University of Sheffield

Artificial Intelligence in Cardiac Magnetic Resonance Imaging to Predict Prognosis and Treatment Response

Dr Samer Alabed

Supervisors: Dr Andrew J Swift and Dr Haiping Lu

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy in Medicine

in the

Department of Infection, Immunity and Cardiovascular Disease

January 17, 2023

Declaration

All sentences or passages quoted in this document from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure.

Name: *Dr Samer Alabed*

Signature: 

Date: *17.01.2023*

بسم الله الرحمن الرحيم

Acknowledgements

I am incredibly grateful to my supervisor, Dr Andy Swift and Dr Haiping Lu, for their invaluable guidance and support throughout my PhD. Their expertise and mentorship were instrumental in helping me to develop a deep understanding of AI applications in cardiac MRI and to complete my thesis in a timely and successful manner.

I would also like to thank Professors David Kiely and Jim Wild and Dr Pankaj Garg for their valuable feedback and suggestions, which greatly improved the quality of my work. I am indebted to Dr Rob van der Geest for developing tools that were paramount to my work.

I am grateful to Professor Paul Griffiths for first hiring me as an Academic Clinical Fellow in Radiology and trusting in my abilities, and to Professor Nigel Hoggard for his support in my academic career. I am also grateful to Dr Dan Connelly for his excellent radiology teaching. My thanks extend to my friends and colleagues at the University of Sheffield and Sheffield Teaching Hospitals. Many come to mind, especially Drs Pete Metherall, Chris Johns, Krit Dwivedi, Mahan Salehi, Faisal Alandejani, Michael Sharkey, Ahmed Maiter and Shuo Zhou.

I would like to express my deepest thanks to Dr Kavita Karunasagaraar for her guidance and mentorship, which has been pivotal in inspiring me to pursue a career in cardiothoracic imaging and to grow as a radiologist and researcher.

I am endlessly grateful to my wife, Aisha, for her unwavering support and love during my PhD journey. Her constant encouragement and belief in me have been my strength and inspiration. I could not have completed this journey without you by my side and I am truly blessed and will be forever thankful for your dedication to our beautiful family. I am also grateful to my parents for their inspiration and support, and a special thanks to my brother, Sami, for teaching me how to code and helping me debug my coding errors in many moments of despair.

Thank you all for supporting me and being a part of my journey.

Synopsis

Background

Pulmonary arterial hypertension (PAH) is a serious disease of the heart and lungs. Its impact on patients can be severe, including limitation of day-to-day activities and high mortality. The diagnosis, treatment and monitoring of PAH are challenging and there is a need for tools that can aid clinical decision-making to optimise patient outcomes.
Cardiac MRI (CMR) provides both qualitative and quantitative information about cardiac function and is an important method for evaluating the severity of PAH. The application of machine learning (ML) tools is of growing interest in medical imaging. ML has the potential to automate complex and repetitive tasks, including the rapid segmentation of anatomical structures on images and extraction of clinically useful information.

Aims

This thesis proposes the combination of CMR with two different ML tools to predict prognosis and treatment response in PAH. The first ML tool involves the automated measurement of different cardiac parameters and assesses their utility in predicting prognosis and treatment response. The second ML tool involves the extraction of imaging features directly without the need for segmentation to predict the risk of mortality.

My Contribution

The ML models in this thesis were developed at the University of Sheffield in collaboration with Leiden University. Sheffield is a centre of excellence in PAH treatment thanks to the Sheffield Pulmonary Vascular Disease Unit, which is one of the largest internationally. Each year, more than 700 PAH patients undergo CMR for diagnosis and monitoring. Additionally, each newly diagnosed patient has accompanying in-depth clinical phenotypic data, including right heart catheterisation, exercise and pulmonary function tests, and quality of life assessment. During my research, I created and curated a dataset combining imaging and time-matched clinical data. I identified eligible CMR scans, landmarked and contoured cardiac chambers on multiple sequences and organised the collaboration with computer scientists at Leiden and Sheffield. I arranged image anonymisation, storage and transfer and advised computer scientists on the clinical relevance of CMR images. I

performed quality control on ML analyses, collated their results, and analysed the data within clinical context. I have written all chapters in this thesis and clarified the roles of my co-authors at the end of each chapter.

Thesis Outline

Chapter 1 provided an overview of the growing role of CMR in the diagnosis and evaluation of PAH. Chapter 2 summarised the prognostic value of CMR measurements in the prediction of clinical worsening and mortality in PAH patients. Chapter 3 illustrated the rapid expansion of research using AI approaches to automate CMR measurements. The quality of the existing literature was reviewed, significant shortcomings in the transparency of studies were identified and solutions were recommended. Chapter 4 showed our experience in developing, validating and testing a fully automatic CMR segmentation tool. Our tool was developed in one of the largest multi-vendor, multi-centre and multi-pathology reported datasets, and included a large group of patients with right heart disease. We implemented the lessons learned in Chapter 3 and provided extensive descriptions of our datasets, ML model and performance. Our model showed excellent reliability, generalisability, agreement with CMR experts and correlation with invasive haemodynamics. Chapter 5 demonstrated that the automatic CMR measurements allowed assessment of patient-orientated outcomes and prediction of mortality. Thresholds of changes in CMR metrics were identified that could inform clinical decisions in the monitoring of PAH patients. Chapter 6 showed promising results of an ML tool to extrapolate prognostic CMR features with incremental value compared to clinical risk scores and volumetric CMR measurements. Finally, Chapter 7 showed that myocardial T1 mapping could potentially add diagnostic and prognostic value in PAH.

