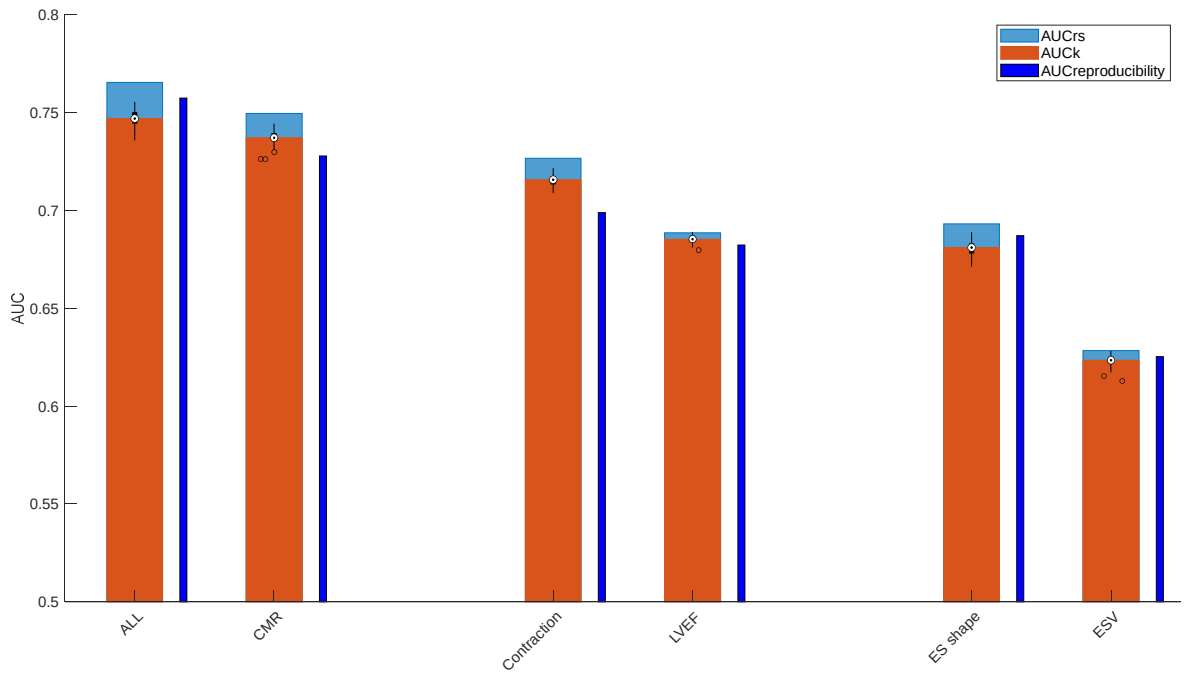


Fig. SVI.2 – Predictive performance of ESV, LVEF, ES shape, contraction and multivariable LDA models resulting from the combination of the 3D descriptors with only CMR biomarkers (CMR) and all the variables of the study (ALL). Results are expressed as AUC re-substitution (AUCrs - light blue), AUC 10-fold cross-validated computed for 100 random data splits (AUCk - mean in orange, distribution in black) and AUC resulting from the reproducibility analysis (AUCreproducibility - dark blue). From any of the AUC metrics it can be concluded that the proposed 3D disentanglement of ESV and LVEF not only outperforms the stand-alone versions, but it also contributes to a significant overall risk management improvement in a multivariable setting including CMR markers, cardiovascular risk factors and basic patient characteristics.

SUPPLEMENTAL MATERIAL VII – Modes Correlation Analysis

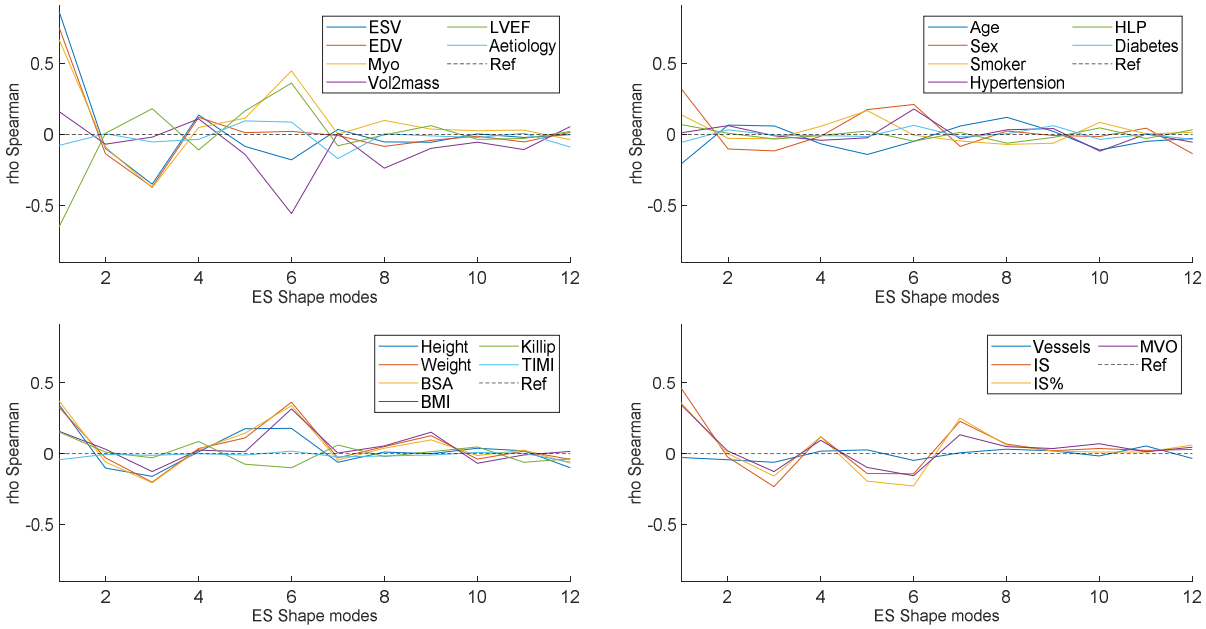


Fig. SVII.1 – Spearman correlation between the first 12 ES shape modes of variation and all the variables included in the study. Myo indicates LV myocardial mass; Vol2mass, EDV to LV mass ratio; Aetiology, infarct aetiology (STEMI vs NSTEMI); HLP, Hyperlipoproteinemia; BSA, body surface area; BMI, body mass index; Killip, Killip class on admission; TIMI, TIMI flow grade post-PCI; Vessels, Nr. of diseased vessels; IS, infarct size; IS% infarct size to LV mass ratio; MVO, microvascular obstruction. Ref is the reference line of zero correlation.

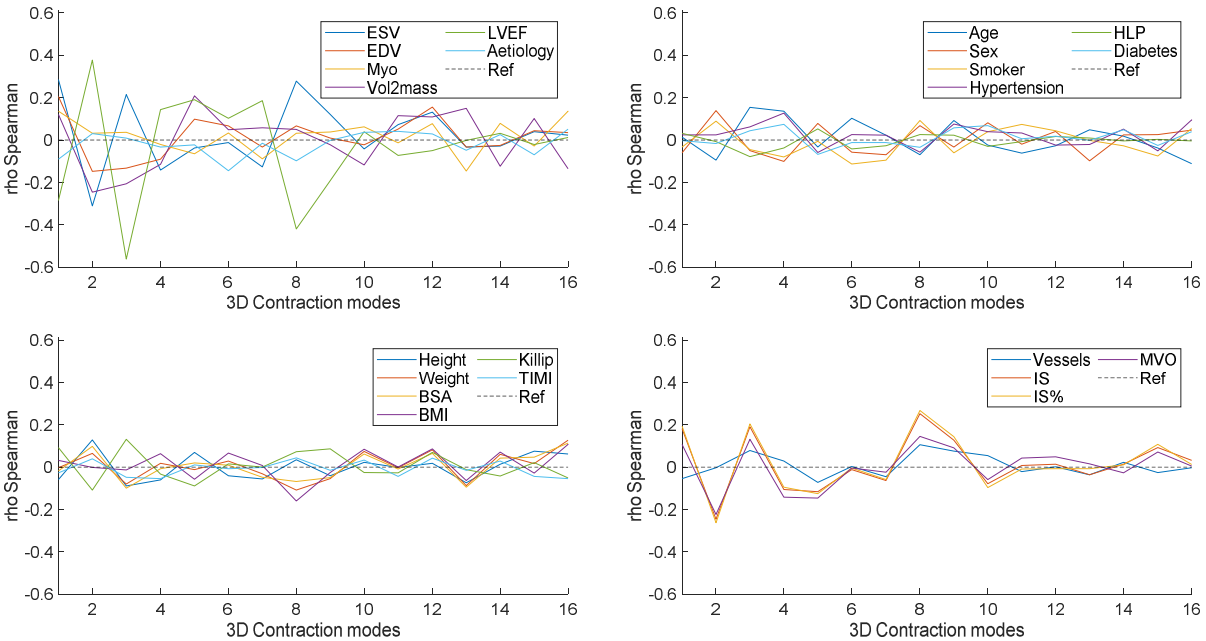


Fig. SVII.2 – Spearman correlation between the first 16 3D contraction modes of variation and all the variables included in the study. Legend reads the same as in Fig. SVII.1 above.

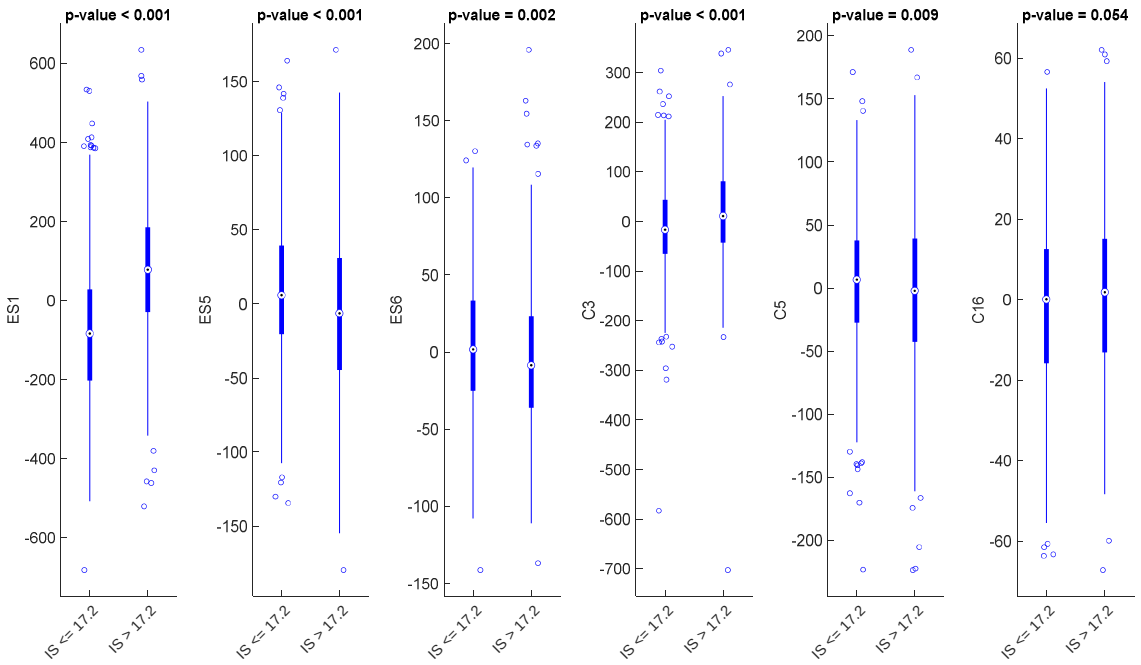


Fig. SVII.3 – Distribution and significance of the identified risk-related shape and contraction variations (ES1, ES5, ES6, C3, C5 and C16) in stratifying **IS** (infarct size) into low/high myocardial damage using the median value of the AMI population as threshold (17.2 mL).

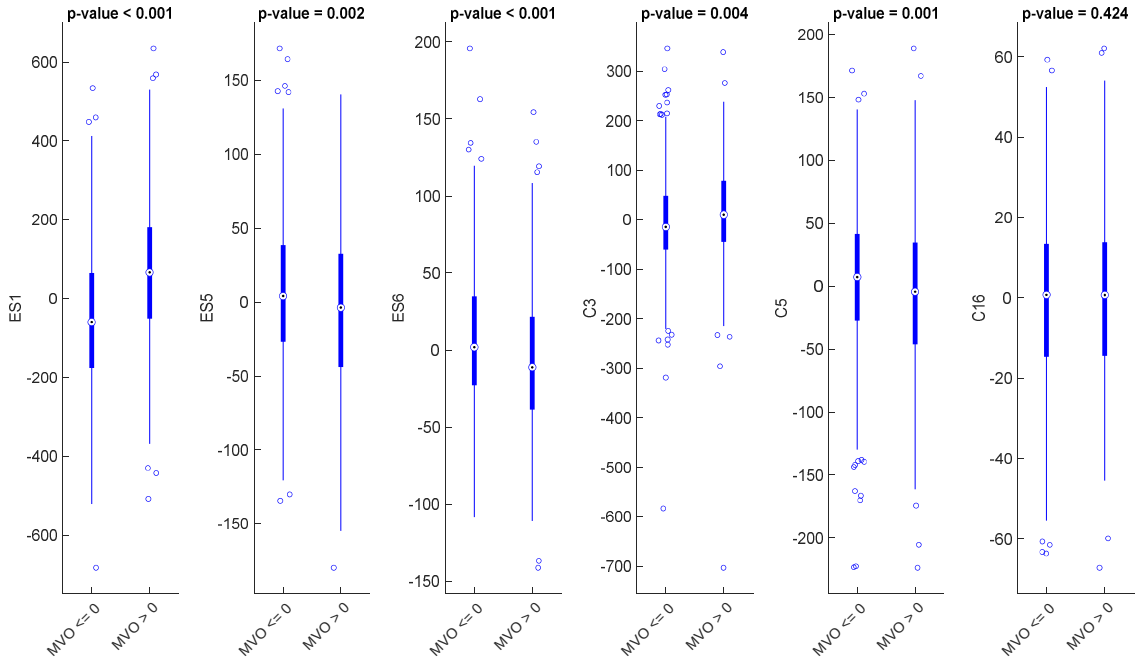


Fig. SVII.4 – Distribution and significance of the identified risk-related shape and contraction variations (ES1, ES5, ES6, C3, C5 and C16) in stratifying **MVO** (microvascular obstruction) into low/high damage using the median value of the AMI population as threshold (0 mL, that is, intact microvasculature).

SUPPLEMENTAL MATERIAL VIII – TRIPOD Reporting

Table SVIII.1 – TRIPOD Checklist

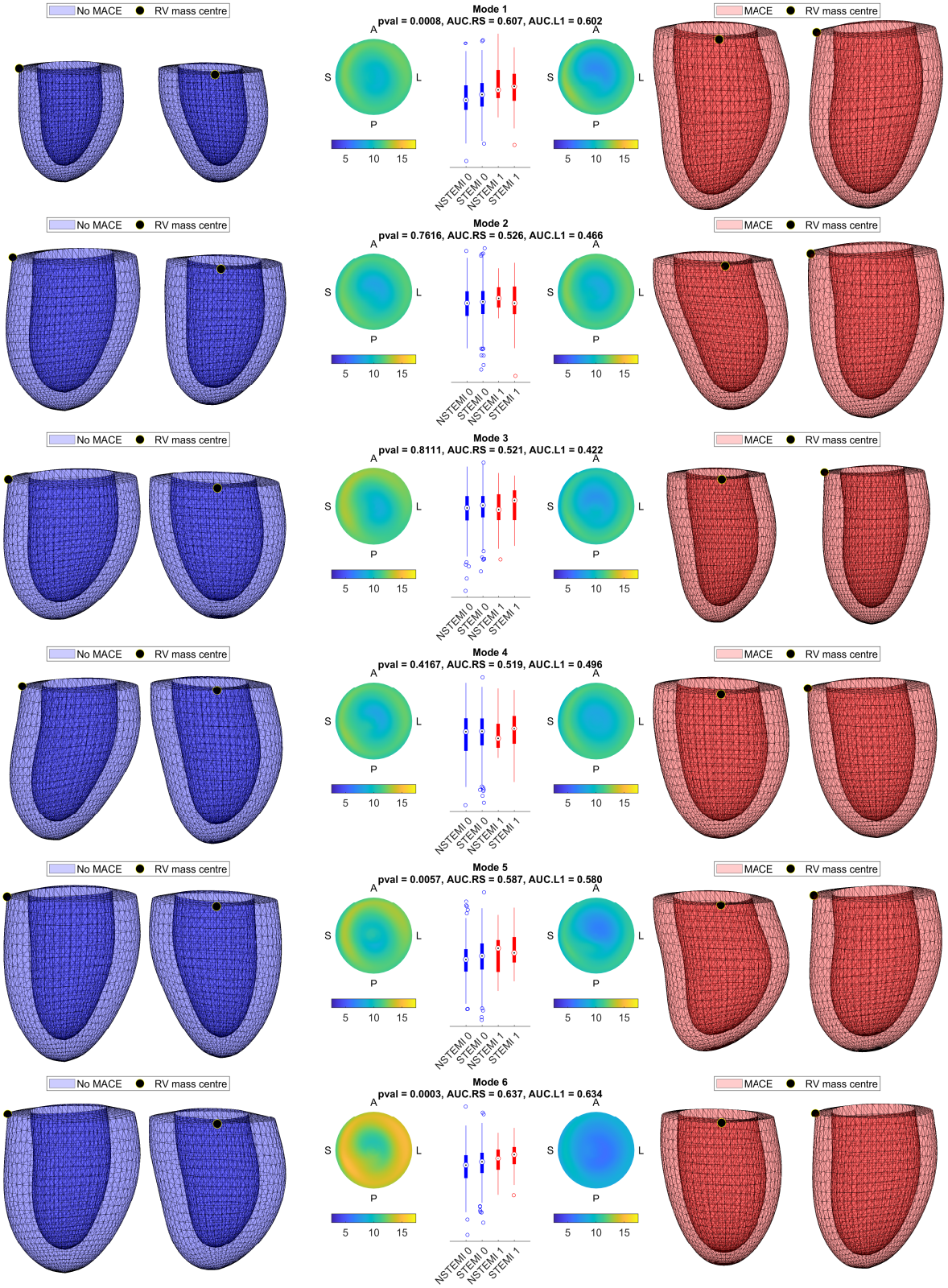
DESCRIPTION	ITEM	CHECKLIST	MANUSCRIPT CHECK
Title & abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted	Title
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions	Abstract
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models	Introduction (par.1-3)
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both	Introduction (par 3)
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable	Methods (Study population)
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up	Methods (Study population and study endpoints)
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centers	Methods (Study population)
	5b	Describe eligibility criteria for participants	
	5c	Give details of treatments received, if relevant	
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed	Methods (Study endpoints)
	6b	Report any actions to blind assessment of the outcome to be predicted	
Predictors	7a	Clearly define all predictors used in developing the multivariable prediction model, including how and when they were measured	Methods (Evaluation experiments and statistical analysis), Table 1
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors	
Sample size	8	Explain how the study size was arrived at	*
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method	Fig.1 and Results (Patients)
Statistical analysis methods	10a	Describe how predictors were handled in the analyses	Methods (Experiments, supervised learning and statistical analysis)
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation	
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models	
Risk groups	11	Provide details on how risk groups were created, if done	Methods (Study endpoints)
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful	Table 1, Fig.1 and Results (Patients)
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome	
Model development	14a	Specify the number of participants and outcome events in each analysis	Table 1, Fig.1 and Results (Patients)
	14b	If done, report the unadjusted association between each candidate predictor and outcome	
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point)	Results and Supplemental Material III
	15b	Explain how to use the prediction model	
Performance	16	Report performance measures (with CIs) for the prediction model	Results and Supplemental Material III

Continued

Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data)	Discussion (limitation)
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence	Discussion
Implications	20	Discuss the potential clinical use of the model and implications for future research	Discussion (towards clinical translation)
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets	Supplemental Material
Funding	22	Give the source of funding and the role of the funders for the present study	Funding

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) checklist, adapted from (17). Only model development items are included. (*) The size of the study is usually determined based on the variables of interest and their standard deviation. The proposed study cannot benefit from this criterion since the variables (modes of variation) were not defined *a priori*. However, aiming at a margin of error of ≤0.05 in the overall outcome proportion estimate, a mean absolute prediction error of 0.05, a desired shrinkage ≤10% and conservative anticipated Cox-Snell R squared statistic of 0.2, and given the approximate MACE incidence of 7% and the number of predictors indicated in the manuscript, the actual size of the study is superior to the trial target size retrospectively determined by any of the 4 proposed calculation methods explained in Riley et al. (13). Par. indicates paragraph.