Impact and Future Direction

In addition to the known advantages of ML for providing rapid results with minimal human involvement, the ML tools developed in this thesis allow visualisation of outcomes and are transparent to the human assessor. ML applications to automate the measurement of CMR metrics and extract prognostic imaging features have potential to add clinical value by (i) streamlining prognostication, (ii) informing treatment selection, (iii) assisting the monitoring of treatment response and (iv) ultimately improving clinical decision-making and patient outcomes. Additionally, these tools could point to new CMR end-points for clinical trials, accelerating the development of new treatments for PAH. ML will likely elevate the role of CMR as a powerful prognostic modality in the years to come. Looking ahead, I hope to combine multi-source clinical, imaging and patient-orientated data from several ML tools into a single package to facilitate the assessment of cardiovascular disease.

Contents

List of Figures

List of Tables

Abbreviations

4-Ch	Four Chamber.
6MWD	6-Min Walk Distance.
6MWT	6-Min Walk Test.
AI	Artificial Intelligence.
AIC	Akaike Information Criterion.
ANGIE	Accelerated And Navigator-Gated Look-Locker Imaging.
ASPIRE	Assessing The Severity Of Pulmonary Hypertension In A Referral Centre.
AUC	Area Under The Curve.
BNP	B-Type Natriuretic Peptide.
BSA	Body Surface Area.
CENTRAL	Cochrane Central Register Of Controlled Trials.
CHD	Congenital Heart Disease.
CI	Confidence Intervals.
CLAIM	Checklist for Artificial Intelligence in Medical Imaging.
CMR	Cardiac Magnetic Resonance.
CNN	Convolutional Neural Network.
CO	Cardiac Output.
COPD	Chronic Obstructive Pulmonary Disease.
CT	Computed Tomography.
CTD	Connective Tissue Disease.
CTEPH	Chronic Thromboembolic Pulmonary Hypertension.
DICOM	Digital Imaging And Communications In Medicine.
DLCO	Diffusing Capacity For Carbon Monoxide.
DSC	Dice similarity coefficient.
E-10	emPHasis-10.
Ea	Arterial Elastance.
ECV	Extracellular Volume.
EDM	End-Diastolic Mass.
EDV	End-Diastolic Volume.
Ees	End-Systolic Elastance.
EF	Ejection Fraction.
eGFR	Estimated Glomerular Filtration Rate.
ERS	European Respiratory Society.
ESC	European Society Of Cardiology.
ESV	End-Systolic Volume.

FDA	Food and Drug Administration.
FEV1	Forced Expiratory Volume In 1 Second.
FOV	Field Of View.
FPHR	French Pulmonary Hypertension Registry.
FVC	Forced Vital Capacity.
GLI	Global Lung Function Initiative.
HFpEF	Heart Failure With Preserved Ejection Fraction.
HITL	Human In The Loop.
HIV	Human Immunodeficiency Virus.
HR	Hazard Ratio.
ICC	Interclass Correlation Coefficient.
IPAH	Idiopathic Pulmonary Arterial Hypertension.
IQR	Interquartile Range.
ISWT	Incremental Shuttle Walking Test.
IVS	Interventricular Septal Angle.
LA	Left Atrium.
LGE	Late Gadolinium Enhancement.
LHD	Left Heart Disease.
LV	Left Ventricle.
LVEF	Left Ventricular Ejection Fraction.
MD	Mean Difference.
MICCAI	Medical Image Computing and Computer-Assisted Intervention.
MID	Minimally Important Difference.
ML	Machine Laerning.
MOLLI	Modified Look-Locker Inversion Recovery.
MPA	Main Pulmonary Artery.
mPAP	Mean Pulmonary Artery Pressure.
MPCA	Multilinear Principal Component Analysis.
mRAP	Mean Right Atrial Pressure.
MRI	Magnetic Resonance Imaging.
Ms	Millisecond.
NHS	National Health Service.
PA	Pulmonary Artery.
PAH	Pulmonary Arterial Hypertension.
PAWP	Pulmonary Artery Wedge Pressure.
PFT	Pulmonary Function Test.
PH	Pulmonary Hypertension.
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
PROSPERO	International Prospective Register Of Systematic Reviews.

PVR Pulmonary Vascular Resistance.

QIPS Quality In Prognosis Studies.

RA Right Atrium.
RAC Relative Area Change.
REVEAL Registry To Evaluate Early And Long-Term Pulmonary Arterial Hypertension Disease Management.
RHC Right Heart Catheterisation.
RIS Radiological Information System.
ROC Receiver Operator Characteristics.
RV Right Ventricle.
RVEDVI Right Ventricle End Diastolic Volume Index.
RVEF Right Ventricle Ejection Fraction.
RVESVI Right Ventricle End Systolic Volume Index.
RVIP Right Ventricle Insertion Point.
RVMI Right Ventricle Mass Index.
RVOT Right Ventricular Outflow Tract.
RVSVI Right Ventricle Stroke Volume Index.

SA Short Axis.
SASHA Saturation Recovery Single-Shot Acquisition.
SD Standard Deviation.
SNR Signal To Noise Ratio.
SPECT Single Photon Emission Computed Tomography.
SPVDU Sheffield Pulmonary Vascular Disease Unit.
STARD Standards for Reporting of Diagnostic Accuracy Studies.
STH Sheffield Teaching Hospitals.
SV Stroke Volume.
SVM Support Vector Machine.
Svo2 Mixed Venous Oxygen Saturation.

T Tesla.
TLCO Transfer Factor Of Carbon Monoxide.

VMI Ventricular Mass Index.

WHO World Health Organisation.
WU Woods Units.

Introduction to background of the thesis

Radiology is transforming, and Artificial Intelligence (AI) will play a significant role in all aspects of medical imaging [1], and the vanguard of this change is cardiac imaging. The advanced technical requirements in capturing the beating heart and cardiac blood flow are pushing the boundaries for AI advancement. However, before the start of my research, the thought of AI in medical imaging filled me with unease. On the one hand, rapid advancements in technology and constant reports of AI outperforming radiologists raised concerns about the future of my profession and its eventual "replacement" by AI. After all, radiology is a highly specialised medical field that took me years of training and involved sitting one of the most challenging post-graduate medical exams. On the other hand, radiology is the fastest-changing and most technologically advanced medical specialty. The detailed images of human anatomy and pathology revealed by constantly evolving imaging techniques are truly fascinating. Radiologists have always adopted and adapted new technologies in their favour to maximise their benefit to patients. As my passion for cardiac imaging grew, I could not help but wonder, could AI be the key to unlocking new levels of accuracy and efficiency? And could it lead to the early detection of abnormalities that might otherwise go unnoticed by the human eye? These questions sparked my curiosity and desire to understand the potential impact of AI in cardiac imaging. The thought of contributing to the development and clinical implementation of AI tools with the ultimate goal of improving patient care motivated me to embark on a PhD journey.

My thesis assesses the role of AI in cardiac MRI in predicting the prognosis of cardiopulmonary disease, mainly pulmonary hypertension (PH) and the first chapter is an overview of the existing cardiac MRI applications in PH. Chapters 2 to 5 assess cardiac MRI volumetric measurements, while chapters 6 and 7 look at the role of myocardial

tissue characterisation and direct imaging features identified by AI. The outline of the chapters of this thesis is shown in (Figure 1).

Sheffield has a strong track record in cardiothoracic imaging in pulmonary hypertension (PH), led by Dr Andy Swift (my supervisor) and by Professor David Kiely and Professor Jim Wild. The Sheffield Pulmonary Vascular Disease Unit (SPVDU) is one of the largest PH centres in the world and one of the pioneers of cardiac MRI in PH. My PhD starts with a scoping literature review on the role of different cardiac MRI techniques in assessing PH [2]. This review, combined with previous experiences at Sheffield, highlighted the promising role of cardiac MRI in the prognostic assessment of PH.

Science should begin and end with up-to-date systematic reviews [3], and so does my thesis. My hypothesis is that cardiac MRI could predict mortality and clinical adverse events based on the volumetric assessment of the cardiac chambers, including the right ventricle (RV). I performed a comprehensive systematic review and meta-analysis that identified all studies assessing the prognostic utility of cardiac MRI in pulmonary arterial hypertension (PAH) [4]. I contacted the authors of all identified studies to obtain volumetric cardiac MRI data in order to create exhaustive evidence. In a notable collaborative spirit, 14 PH centres around the world responded with additional unpublished data and highlighted the PH community's interest in better defining the role of cardiac MRI for the prognostication purposes in PH. The systematic review confirmed the significant role of cardiac MRI measurements in predicting clinical worsening and mortality in PAH and is now cited in the European guidelines for managing PH.

The systematic review generated further questions. Firstly, what is the role of AI in obtaining cardiac MRI volumetric measurements? Secondly, what has been achieved and what is still to be done in AI measurements? And thirdly, how effective are AI measurements in predicting prognosis in PH? To answer the first two questions and to help plan a study to answer the third question, I conducted a further systematic review that aimed to identify all studies using AI to segment cardiac MRI images [5]. Systematic reviews are essential tools for designing studies that identify gaps in the existing evidence and provide a basis for informing new research [6]. My systematic review has shown that