SUPPLEMENTAL MATERIAL IX – ES Shape Modes



Continued in next page

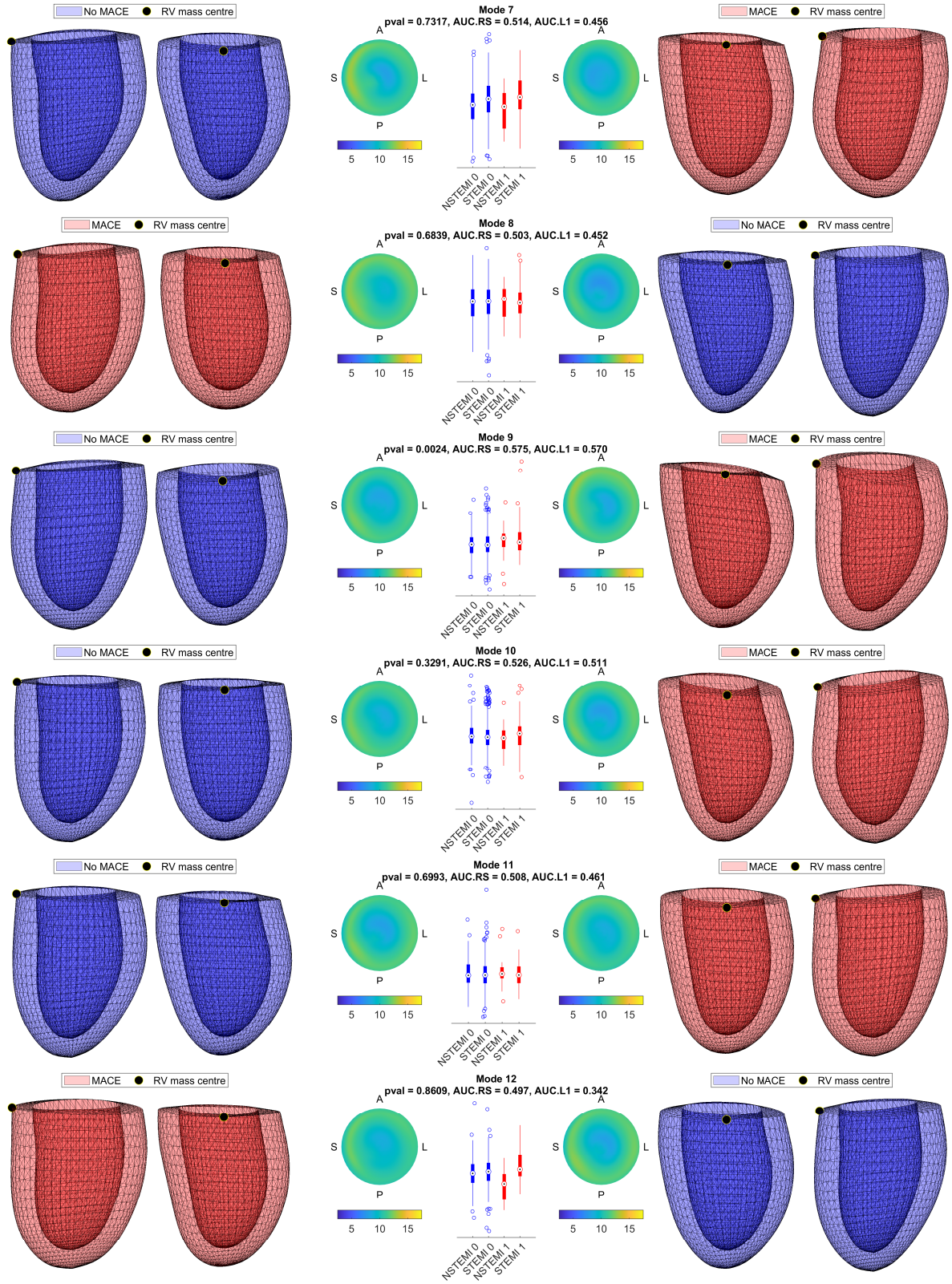
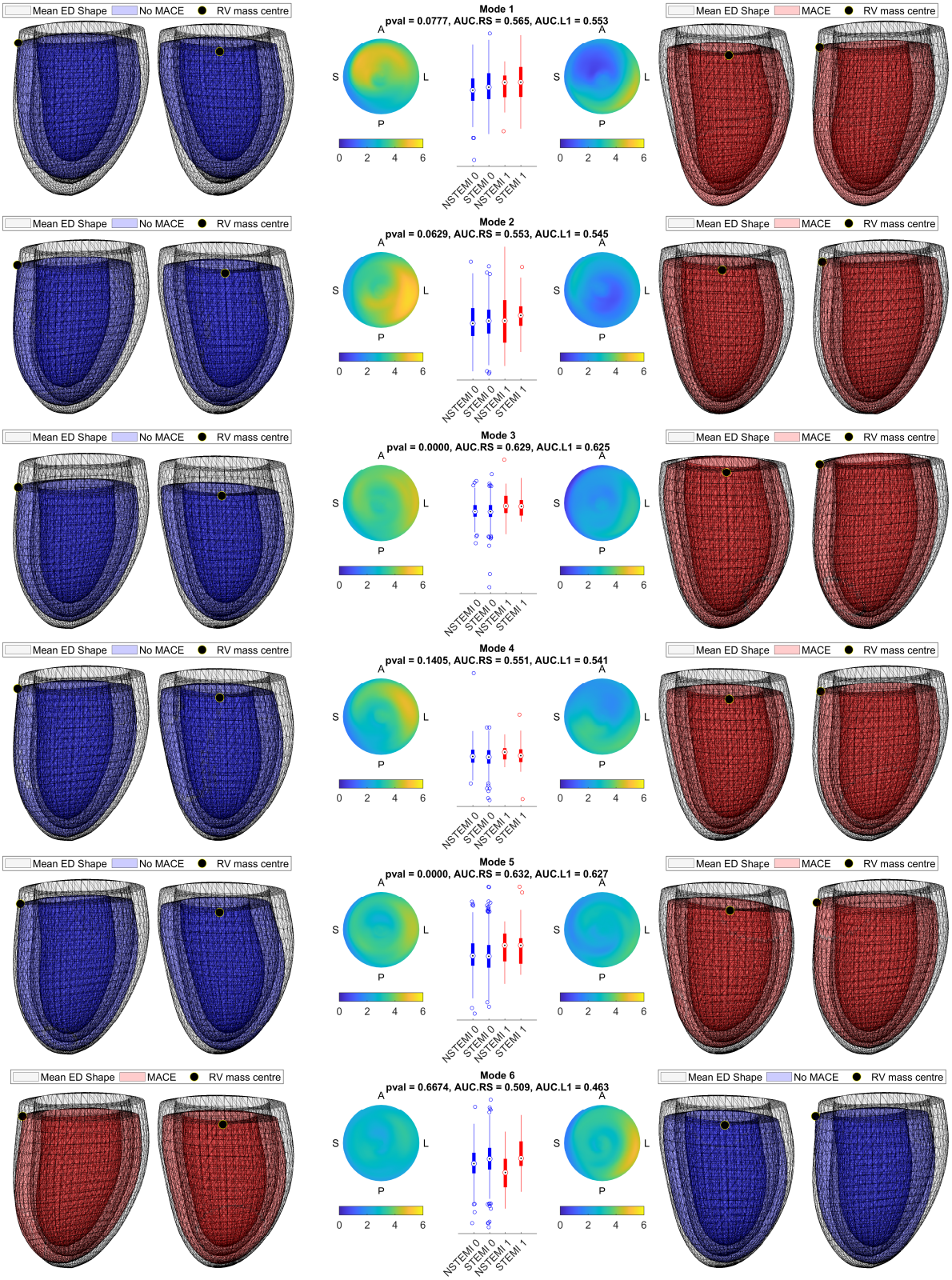
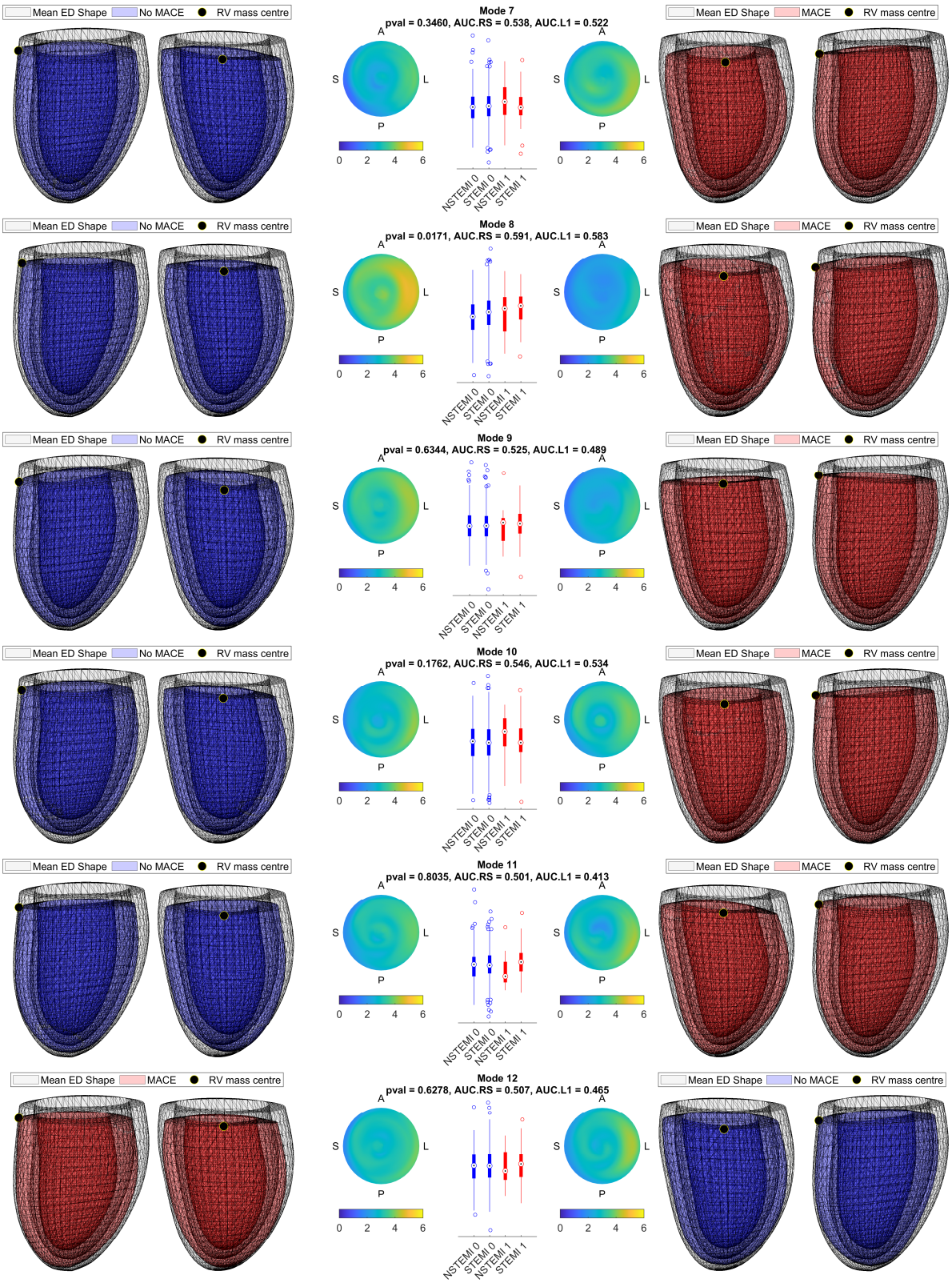


Fig. SIX.1 – ES shape main modes of variation accounting for 95% of the population variance. The meshes are presented in anterior and septal views and in the form of thickness maps on a circumferential polar plot as in the 17-segments AHA model. The MACE (red, class 1) and No MACE (blue, class 0) representations correspond to the 5th and 95th percentiles in the mode direction. *P* value and re-substitution and leave-one-out AUCs are presented along with the MACE and no MACE distributions, further stratified into infarct aetiology (STEMI and NSTEMI).

SUPPLEMENTAL MATERIAL X – 3D Contraction Modes



Continued in next page



Continued in next page

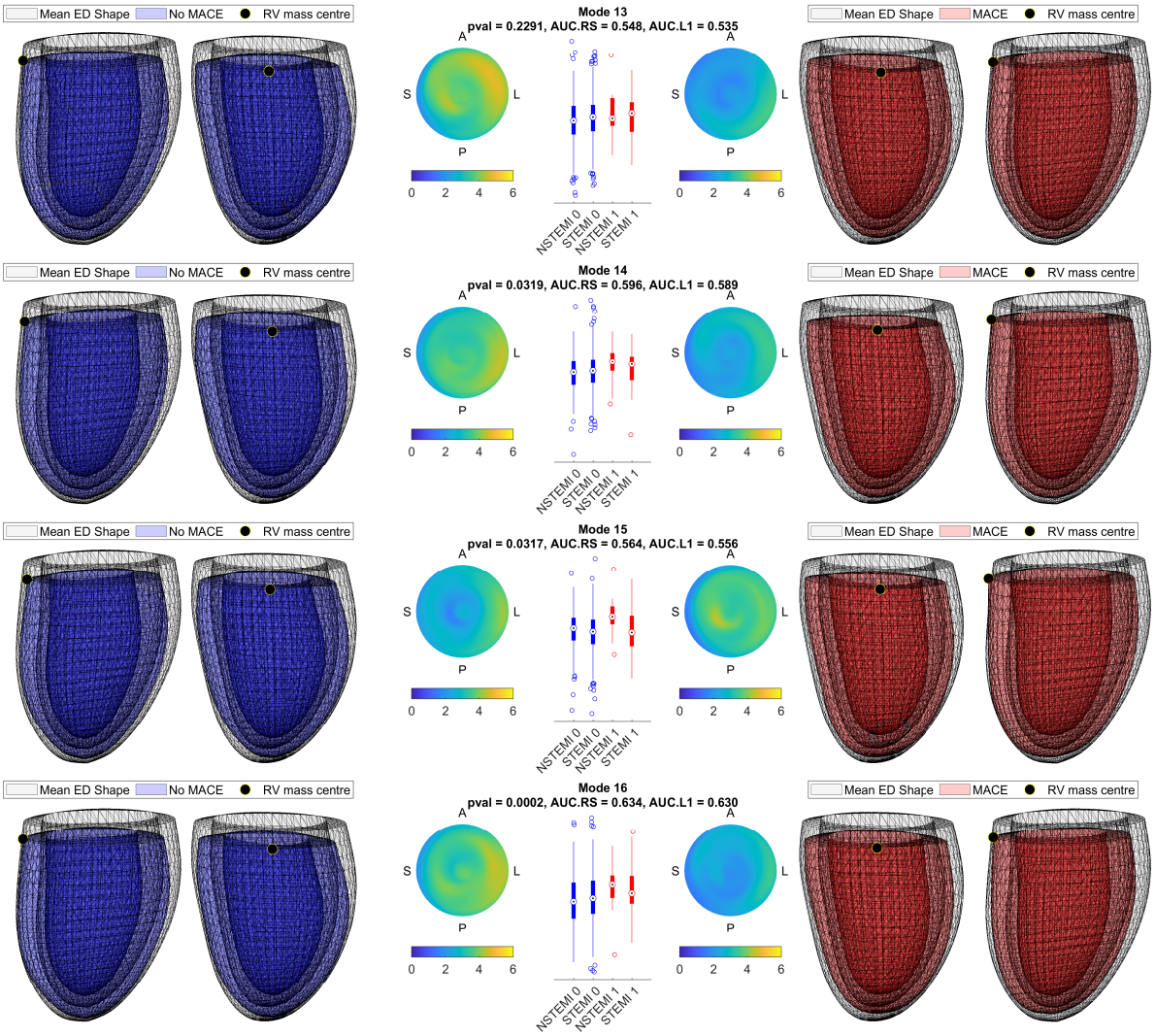


Fig. SX.1 – LV contraction modes 1 to 16, accounting for 93.2% of the population variance. The meshes are presented in anterior and septal views and in the form of differential thickness maps (ED minus ES thickness) on a circumferential polar plot as in the AHA model. To facilitate comparisons, the contractions are applied on the mean ED shape and visualized as resulting ES shapes. The MACE (red, class 1) and No MACE (blue, class 0) representations correspond to the 5th and 95th percentiles in the mode direction. *P* value and re-substitution and leave-one-out AUCs are presented along with the MACE and no MACE distributions, further stratified into infarct aetiology (STEMI and NSTEMI). The mean ED shape (transparent surface) is included as reference.

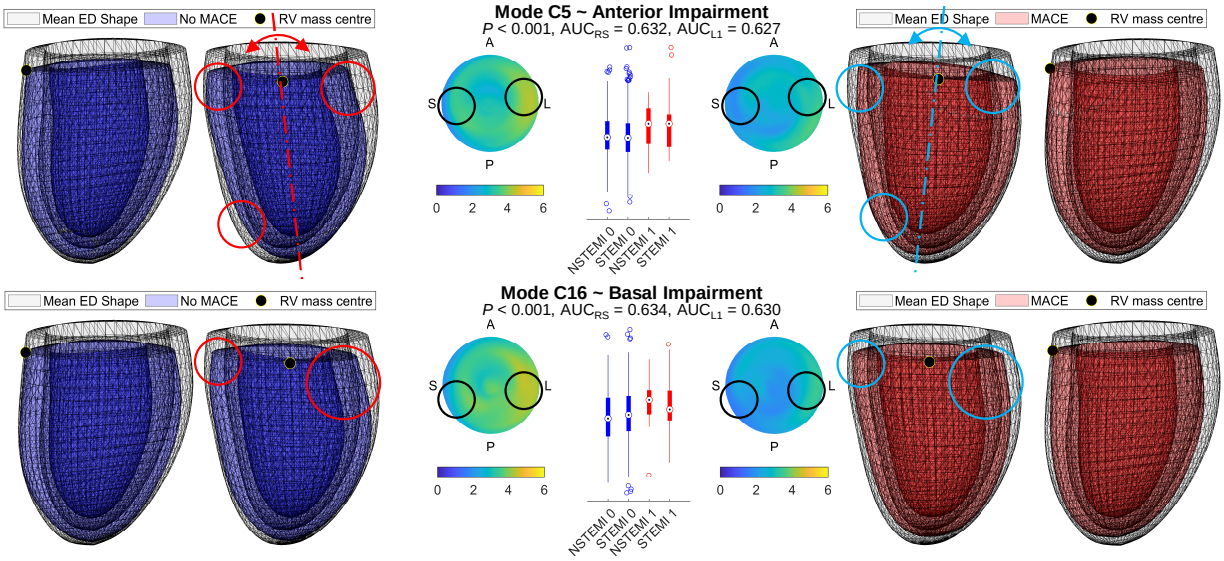


Fig. SX.2 – Repetition of the contraction modes 5 and 16, shown in the previous figure and proven relevant to AMI prognosis (see main manuscript), highlighting the differences between the MACE and No MACE contractions. While C5 shows an anterior impairment, the main difference in C16 is the basal impairment between the 2 configurations, particularly visible in the anterior wall, but also affecting the posterior one. A difference in contraction thickening of 4 vs 2 mm is also observed in the lateral and septal walls as a secondary change.

Appendix B:

Supplemental Material Paper VI

SUPPLEMENTAL MATERIAL

Comprehensive Characterisation of Cardiac Function for Improved Post-Infarction Risk Assessment

Jorge Corral Acero, Pablo Lamata, Ingo Eitel, Ernesto Zacur, Ruben Evertz, Torben Lange, Sören J. Backhaus, Thomas Stiermaier, Holger Thiele, Alfonso Bueno-Orovio, Andreas Schuster, Vicente Grau.

Table of Contents

SUPPLEMENTAL MATERIAL I – LV Segmentation 2
SUPPLEMENTAL MATERIAL II – Volume Transient Normalisation 3
SUPPLEMENTAL MATERIAL III – Dimensionality Reduction 4
SUPPLEMENTAL MATERIAL IV – Additional Patient Characteristics 7
SUPPLEMENTAL MATERIAL V – Endpoint Prediction Additional Results 10
SUPPLEMENTAL MATERIAL VI – Modes Correlation Additional Results 13
SUPPLEMENTAL MATERIAL VII – Systolic vs Diastolic Components 14
SUPPLEMENTAL MATERIAL VIII – TRIPOD reporting 16

SUPPLEMENTAL MATERIAL I – LV Segmentation

A 2-step state-of-art approach was applied to segment the LV endo- and epicardium, from which the LV cavity and myocardium can be estimated (28,29):

- In the first step, pre-processing, the images are re-oriented, cropped and normalised. Considering the mid-ventricular SAx slice as reference, the first neural network (NN) detects the position of the heart and defines a region of interest (ROI) of 139.7x139.7 mm cantered in the LV. Based on these LV centroid and ROI, the LV is aligned to a canonical position by a rigid registration of the same SAx reference slice to the atlas built in (29). The SAx is then cropped accordingly, and the intensities are normalised.
- The second step, fine segmentation, applies another NN to the pre-processed images to regress the enhanced LV segmentation. A final postprocessing is applied to binarized the segmentation predictions and improve the segmentation quality.

A cohort of 100 patients of the study, manually segmented with random resolutions and endpoints, was used to train the 2 NNs, following a 5-fold cross-validation strategy and using the same training-validation-testing split ratio as in (29). Translation, rotation and flipping were used for augmentation. Architectures and implementation are detailed in (28,29). Segmentation performance assessment is based on endocardium and epicardium gold standard Dice scores. This metric accounts for the overlap between manual segmentation and automated prediction and varies between 0 and 1, with 1 corresponding to a perfect match. Fig. SI.1 provides a visual sample of the segmentation results.

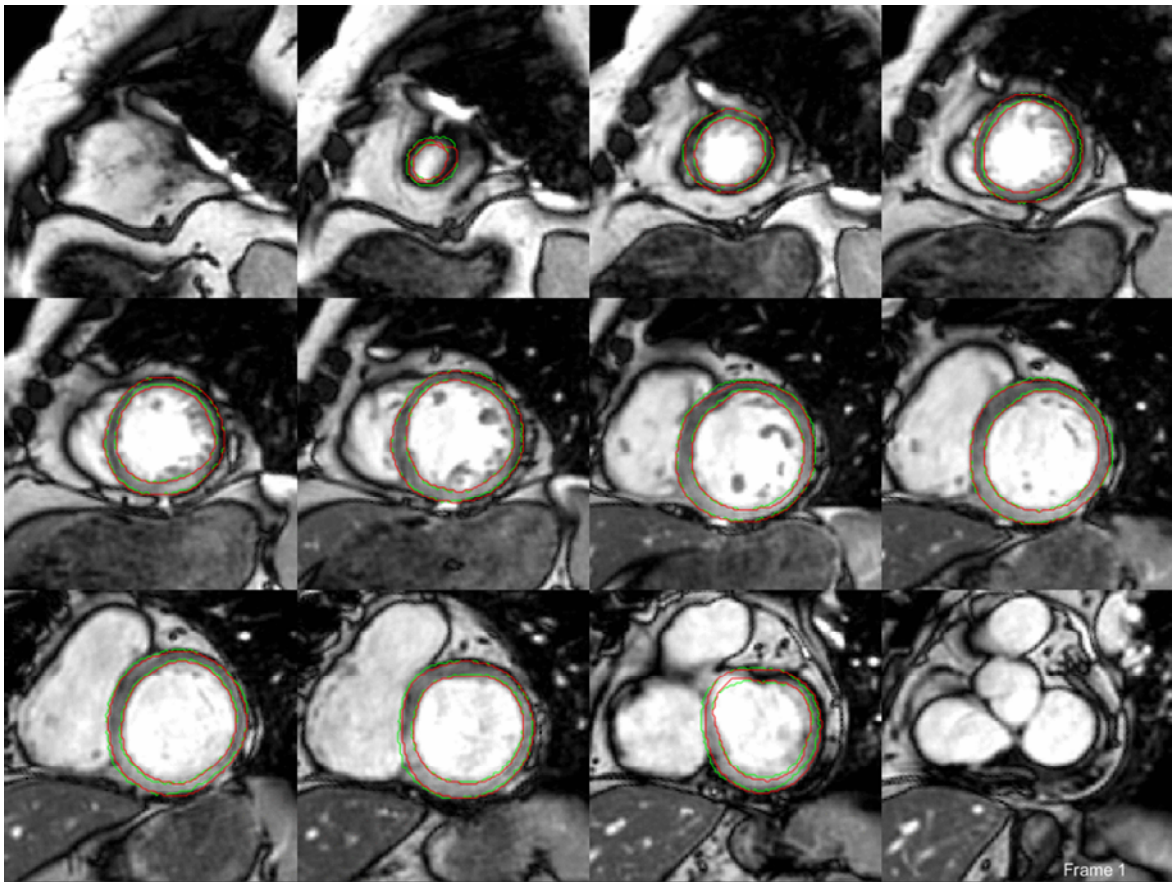


Fig. SI.1 – Segmentation results of a representative patient (median Dice) at ED. The green contours correspond to the LV reference segmentation (manual segmentation); and the red contours, to the prediction results based on our proposed 2-step deep learning approach.

SUPPLEMENTAL MATERIAL II – Volume Transient Normalisation

The LV volume temporal transients are described by the magnitude of the curve (i.e. maximum and minimum LV volumes) and by its shape (i.e. relative filling contributions, presence or absence of diastasis, etc.). While the latter rather involves subtle changes and its understanding is the aim of this study, the former, that contains information about ventricle size and amplitude of contraction or stroke volume, is a major source of variability and can be completely described by the well-established EDV, ESV and LVEF markers. This motivates the standardisation in magnitude, to remove cofounding factors, towards facilitating multivariate models, and variability noise, so that the PCA analysis concentrates on these subtle transient shape changes. Thus, a standard minmax normalisation is proposed:

$$\bar{V}(t) = \frac{V(t) - ESV}{SV},$$

where $V(t)$ represents the original LV volume transient (mL/s); $\bar{V}(t)$, the normalised LV volume transient (%/s); ESV, the end-systolic volume (mL); and SV, the stroke volume (mL), calculated as end-systolic minus end-diastolic volume.

Fig. SII.1 illustrates two LV volume temporal transients from patients of a similar ventricle size before and after the proposed normalisation. The intra-relationships within a transient are preserved, that is, the active versus passive contribution balance or the systolic and diastolic velocities ratio, for instance, are constant before and after normalisation. However, the volumes are no longer absolute values but expressed as percentage of SV, enabling for direct inter-comparison between patients of a very different ventricle size or LVEF. This should be considered when drawing conclusions on the resulting transient patterns, presented in the main manuscript (see Fig. 2). Thus, a plausible conclusion from Fig. SII.1 is not that the passive filling is faster in patient 1 but that, considering the rest of the transient velocities, the passive filling is relatively faster in patient 1.

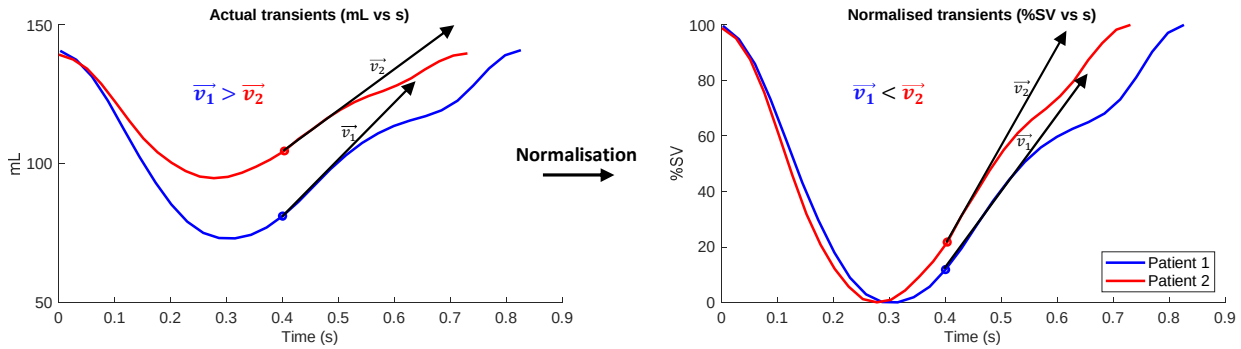


Fig. SII.1 – LV volume temporal transient normalisation for 2 patients. **LEFT**: LV volume temporal transients expressed in absolute volumes (mL/s). **RIGHT**: Normalised LV Volume temporal transients (%SV/s). The normalisation standardises for ventricle size and stroke volume, enabling for transient 'shape' inter-comparisons. In consequence, the passive filling peak velocity, considering its smaller SV and the rest of the transient velocities, is relatively faster for patient 2 ($\vec{v_2} > \vec{v_1}$), while, in absolute terms, this is the opposite ($\vec{v_1} > \vec{v_2}$).