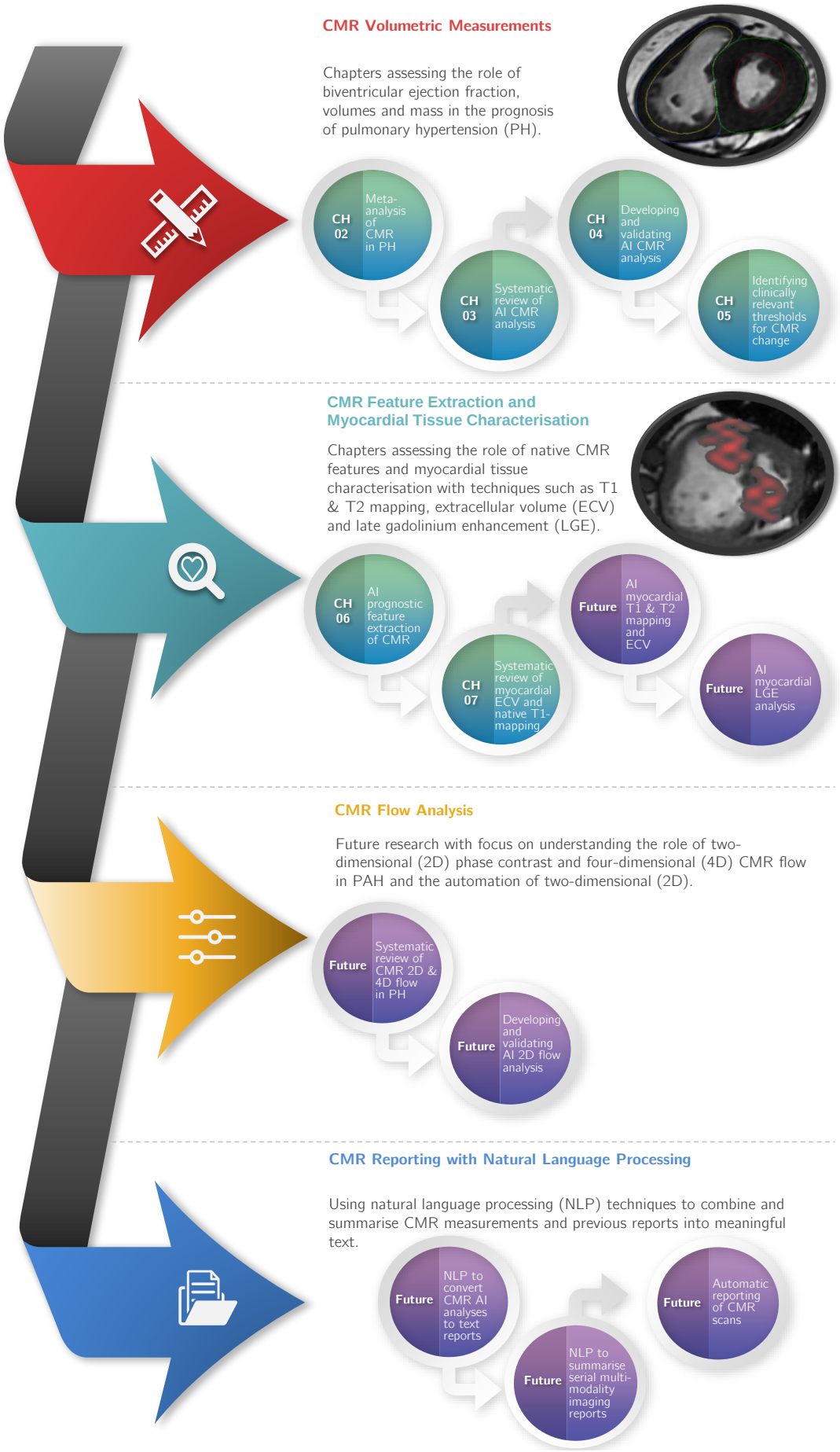


Figure 1: Thesis outline. Green represents the chapters in this thesis. Purple indicates areas of future research.

although more than 200 AI studies in cardiac MRI segmentation have been conducted, more progress has yet to be made in evaluating the RV or showing a clinical impact of AI measurements. Surprisingly, right ventricular disease was ignored from inclusion in training or testing datasets, and the clinical characteristics of patients were poorly reported. External validation of performance and failure assessment was only performed in a minority of studies. For the AI research community, my systematic review promotes improvement in the quality of reports and provides a reference to help avoid wasted resources in AI health research. The results of the review set the context of a study to automate cardiac MRI ventricular and atrial segmentation in right heart abnormalities.

While most AI cardiac MRI segmentation studies focused on the left ventricle, there remained uncertainty in the automatic assessment of the RV, particularly in right ventricular diseases such as PH. In addition, most algorithms for the RV only segment the inner surface, the endocardium. As such, no AI algorithm to date has managed to measure the RV muscle mass which is an important measurement in suspected PH. In a research collaboration with computer scientists at the University of Leiden (led by Dr Rob van der Geest) we aimed to develop an AI segmentation tool to assess the four chambers of the heart on short-axis, long-axis (two-chamber) and four-chamber imaging. Our research was driven by lessons learnt from the AI segmentation systematic review that revealed shortcomings of other cardiac MRI AI research. We aimed to address the important uncertainties of AI segmentation, including assessment of accuracy against invasive haemodynamics, the robustness of findings in scan-rescan repeat imaging and external validation in multi-centre and multi-vendor datasets [7]. In addition, our study included one of the largest AI cardiac MRI segmentation datasets (> 5600 scans and $> 200,000$ images). The results showed that AI segmentation is more accurate and repeatable than manual assessment. Additionally, we have shown that cardiac MRI measurements from AI contours can predict mortality in PAH.

Cardiac MRI in PH is typically used to monitor progression and treatment response [2, 8]. While Chapters 2 and 4 show that cardiac MRI measurements could predict clinical worsening and mortality, the threshold of important change was poorly defined.

The change in cardiac volumes indicating a change in PH disease state effect needed to be addressed. I chose a patient-centred approach to identify the minimally important differences in cardiac volumes using quality of life and exercise tests as anchors. I identified treatment response thresholds that can be applied in everyday clinical practice from paired follow-up data and those can guide radiologists and clinicians in assessing change based on cardiac MRI. The novelty of my approach is to benchmark changes in cardiac MRI measurements to changes in health-related quality of life so as to link how a patient feels to the imaging findings.

I have also explored the prognostic utilities of cardiac MRI beyond AI segmentation in Chapters 6 and 7. A tensor-based machine learning framework is applied in the prognostic assessment of PAH in chapter 6. This approach deals with MRI images as multi-dimensional data blocks called tensors allowing for better data structure preservation than traditional linear machine learning models [9]. The novelty of this approach is in extracting prognostic features from cardiac MRI images directly without prior segmentation while allowing the visualisation of these prognostic features. Thus, achieving a transparent way of implementing AI in cardiac MRI prognostication.