SUPPLEMENTAL MATERIAL III – Dimensionality Reduction

The LV volume temporal transients are obtained by LV cavity volume integration in each frame of the cardiac cycle, as explained in the manuscript (see Methods). Therefore, the transients are described as a collection of 2D points, where the 'x' dimension corresponds to the trigger time of the frame, or time-point within the cardiac cycle, and the 'y' dimension represents the integrated cavity volume. Given the multicentre and multi-scanner nature of the study, patients are imaged at different resolutions, which results in 2D collection of points of different sizes (i.e. from 20 frames, and, in consequence, 20 pairs of time and volume 2D points, to 50 frames). Thus, a first step resamples the obtained volume transients, via splines, to ensure that all of them are described with the same number of points, that is, 30 points (evenly distributed in the 'x' or time dimension). This effectively means that the LV volume transient of each patient is described by 60 variables (30 points x 2 dimensions), which take a certain different value per patient. A second step normalises the 'y' or volume dimension, as explained in Supplemental Material II, to standardise in ventricle size and SV.

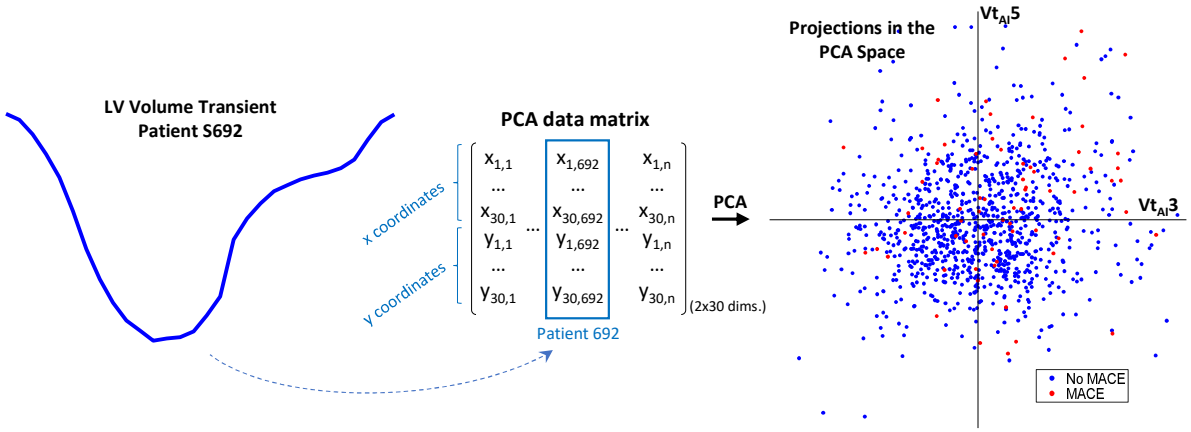


Fig. SIII.1 – **LEFT**: LV Volume transient of patient 692. **MIDDLE**: The 60 variables that describe the patient volume transient are sorted in columns, one per patient, to apply PCA. **RIGHT**: Projection of the patients in the 2 PCA directions that contributed the most to multivariate models (see Table 3) in order to maximise MACE (red) vs No MACE (blue) differences, that it, $Vt_{AI}3$ and $Vt_{AI}5$. We moved from 60 to just 2 transient variables per patient, which, additionally, are interpretable.

PCA, the dimensionality reduction technique applied in this work, is able to encode the information contained in these 60 variables (LV volume transient variability) into a few variables, the PCA modes, that describe the main LV volume transient variations. Mathematically, this is done by finding the vector space whose basis are the orthogonal directions that maximise the variance of the data, and subsequently projecting the data in this new space (See Fig. SIII.1).These directions that maximise the variance are, precisely, the PCA modes of variation that represent the way in which the LV volume transient varies in the population (i.e. RR-interval length, passive vs active filling relative contributions, etc.). As illustrated in Fig. SIII.2, if we move in the direction of a mode, we can see how the mean transient curve deforms in its particular way. This allows to describe each LV volume transient as a mean volume transient curve plus the variations encoded by each linear anatomical mode times the amount and direction of variation, or PCA coefficients (See Fig. SIII.4):

$$LV\ Volume\ transient = \varphi_0 + \sum_i Vt_{AI}i \cdot \varphi_i,$$

where φ_0 represents the mean volume transient, φ_i the anatomical PCA modes, and $Vt_{AI}i$ their respective PCA coefficients. Each of the modes, i, is therefore a continuous variable, that accounts for

a particular shape variation, which has a certain value for each patient, $Vt_{AI}i$, and whose MACE predictive power can be analysed. In other words, we can investigate which LV transient features or contraction patterns (modes) are related to AMI prognosis.

The modes are sorted in descending order of importance according to the amount of variability of the population that they explain (i.e. in our results 51.6% variability is described by mode 1; 18.4%, by mode 2; etc.). As we progressively incorporate modes, the LV transient reconstruction improves to the point that with only few modes we can accurately approximate any volume transient (i.e. 95.99% of LV transient variance described by first 6 modes), and hence the dimensionality reduction is achieved (See Fig. SIII.3).

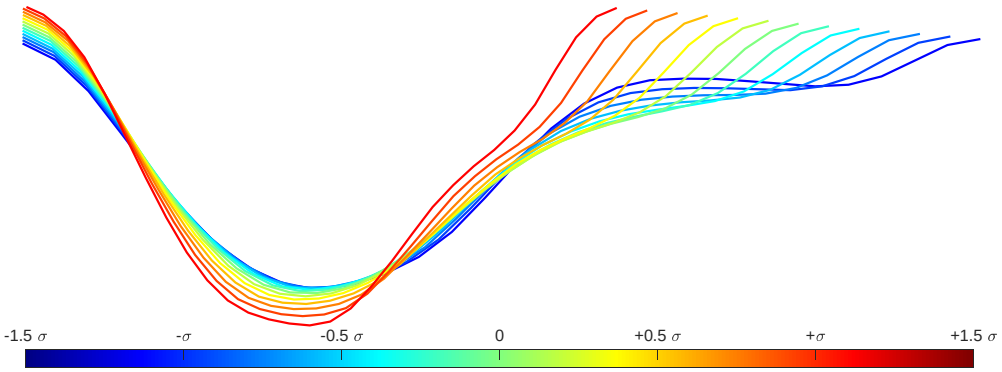


Fig. SIII.2 – The PCA direction of maximum variability corresponds to mode 1. In our results, the shape variation that it encodes is interpreted as mainly RR-interval length, although some other subsequent changes in transient morphology (i.e. diastasis) can be appreciated. As we move along the direction of mode 1, the RR-interval of the mean LV volume transient is reduced (positive $Vt_{AI}1$) or increased (negative $Vt_{AI}1$), proportionally to the value of its PCA coefficient $Vt_{AI}1$.

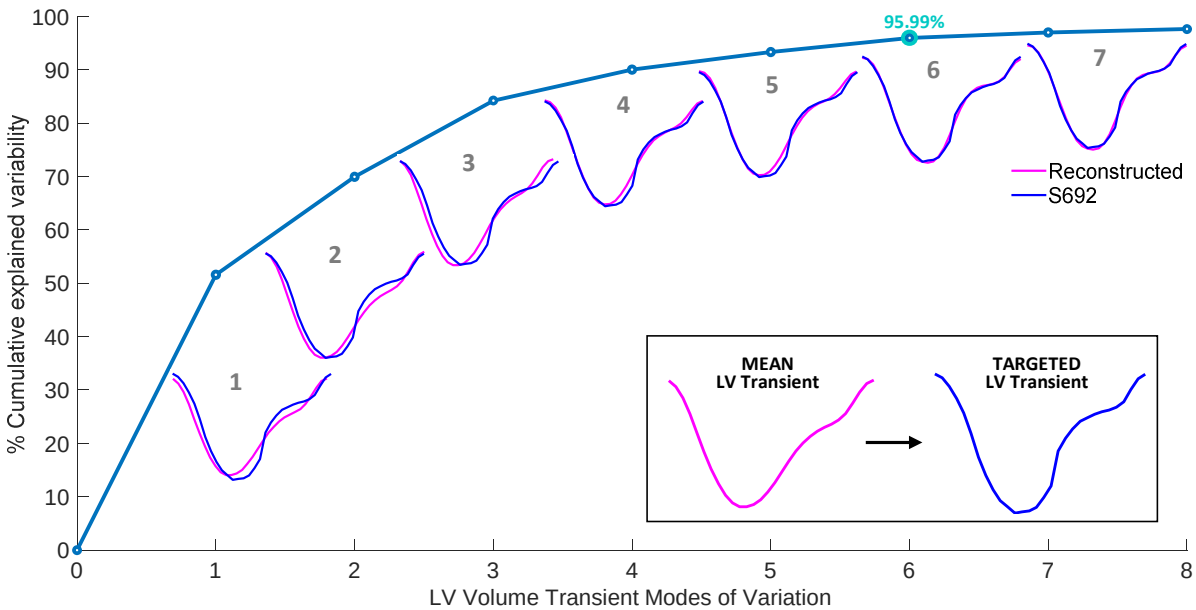


Fig. SIII.3 – Cumulative % of the population variance explained by the modes. As we incorporate more modes, we account for more variance and we accurately approximate the target shape (Patient 692). The first few modes account for the majority of the variance, and as we move to latter modes the improvement in reconstruction is smaller (i.e. mode 6 vs mode 7).

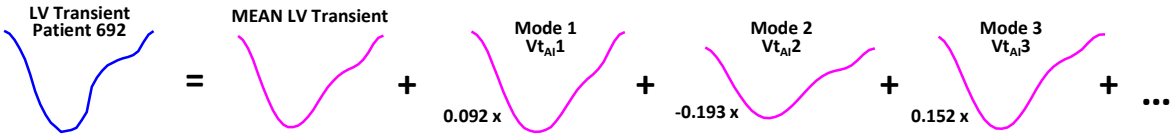


Fig. SIII.4 – Following PCA application, each LV volume temporal transient is decomposed into the mean volume transient (average of population of LV transients) plus the anatomical modes (transient variations, illustrated here as mean plus the positive extreme) times the corresponding PCA coefficient, $Vt_{AI}i$. The figure illustrates the decomposition of patient 692.

SUPPLEMENTAL MATERIAL IV – Additional Patient Characteristics

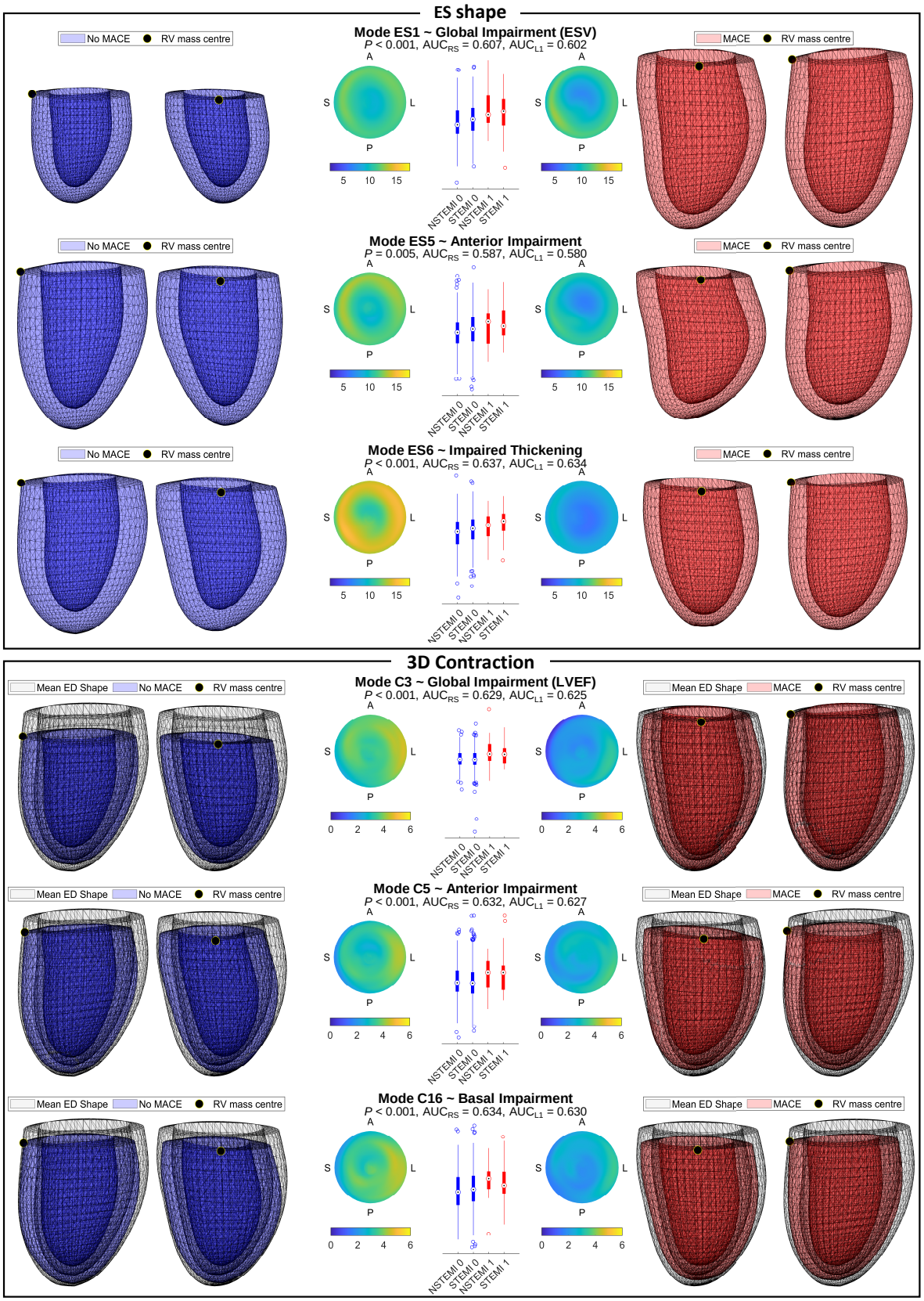


Fig. SIV.1 – The 3 main ES shape features (ES1, ES5, ES6) and contraction patterns (C3, C5, C16) related to MACE occurrence, resulting from the shape analysis explained in detail in (17), are

illustrated. Meshes shown in anterior and septal views, and as differential thickness maps (ED-ES thickness) on polar plots of the AHA model. To facilitate comparisons, the contractions are applied on the mean ED shape (reference transparent surface) and visualised as resulting ES shapes. No MACE (blue, class 0) and MACE (red, class 1) representations correspond to the 5th and 95th percentiles in the LDA direction. *P* values and both resubstitution (RS) and leave-one-out (L1) AUCs shown along MACE and no MACE distributions, further stratified into infarct aetiology (STEMI and NSTEMI). Reproduced from (17).

Table SIV.1 – Basic patient characteristics, cardiovascular risk factors and 3D LV biomarkers

Variable	ALL Patients	MACE (n = 73)	No MACE (n = 948)	AUC_k	*P* value	HR	HR *P*-val
Age	63 (52 - 72)	72 (61 - 77)	63 (52 - 72)	0.659	<0.001	1.80 (1.39 - 2.32)	<0.001
Sex	753/1011 (74.5)	46/71 (64.8)	707/940 (75.2)	0.505	0.052	0.81 (0.65 - 1.00)	0.050
Height, cm	2 (1 - 2)	2 (1 - 3)	1 (1 - 2)	0.597	0.003	0.69 (0.55 - 0.86)	0.001
Weight, kg	81 (72 - 90)	76 (70 - 86)	82 (73 - 90)	0.568	0.035	0.82 (0.64 - 1.05)	0.110
Cardiovascular risk factors							
Current smoking	405/935 (43.3)	19/63 (30.2)	386/872 (44.3)	0.522	0.029	0.75 (0.57 - 0.97)	0.032
Hypertension	716/1010 (70.9)	61/71 (85.9)	655/939 (69.8)	0.545	0.004	1.53 (1.13 - 2.08)	0.006
Hyperlipoproteinemia	624/1005 (62.1)	45/71 (63.4)	579/934 (62.0)	<0.5	0.816	1.03 (0.81 - 1.30)	0.824
Diabetes mellitus	231/1010 (22.9)	26/71 (36.6)	205/939 (21.8)	0.526	0.004	1.34 (1.09 - 1.64)	0.005
Body mass index, kg/m^2	27.4 (25.0 - 30.4)	27.0 (25.2 - 30.8)	27.4 (25.0 - 30.3)	<0.5	0.959	1.03 (0.82 - 1.30)	0.773
Body surface area, m2	1.95 (1.83 - 2.08)	1.88 (1.76 - 2.00)	1.96 (1.83 - 2.08)	0.593	0.006	0.74 (0.59 - 0.94)	0.014
Killip class on admission				0.573	<0.001	0.52 (0.41-0.65)	<0.001
1	899/1011 (88.9)	49/71 (69.0)	850/940 (90.4)				
2	76/1011 (7.5)	13/71 (18.3)	63/940 (6.7)				
3	20/1011 (2.0)	4/71 (5.6)	16/940 (1.7)				
4	16/1011 (1.6)	5/71 (7.0)	11/940 (1.2)				
Nr. of diseased vessels				0.567	0.003	1.40 (1.12 - 1.75)	0.003
1	502/1011 (49.7)	25/71 (35.2)	477/940 (50.7)				
2	310/1011 (30.7)	23/71 (32.4)	287/940 (30.5)				
3	199/1011 (19.7)	23/71 (32.4)	176/940 (18.7)				
TIMI flow grade post-PCI				<0.5	0.318	0.95 (0.77 - 1.16)	0.598
0	19/1011 (1.9)	1/71 (1.4)	18/940 (1.9)				
1	21/1011 (2.1)	2/71 (2.8)	19/940 (2.0)				
2	78/1011 (7.7)	8/71 (11.3)	70/940 (7.4)				
3	893/1011 (88.3)	60/71 (84.5)	833/940 (88.6)				
CMR biomarkers							
LV ESV, mL	70 (53 - 91)	86 (60 - 110)	69 (53 - 90)	0.599	0.004	1.43 (1.18 - 1.73)	<0.001
LV EDV, mL	144 (117 - 171)	145 (121 - 170)	144 (117 - 172)	<0.5	0.987	1.05 (0.83 - 1.33)	0.679
LVEF (%)	50.5 (43.3 - 57.3)	40.6 (33.1 - 52.2)	50.8 (44.0 - 57.5)	0.683	<0.001	0.80 (0.74 - 0.87)	<0.001
Infarct size, mL	17.2 (6.4 - 30.2)	24.6 (9.7 - 36.4)	16.7 (6.0 - 29.9)	0.591	0.006	1.29 (1.08 - 1.53)	0.005
Infarct size (% LV mass)	13.4 (5.4 - 21.8)	20.3 (9.6 - 28.9)	13.1 (5.3 - 21.4)	0.609	0.001	1.44 (1.18 - 1.76)	<0.001
MVO, mL	0.00 (0.00 - 1.90)	0.40 (0.00 - 3.00)	0.00 (0.00 - 1.80)	0.543	0.060	1.27 (1.11 - 1.45)	<0.001
MVO (% LV mass)	0.00 (0.00 - 1.39)	0.32 (0.00 - 2.15)	0.00 (0.00 - 1.27)	0.547	0.044	1.26 (1.09 - 1.46)	0.002
ES Shape				0.680			
Mode 1	-5 (-135 - 126)	62 (-61 - 233)	-10 (-136 - 119)		0.002	1.49 (1.18 - 1.87)	<0.001
Mode 5	0 (-34 - 36)	-11 (-51 - 22)	1 (-33 - 37)		0.014	0.73 (0.58 - 0.91)	0.006
Mode 6	-3 (-30 - 28)	-21 (-44 - 3)	-2 (-29 - 30)		<0.001	0.64 (0.50 - 0.81)	<0.001
Contraction Displacement				0.716			
Mode 3	-4 (-54 - 62)	49 (-29 - 106)	-6 (-55 - 59)		<0.001	1.70 (1.36 - 2.12)	<0.001
Mode 5	3 (-37 - 39)	-30 (-57 - 23)	4 (-33 - 40)		<0.001	0.63 (0.51 - 0.78)	<0.001
Mode 16	1 (-15 - 14)	-9 (-19 - 5)	1 (-14 - 15)		<0.001	0.65 (0.52 - 0.82)	<0.001

Data presented as n/N (%) or median (interquartile range). *P* values calculated between MACE/No-MACE groups. Hazard ratios (HR) presented with 95% confidence intervals and predictor significance. AUC_k provides the predictive power of each biomarker, assessed via LDA (median AUC, 10-cross-fold validated, 100 random data splits). MACE: major adverse cardiac events; PCI: percutaneous coronary intervention; TIMI: Thrombolysis in Myocardial Infarction. Reproduced from (17).

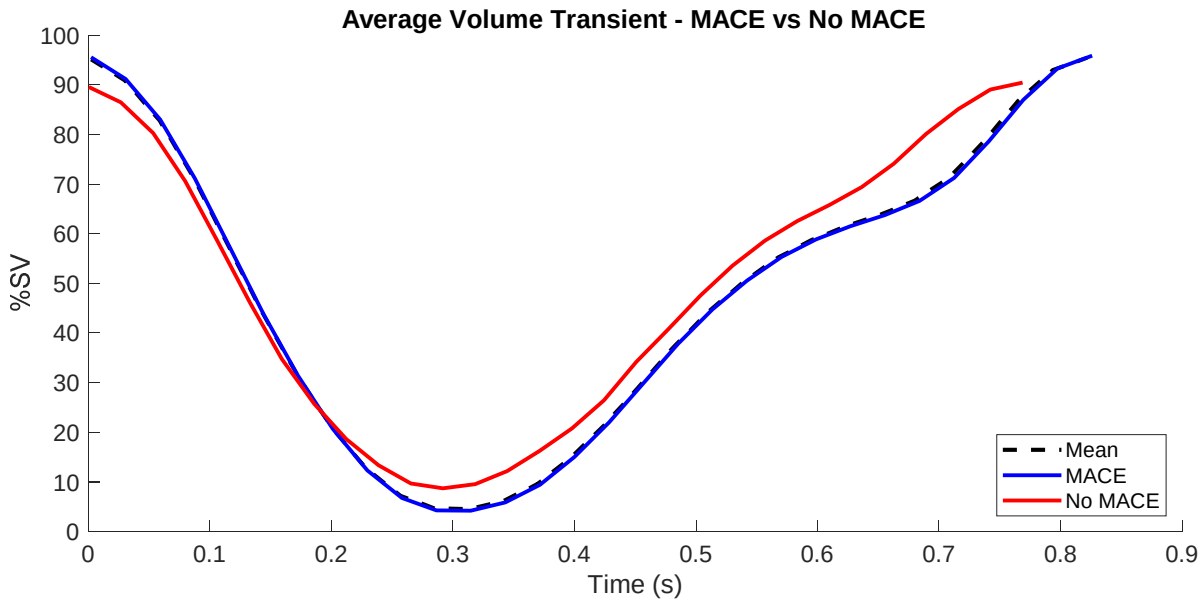


Fig. SIV.2 – Average LV volume transient, normalised by stroke volume, of the AMI cohort (dashed line), along with the average transient stratifying by MACE (red) and no MACE (blue).

SUPPLEMENTAL MATERIAL V – Endpoint Prediction Additional Results

Table SV.1 – LDA models coefficients

MODEL	VARIABLES	LDA COEFFICIENTS	CONSTANT
LVEF	LVEF	-0.65	-2.77
CMR	ESV, EDV	1.91, -1.51	-2.93
CMR + Vt	ESV, EDV, RR, $t_{diastolic}$	1.90, -1.48, -0.73, 0.69	-3.00
CMR + Vt_{AI}	ESV, EDV, $Vt_{AI}2$, $Vt_{AI}3$, $Vt_{AI}5$	1.79, -1.42, 0.36, 0.28, 0.37	-3.07
CMR + Strain	ESV, EDV, GLS	1.31, -1.14, 0.52	-3.00
CMR + LV_{3D}	ESV, EDV, C5, C16	1.76, -1.34, -0.33, -0.47	-3.07
ALL	ESV, EDV, Age, Killip	1.55, -1.14, 0.50, 0.53	-3.16
ALL + Vt	ESV, Age, Killip, BSA, RR, $t_{diastolic}$	0.65, 0.58, 0.53, -0.32, -0.91, 0.65	-3.15
ALL + Vt_{AI}	ESV, EDV, Age, Killip, $Vt_{AI}3$, $Vt_{AI}5$	1.52, -1.17, 0.55, 0.51, 0.36, 0.35	-3.26
ALL + Strain	ESV, Age, Killip, BSA, Weight, Vessels, GLS	0.41, 0.44, 0.43, -1.19, 0.83, 0.21, 0.63	-3.24
ALL + LV_{3D}	ESV, EDV, Age, Killip, C5, C16	1.40, -0.98, 0.48, -0.34, -0.43, 0.51	-3.28
All + Vt_{AI} + Strain + LV_{3D}	ESV, EDV, Age, C16, GLS, $Vt_{AI}3$, $Vt_{AI}5$	1.09, -0.88, 0.58, -0.47, 0.58, 0.37, 0.42	-3.35

Coefficients and constant of the resulting LDA models following the backward stepwise variable selection, to assess the additional prognostic contribution of LV contraction unravelling via conventional (Vt) and AI-derived (Vt_{AI}) volume transient; CMR-FT strains (strains); and LV 3D detail patterns (LV_{3D}) metrics on top of considering only CMR biomarkers (CMR) or all the cardiovascular risk factors and patient characteristics of the study (ALL). All the variables are normalised to zero mean and unit variance prior to LDA fitting. The performance of these models is reported in the main manuscript (Table 3).