My scoping literature review in Chapter 1 identified cardiac MRI techniques that may play a role in PH assessment, such as tissue characterisation with (T_1) mapping and extracellular volume assessment. I performed a systematic review and meta-analysis of all existing studies using cardiac MRI myocardial tissue characterisation techniques in the diagnostic and prognostic assessment of PAH [10]. While (T_1) mapping was significantly elevated in PAH, only one study assessed the prognostic impact and found no strong correlation with mortality. My plan was to develop an AI method to automatically extract (T_1) mapping values. An initial model was developed, and early promising findings were presented at the European Congress of Radiology 2021 [11]. However, due to the onset of the Covid-19 pandemic and the poor prognostic signal identified in my systematic review, I had to prioritise other research, and the development of the AI (T_1) mapping tool was put on hold. In the future, the AI (T_1) mapping tool will be incorporated into a fully automated cardiac MRI assessment package. However, further development and validation are first required.

CHAPTER 1

Cardiac Magnetic Resonance in Pulmonary Hypertension

Samer Alabed[1], Pankaj Garg[1], Christopher Johns[1], Faisal Alandejani[1], Yousef Shahin[1], Krit Dwivedi[1], Hamza Zafar[1], Jim M. Wild[1], David G. Kiely[2] and Andy J. Swift[1]]

[1]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield

[2]Sheffield Pulmonary Vascular Disease Unit, Sheffield Teaching Hospitals

This chapter is based on a publication in *Current Cardiovascular Imaging Reports*. 2020 Nov;13(12):30.

Abstract

Purpose of Review

This article reviews advances in cardiac magnetic resonance (CMR) imaging in pulmonary hypertension (PH). We aim to bring the reader up-to-date with CMR applications in diagnosis, prognosis, 4D flow, strain analysis, T_1 mapping, machine learning and ongoing research.

Recent Findings

CMR volumetric and functional metrics are now established as valuable prognostic markers in PH. This imaging modality is increasingly used to assess treatment response and improves risk stratification when incorporated into PH risk scores. Emerging techniques such as myocardial T_1-mapping may play a role in the follow-up of selected patients. Myocardial strain may be used as an early marker for right and left ventricular dysfunction and a predictor for mortality. Machine learning has offered a glimpse into future possibilities.

Summary

Several recent studies have shown the diagnostic and prognostic value of CMR in patients with PH and clinical trials of PH therapies are increasingly considering it as an endpoint. Machine learning approaches to improve automation and accuracy of CMR metrics and identify imaging features of PH is an area of active research area with promising clinical utility.

1.1 Introduction

1.1.1 Definition

Pulmonary hypertension (PH) is a heterogeneous group of diseases that cause an elevated pulmonary artery pressure [12, 13]. Until recently PH was defined by an increase in the mean pulmonary artery pressure (mPAP) $\geq$ 25 mmHg measured during right heart catheterisation (RHC), however a consensus in the PH community has reduce the threshold to $\geq$ 20 mmHg to facilitate earlier detection and therefore treatment of the disease. Chronic heart and lung diseases are the underlying causes of most PH cases in the Western world [14] such as systemic hypertension and chronic obstructive pulmonary disease (COPD). Since chronic comorbid cardiac and pulmonary diseases are increasingly common in an ageing population [15–17], the prevalence of PH has als increased by almost 30% over the last 30 years [14]. PH is now estimated to have a prevalence of 1% of the world's population and might be the fourth most prevalent cardiovascular disease [16, 18]. In resource-limited countries, however, living at high-altitude and schistosomiasis infection are more still the most common cause of PH [19]. Treatable causes of PH include chronic thromboembolic disease (CTEPH) and pulmonary arterial hypertension (PAH) [20].

1.1.2 History of PH

PH is a relatively young diagnosis. The diagnostic test RHC was first performed in 1944 by Cournand and Richards [21]. Hypertensive pulmonary vascular disease cases were first described by Donald Heath in 1956, a University of Sheffield graduate [22, 23], and the disease of PH was formally acknowledged by the World Health Organisation (WHO) in 1973 [24] after a surge of cases in the 1960s secondary to the use of Aminorex, an appetite suppressant [25]. PH was historically considered as a rare and difficult disease to diagnose and until the advancement of lung transplantation in the 1980s, PH was considered untreatable [18, 26]. However, our understanding of the epidemiology, diagnosis and treatment of PH has changed over the past two decades. Particularly the management of a subtype of PH, pulmonary arterial hypertension (PAH), has revolutionised. While untreated PAH had a mortality rate close to lung cancer, advancement in therapy options over the past two decades have led to a dramatic improvement in survival [14, 27]. In 1985 CMR imaging was first used in the diagnosis of PH [28] and in 1998, the WHO

recognised the importance of PH as a global burden of disease and categorised PH into five groups (Table 1.1). In 2001, the Sheffield Pulmonary Vascular Disease Unit (SPVDU) was established and has rapidly become one of the largest PH centres in the world and by far the largest in the UK. The SPVDU currently manages one in four of the active PH patients in the UK and covers a referral population of 15 million [29]. The SPVDU since its establishment has been continuously growing and the number of patients managed has more than doubled over the last decade from just less than a 1,000 in 2010 to almost 2,200 in 2019.