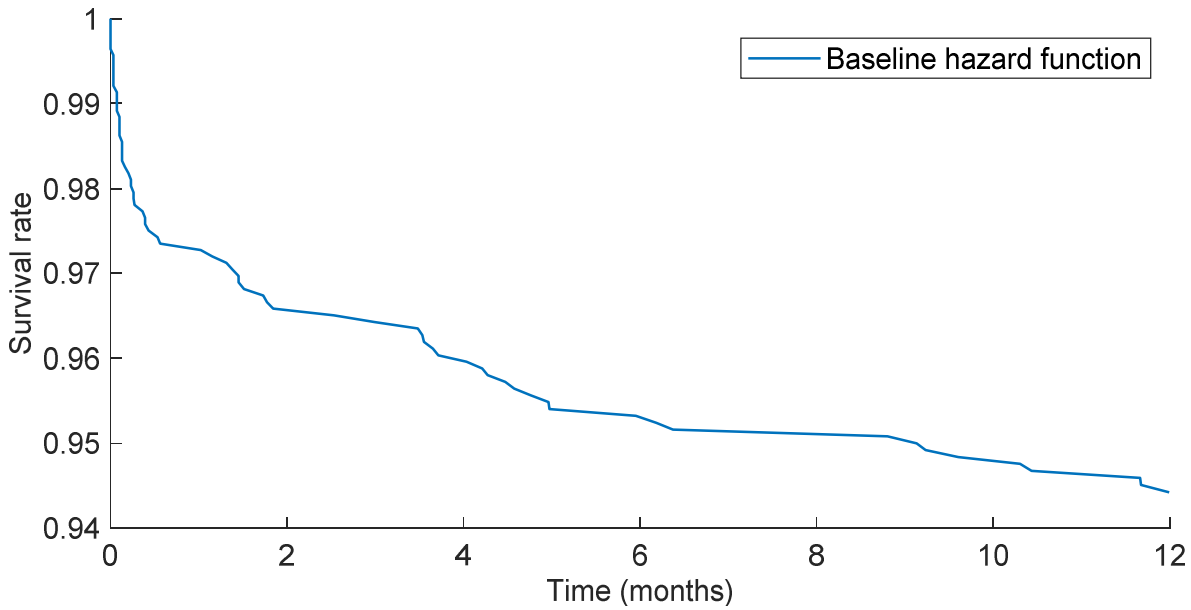


Fig. SV.1 – Underlying baseline hazard function

Table SV.2 – Cox models hazard ratios

MODEL	VARIABLES	HR	P-value
LVEF	LVEF	0.80 (0.74 - 0.87)	< 0.001
CMR	ESV	4.45 (3.01 - 6.58)	< 0.001
	EDV	0.28 (0.18 - 0.43)	< 0.001
CMR + Vt	ESV	4.45 (3.01 - 6.58)	< 0.001
	EDV	0.28 (0.18 - 0.43)	< 0.001
CMR + Vt_{AI}	ESV	4.51 (2.95 - 6.89)	< 0.001
	EDV	0.26 (0.16 - 0.42)	< 0.001
	$Vt_{AI}3$	1.28 (1.05 - 1.55)	0.013
	$Vt_{AI}5$	1.28 (1.05 - 1.56)	0.015
CMR + Strain	ESV	2.79 (1.73 - 4.52)	< 0.001
	EDV	0.38 (0.23 - 0.63)	< 0.001
	GLS	1.69 (1.28 - 2.23)	< 0.001
CMR + LV_{3D}	ESV	4.23 (2.72 - 6.60)	< 0.001
	EDV	0.29 (0.18 - 0.48)	< 0.001
	C16	0.64 (0.51 - 0.80)	< 0.001
	C5	0.76 (0.61 - 0.95)	0.018
ALL	ESV	3.22 (2.12 - 4.91)	< 0.001
	EDV	0.39 (0.24 - 0.63)	< 0.001
	Age	1.55 (1.19 - 2.01)	0.001
	Killip	1.26 (1.10 - 1.45)	0.001
ALL + Vt	ESV	1.56 (1.29 - 1.89)	< 0.001
	Age	1.66 (1.27 - 2.17)	< 0.001
	RR	0.58 (0.44 - 0.78)	< 0.001
	$v_{diastolic_avg}$	0.09 (0.02 - 0.43)	0.003
	Killip	1.22 (1.05 - 1.41)	0.008
ALL + Vt_{AI}	ESV	3.32 (2.12 - 5.19)	< 0.001
	EDV	0.36 (0.21 - 0.60)	< 0.001
	Age	1.59 (1.21 - 2.07)	< 0.001
	Killip	1.25 (1.08 - 1.43)	0.002
	$Vt_{AI}3$	1.29 (1.06 - 1.57)	0.011
	$Vt_{AI}5$	1.26 (1.03 - 1.54)	0.025
ALL + Strain	GLS	1.78 (1.38 - 2.30)	< 0.001
	Age	1.58 (1.21 - 2.05)	< 0.001
	Killip	1.22 (1.04 - 1.43)	0.013
	BSA	0.72 (0.55 - 0.93)	0.014
	ESV	1.31 (1.06 - 1.63)	0.014
ALL + LV_{3D}	ESV	3.41 (2.15 - 5.40)	< 0.001
	C16	0.65 (0.51 - 0.81)	< 0.001
	EDV	0.39 (0.23 - 0.65)	< 0.001
	Age	1.55 (1.19 - 2.02)	0.001
	C5	0.76 (0.60 - 0.95)	0.019
ALL + Vt_{AI} + Strain + LV_{3D}	GLS	1.78 (1.38 - 2.30)	< 0.001
	C16	0.61 (0.48 - 0.78)	< 0.001
	ESV	2.36 (1.42 - 3.92)	< 0.001
	Age	1.49 (1.15 - 1.95)	0.003
	EDV	0.47 (0.27 - 0.81)	0.007
	$Vt_{AI}5$	1.25 (1.05 - 1.49)	0.012

Hazard ratios, HR (95% confidence interval), and predictor significance, *P*-value, of the resulting Cox multivariate models following the backward stepwise variable selection, to assess the additional prognostic contribution of LV contraction unravelling via conventional (Vt) and AI-derived (Vt_{AI}) volume transient; CMR-FT strains (strains); and LV 3D detail patterns (LV_{3D}) metrics on top of considering only CMR biomarkers (CMR) or the cardiovascular risk factors and patient characteristics of the study (ALL). The performance of these models is summarised in the manuscript (Table 3).

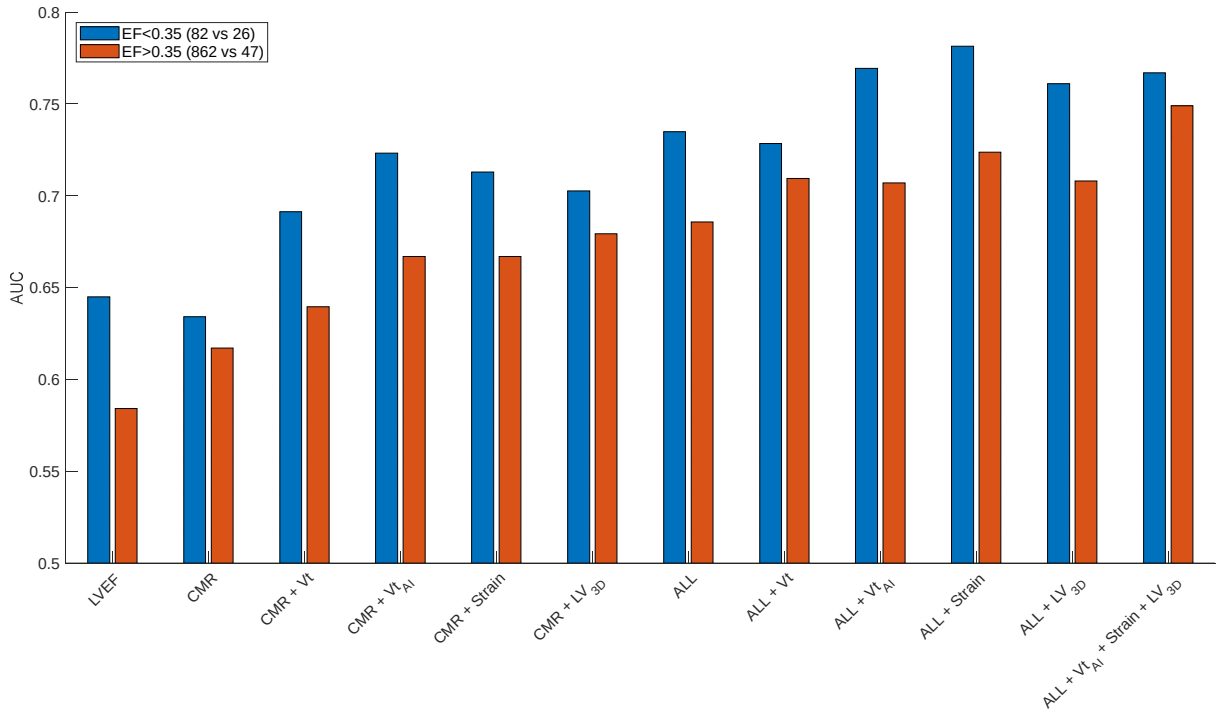


Fig. SV.2 – AUC endpoints prediction results stratifying by LVEF (threshold: 0.35) and applying the LDA models resulting from the backward stepwise analysis (See Table 3 and Table SV.1). The labels are presented as subgroup (MACE vs No MACE cases within the subgroup).

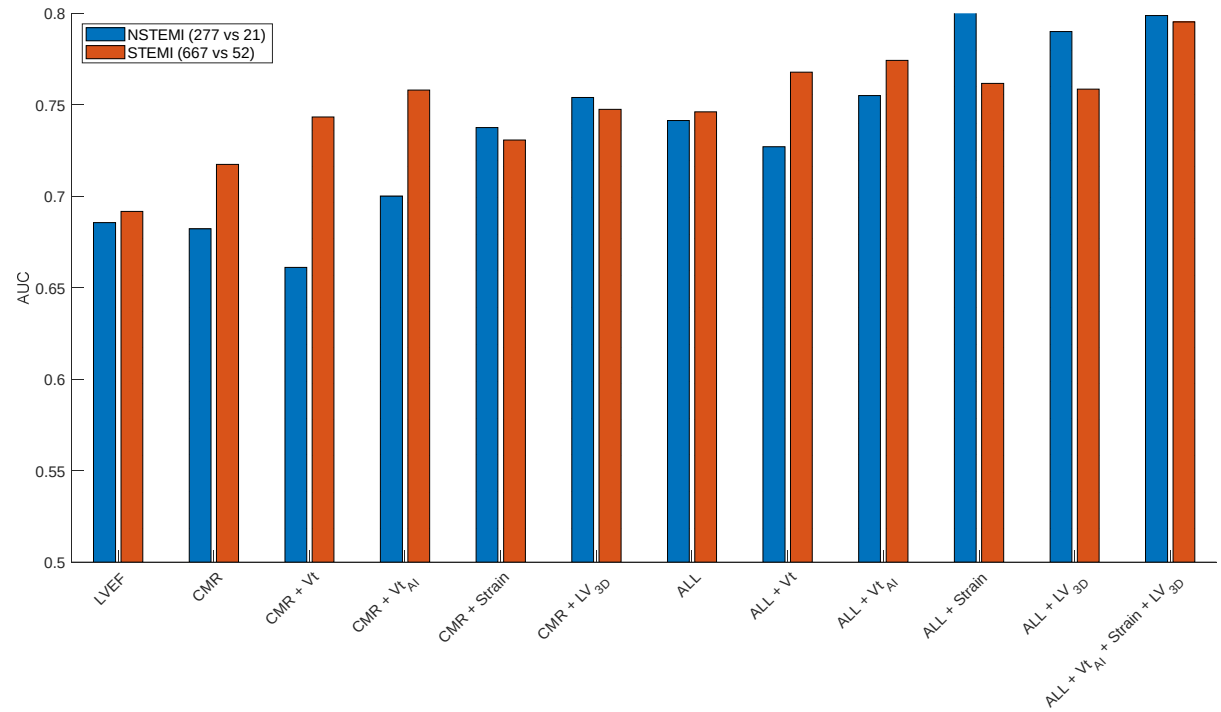


Fig. SV.3 – AUC endpoints prediction results stratifying by infarct aetiology (STEMI vs NSTEMI) and applying the LDA models resulting from the backward stepwise analysis (See Table 3 and Table SV.1). The labels are presented as subgroup (MACE vs No MACE cases within the subgroup).