Table 1.1: World Health Organisation classification of pulmonary hypertension

Group	WHO pulmonary hypertension (PH) classification
I	Pulmonary arterial hypertension (PAH)
II	PH secondary to chronic heart disease
II	PH secondary to chronic lung disease
IV	PH secondary to chronic thromboembolic disease (CTEPH)
V	PH secondary to multifactorial mechanisms

1.1.3 Pulmonary arterial hypertension

Pulmonary arterial hypertension (PAH) is one form of PH that is a primary vasculopathy of the small pulmonary arterioles, characterised by progressive vascular remodelling that starts with vascular wall hypertrophy and leads to plexiform fibrotic lesions [26, 30]. PAH gradually leads to stiffening and narrowing in the small pulmonary arteries resulting in restricted blood flow and increased pulmonary vascular resistance and pressure [31]. Over the course of its disease the elevated pulmonary arterial pressure causes the RV to dilate and hypertrophy, eventually leading to increased RV pressure and subsequent right heart failure; the main cause of death in PAH [26, 30]. In addition to a raised mPAP, PAH diagnostic criteria also include a pulmonary vascular resistance (PVR) of $\geq$ 3 Wood units and a pulmonary artery wedge pressure (PAWP) of $\leq$ 15 mmHg [20]. PAH can be idiopathic (IPAH), familial, associated with exposure to drugs and toxins or associated with underlying conditions such connective tissue disease (CTD), congenital heart disease (CHD), HIV infection or portal hypertension [20]. PAH affects on average 15 per million

people and every year up to 8 in a million people are newly diagnosed with PAH [14, 17, 32]. The mean age at diagnosis ranges between 50-65 and middle-aged women account for 60-80% of the PAH population, however, it can develop in any age group [14, 17, 32].

1.1.4 Clinical Pathway

PAH often presents with non-specific symptoms including shortness of breath, fatigue, chest pain, and dizziness, which can be seen in various other respiratory and cardiovascular conditions. In addition, the presentation can vary, with some patients having very mild symptoms that are easy to overlook. In contrast, others may have severe symptoms that lead to rapid deterioration. Because of the combination of non-specific symptoms and the heterogeneity in presentation, PAH may not be suspected initially. As a result, misdiagnosis or delayed diagnosis are common and lead to delays in the referral process. In patients with a higher risk of PAH, such as a background of connective tissue or congenital heart diseases, signs and symptoms of PH might be evaluated earlier, leading to earlier detection and treatment. Once PAH is suspected by a general practitioner, cardiologist or respiratory physician, the PH referral pathway is initiated. Typically, patients would have had initial tests such as chest X-ray, ECG, echocardiogram, or pulmonary function tests by this point. At the PH referral centre further management is based on ascertaining the diagnosis of PAH, discerning the underlying cause, assessing disease severity and monitoring response to treatment [12, 13]. Patients are examined during the first visit with blood tests, functional tests (exercise capacity and pulmonary function tests) and imaging (chest x-ray and echocardiogram). If PH is suspected, patients are admitted on a second visit for further investigations including RHC, computed tomography (CT), CMR and baseline quality of life assessment. The results of all these examinations are reviewed at a multidisciplinary meeting where management options are reviewed. Finally, targeted PH therapy is initiated where required and regular follow-up is planned with repeated investigations to assess treatment response (Figure 1.1).

This chapter reviews the current developments in CMR applications in PH including diagnosis, prognosis, 4D flow, strain analysis, T_1 mapping, machine learning and interesting ongoing research.

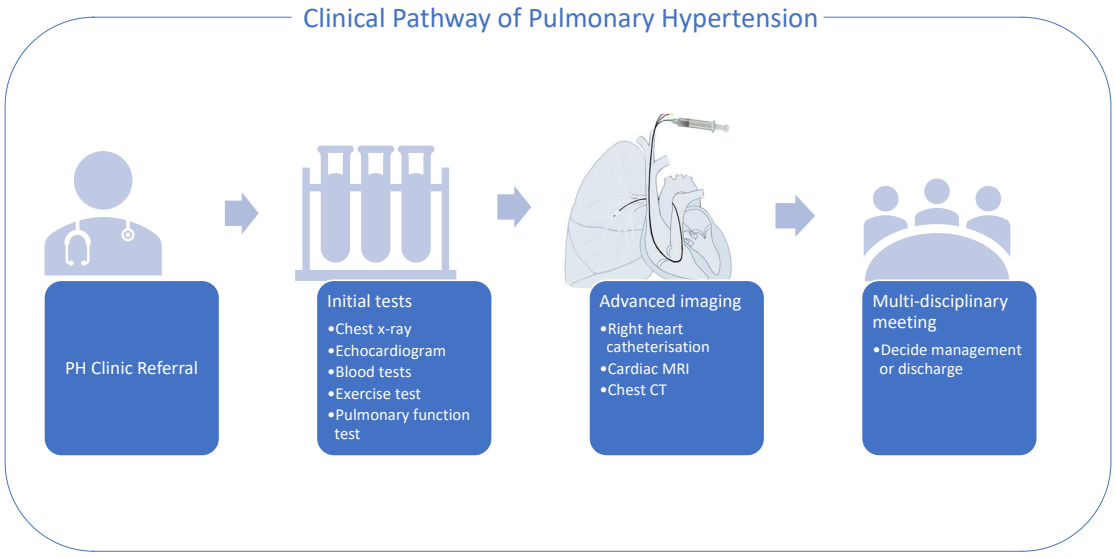


Figure 1.1: Clinical pathway of pulmonary hypertension

1.2 Diagnosis

PH is diagnosed by right heart catheter (RHC). The haemodynamic criteria for a diagnosis of PH from international guidelines are an elevated mean pulmonary artery pressure (mPAP) of $\geq$ 25 mmHg and pulmonary vascular resistance (PVR) of $\geq$ 3 Wood units [20]. A new mPAP threshold of $>$ 20 mmHg has recently been proposed as a more accurate criterion, being two standard deviations above the normal threshold [20]. In a breathless patient, PH is increasingly suggested on common investigations such as chest radiography, echocardiogram or computer tomography [12, 30]. New advances in noninvasive cardiac imaging techniques such as cardiac MRI (CMR) have helped to establish the diagnosis of PH sooner in the course of the disease [33, 34]. CMR has a spatial and temporal resolution that allows detailed visualisation of the heart and surrounding structures without exposure to radiation or invasive heart catheterisation. A wealth of quantitative information on the function and structure of the heart can be obtained after post-processing images by contouring the endo- and epicardial surfaces. This has made CMR the gold standard for quantifying cardiac chamber sizes, ventricular function and mass (Figure 1.2) [12, 35] .

CMR also allows visualisation of wall motion abnormalities and myocardial tissue characterization using myocardial T_1 or T_2 mapping and late-gadolinium contrast enhancement. It allows accurate detection of myocardial abnormalities that insult cardiac myocytes such as fibrosis by detecting accumulation of gadolinium contrast in the extracellular

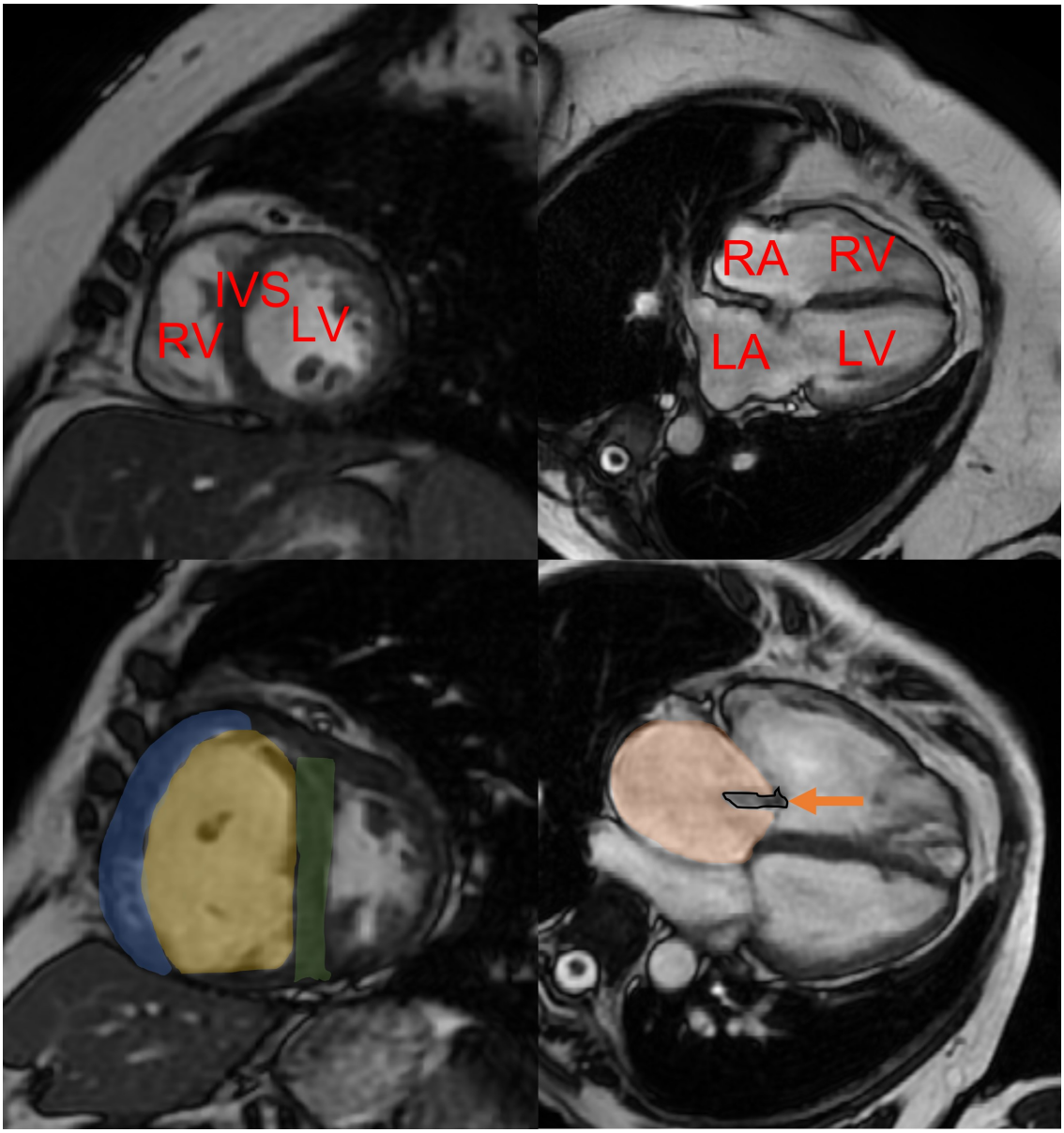


Figure 1.2: CMR images labelled with findings of PH.
Top row: Normal CMR. RV; right ventricle, LV; left ventricle, IVS; interventricular septum, RA; right atrium, LA; left atrium
Bottom row: Pulmonary arterial hypertension features including a hypertrophied RV myocardium (blue), dilated RV chamber (yellow), IVS straightening (green), RA enlargement (orange) and tricuspid regurgitation jet (arrow).