SUPPLEMENTAL MATERIAL VI – Modes Correlation Additional Results

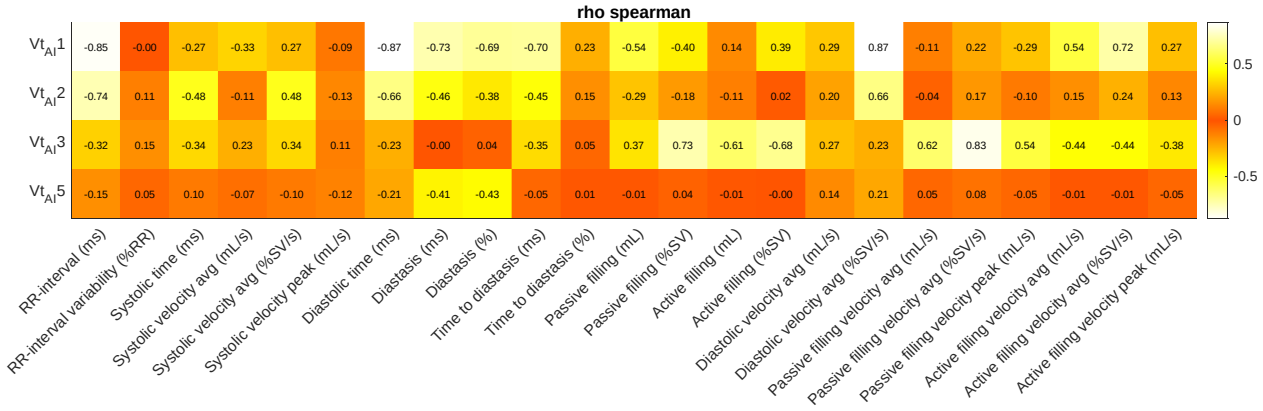


Fig. SVI.1 – Heat map of Spearman correlation coefficients between AI-derived (columns) and conventional (rows) volume transient metrics.

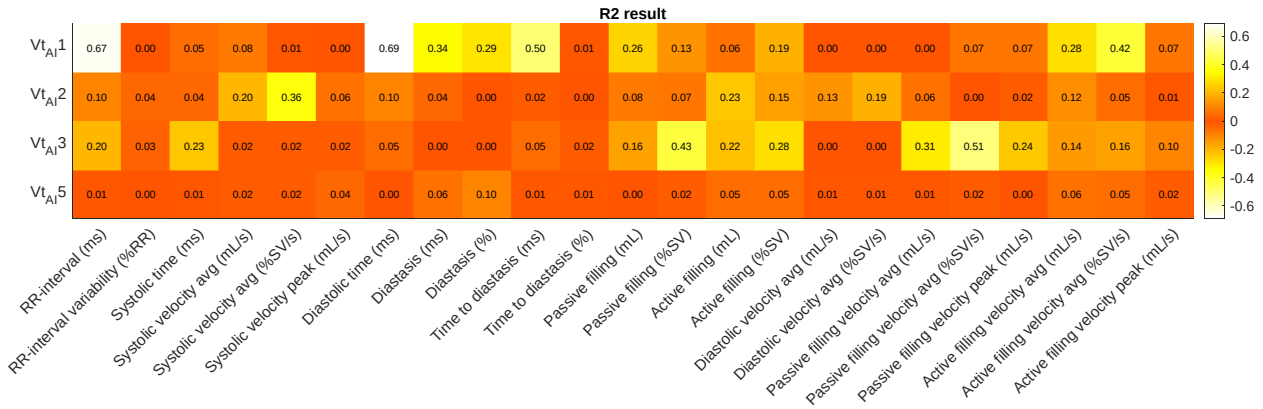


Fig. SVI.2 – Heat map of R2 correlation coefficients between AI-derived (columns) and conventional (rows) volume transient metrics.

SUPPLEMENTAL MATERIAL VII – Systolic vs Diastolic Components

The LV volume temporal transients are split into their systolic and diastolic components using the systolic time as reference. The analysis is repeated for each of the two components independently as described in Methods and Supplemental Material III.

The 95% of the population variance was explained by the first 4 systolic modes of variation (Vt-S_{AI}) and the first 5 diastolic modes of variation (Vt-D_{AI}), respectively. Among them, the LDA stepwise analysis determined modes 2, 4 and 5 (Vt-D_{AI}2, Vt-D_{AI}4 and Vt-D_{AI}5) as relevant to MACE in the diastolic component analysis; and only mode 2 (Vt-S_{AI}2,) in the systolic component scenario (see Fig. SVII.1).

The experiments summarised in Table SVII.1 show that any of the two components significantly contributes to the baseline model (cardiovascular risk factors, basic patient characteristics and established CMR markers), as the LDA stepwise selection demonstrates. This is particularly interesting in the systolic component case study, where the mode Vt-S_{AI}2 is significantly related to MACE in combination with other variables but not if analysed individually (p = 0.148, see Fig. SVII.1). Nevertheless, only the diastolic component contributes to a significant improvement in prediction performance. This is in line with the results of the entire transient analysis (Vt_{AI}), presented in the main manuscript (see Table 3), where the two modes that contributed the most to the multivariate models were related to diastolic function, that is, Vt_{AI}3 and Vt_{AI}5. It is also sensible that the inclusion of the entire transient provides more additional prognostic value than any of the two components individually analysed.

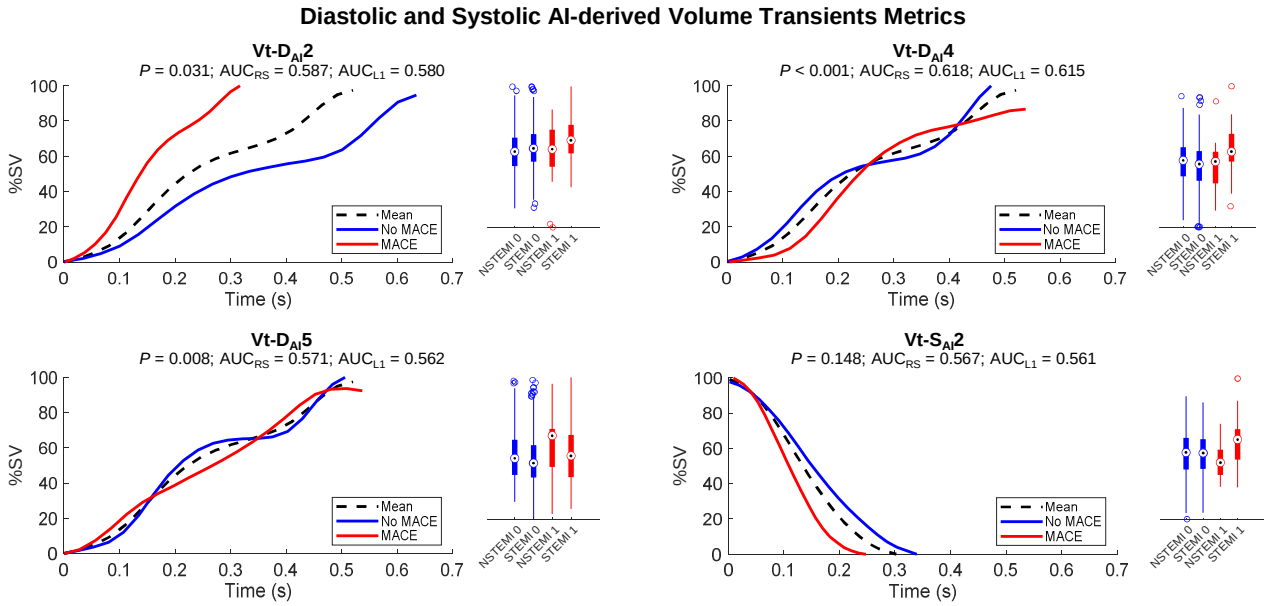


Fig. SVII.1 –AI-derive volume transient features relevant to MACE occurrence prediction, resulting from diastolic component unsupervised learning (Vt-D_{AI}2, Vt-D_{AI} 4, and Vt-D_{AI} 5) and systolic component analysis (Vt-S_{AI}5). The MACE (red, class 1) and No MACE (blue, class 0) representations correspond to the 10th and 90th percentiles in the LDA direction. This allows to visualise the particular pattern or change encode by each of the unsupervised variables (RR-interval, diastasis, etc.) as well as to describe how a representative MACE and No-MACE volume transient components would theoretically look like according to each of these four unsupervised variables. The *P* value, re-substitution and leave-one-out AUCs are presented along each mode as MACE and No-MACE distributions, further stratified into infarct aetiology (STEMI and NSTEMI).

Table SVII.1 – Additional prognostic contribution of AI volume transient components

MODEL	LINEAR SELECTION	AUC_k	AUC_{RS}
ALL	ESV, EDV, Age, Killip	0.729 (0.727 - 0.733)	0.745
ALL + Vt_{AI}	ESV, EDV, Age, Killip, $Vt_{AI}3$, $Vt_{AI}5$	0.746 (0.743 - 0.749)	0.769
ALL + Vt-D_{AI}	ESV, EDV, Age, Killip, Vt-$D_{AI}4$	0.736 (0.733 - 0.738)	0.751
ALL + Vt-S_{AI}	ESV, EDV, Age, Killip, Vt-$S_{AI}2$	0.726 (0.723 - 0.730)	0.743

Backward stepwise LDA results of the additional prognostic contribution of the AI-derived volume transient metrics, considering the entire transient (Vt_{AI}), the diastolic component (Vt-D_{AI}) or the systolic component (Vt-S_{AI}), on top of all the cardiovascular risk factors and patient characteristics of the study (ALL). The resulting significant selection of variables is reported along with the predictive performance, expressed as AUC re-substitution (RS) and 10-fold cross-validated (K), computed for a 100 random data splits and presented as median (interquartile range). Killip indicates Killip class on admission.

SUPPLEMENTAL MATERIAL VIII – TRIPOD reporting

Table SVIII.1 – TRIPOD Checklist

DESCRIPTION	ITEM	CHECKLIST	MANUSCRIPT CHECK
Title & abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted	Title
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions	Abstract
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models	Introduction (par.1-3)
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both	Introduction (par. 4)
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable	Methods (Study Population)
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up	Methods (Study Population & Study Endpoints)
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centers	Methods (Study Population & CMR Imaging Protocol)
	5b	Describe eligibility criteria for participants	
	5c	Give details of treatments received, if relevant	
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed	Methods (Study Endpoints)
	6b	Report any actions to blind assessment of the outcome to be predicted	
Predictors	7a	Clearly define all predictors used in developing the multivariable prediction model, including how and when they were measured	Methods (Prognostic Value Assessment & Statistical Analysis), Tables 1 and 2
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors	
Sample size	8	Explain how the study size was arrived at	*
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method	Fig.1 and Results (Patients)
Statistical analysis methods	10a	Describe how predictors were handled in the analyses	Methods (Prognostic Value Assessment & Statistical Analysis)
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation	
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models	
Risk groups	11	Provide details on how risk groups were created, if done	Methods (Study Endpoints)
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful	Table 2, Table SIV.1, Fig.1 and Results (Patients)
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome	
Model development	14a	Specify the number of participants and outcome events in each analysis	Table 2, Fig.1 and Results (Patients)
	14b	If done, report the unadjusted association between each candidate predictor and outcome	
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point)	Results and Supplemental Materia V
	15b	Explain how to use the prediction model	
Performance	16	Report performance measures (with CIs) for the prediction model	Table3, Results, and Supplementary Data V

Continued

Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data)	Discussion (Limitations)
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence	Discussion
Implications	20	Discuss the potential clinical use of the model and implications for future research	Discussion (Impact and Clinical Translation)
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets	Supplemental Material
Funding	22	Give the source of funding and the role of the funders for the present study	Funding

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) checklist, adapted from (33). Only model development items are included. (*) The size of the study is usually determined based on the variables of interest and their standard deviation. The proposed study cannot benefit from this criterion since the variables (modes of variation) were not defined *a priori*. However, aiming at a margin of error of ≤0.05 in the overall outcome proportion estimate, a mean absolute prediction error of 0.05, a desired shrinkage ≤10% and conservative anticipated Cox-Snell R squared statistic of 0.2, and given the approximate MACE incidence of 7% and the number of predictors indicated in the manuscript, the actual size of the study is superior to the trial target size retrospectively determined by any of the 4 proposed calculation methods explained in Riley et al. (24). Par. indicates paragraph.