space. Velocity information through the vessels or valves can also be assessed on CMR using either black blood or phase-contrast imaging. In addition, the high reproducibility of CMR findings and measurements makes it an ideal tool for follow-up imaging and disease monitoring. Cardiac MRI has been used to estimate mPAP to provide diagnosis through noninvasive means. A recent study developed regression models to predict mPAP based on cardiac MRI [36]. A cohort of 600 patients who had both CMR and RHC was retrospectively included. The cohort was divided in half, the first half to derive the regression model and the second half to validate its performance. CMR parameters included in the linear regression model were the interventricular septum angle (IVS), ventricular mass index (VMI) and black blood slow flow. The model had a sensitivity of 93% and specificity 79% of detecting PH, allowing for an accurate non-invasive diagnosis. The model highly correlated with mPAP and had good interobserver reproducibility. The authors also recently updated their model in line with the new suggested mPAP threshold of > 20 mmHg [37]. The same group also developed a diagnostic and prognostic CMR model for patients with PH secondary to COPD [38]. Pulmonary artery (PA) indices of the diastolic area and relative area change in addition to the IVS and VMI were included in the COPD model. This model had a 92% sensitivity and 80% specificity in detecting PH in COPD patients. PA systolic and diastolic areas also had a high diagnostic accuracy in detecting PH in patients with interstitial lung disease [39]. CMR can be used as a one-stop study to provide functional, aetiological and prognostic information [40]. CMR showed a good correlation with RHC parameters and had high sensitivity and specificity in identifying the underlying causes of PH [40]. CMR-guided RHC is a recent technique that allows performing RHC in a CMR suite. It combines the benefits of radiation-free CMR information to haemodynamics in a single sitting. CMR-guided RHC has a rare failure rate and an acceptable procedure time that is comparable to a standard CMR study [41–43].

1.3 Prognosis and therapy response

PH is a chronic, progressive and mostly incurable disease with high morbidity and mortality. A new diagnosis of PH increases the risk of death at one year by sevenfold [14]. Mortality is attributed to right heart failure resulting from the increased afterload secondary to

elevated pulmonary arterial pressures [13]. PAH prognosis, has significantly improved with the advancement of treatment and the median survival has increased from three to seven years over the last 20 years [32, 44]. Despite this, the annual mortality rate of PAH remains high at 12 - 15% [14, 27]. Predicting worsening of disease and response to treatment at diagnosis is a key part of the clinical assessment in particular when discussing treatment options with patients or referring to lung transplantation. However, determining prognosis remains a major challenge and is far from straightforward [45].

The ESC/ERS guidelines describe the prognostic factors including imaging parameters in PAH [13]. Large right atrial size and the presence of pericardial effusion on echocardiogram or CMR imaging feature as prognostic markers in the ESC/ERS traffic light system representing low, intermediate and high risk. There is a wider range of prognostic factors, including the right ventricular metrics that are considered crucially important in prognostication. The prognostic role of CMR in PAH in the guidelines is undervalued and has been extensively studied over the last decade and is increasingly established in clinical practice. The ability of CMR to accurately, reproducibly and non-invasively detect changes in ventricular function and volumetric measurements, even over long periods of follow-up, makes it an ideal prognostic marker [46–48]. Further work to determine the most potent CMR prognostic factors and their incremental value in prognostic equations is required.

The most recent statement on imaging and PH from the Pulmonary Vascular Research Institute recommends CMR to monitor right ventricular (RV) function [12]. CMR can be used at follow-up to assess disease progression and treatment response and provide prognostic information. A recent systematic review, including 1,938 patients, has shown that CMR is a powerful predictor of clinical worsening and mortality in PH [4]. In particular, worse RV function and larger RV volume are associated with a worse outcome. A large study in 2017 assessed CMR prognostic features in 576 patients [46]. The study has shown that RV end-systolic volume and pulmonary artery (PA) relative area change have incremental prognostic value over clinical parameters. Pulmonary artery stiffness assessed by a low relative area change and distensibility is associated with a more severe PH and a higher risk of mortality [46, 49, 50]. The prognostic features in patients with connective tissue disease (CTD) were shown to be different to other PAH subgroups. In

CTD patients, metrics such as ventricular-vascular coupling (Ees/Ea) and RV mass appear to be more significant than function and volume [46, 51–53]. New therapies have shown to improve RV contractility and reduce RV mass in CTD PAH, and CMR can play an important role in assessing their treatment response [54, 55].

Cardiac MRI assessment of the interventricular septal (IVS) angle helps to differentiate pre- and postcapillary PH from isolated postcapillary PH. An increased angle of $\geq 160°$ is associated with pre- and postcapillary PH and with a poorer prognosis [56]. In addition, increased trabeculation at the marginal IVS is associated with severe PH, reduced RV ejection fraction (RVEF) and exercise tolerance [57, 58].

A large study has set the prognostic thresholds for CMR indices [59]. This study assessed the added value of CMR to the validated prognostic calculators such as the Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) and the modified French Pulmonary Hypertension Registry (FPHR). The age- and sex-adjusted RV end-systolic volume index improved prognostication when combined with a risk score. The prognostic thresholds will serve as an important guide for CMR risk stratification in PAH. Notably, one of the most significant predictors for a worse outcome in IPAH was a background of even minor or mild parenchymal lung abnormality. A background of mild fibrosis or emphysema was associated with a 5-year survival of 22% compared to 78% in IPAH without any lung disease [38, 60]. IPAH with lung disease has, therefore, been suggested to be a separate phenotype of PAH [61].

1.4 Other cardiac MRI applications in PH

Myocardial strain analysis

Strain analysis is an established CMR technique for the quantification of myocardial deformation and assessment of wall motion [62]. Feature tracking is one method of strain analysis which follows cardiac borders throughout the cardiac cycle on cine images (Figure 1.3). Strain analysis in CMR uses similar assumptions to speckle tracking on echocardiogram with good agreement between the two modalities [63, 64]. Biventricular strain is significantly impaired in PH and could assist in the early detection of right and left heart dysfunction [65–67]. Besides, feature tracking technology has been used to

predict outcome in patients with PH. A reduced RV circumferential and longitudinal strain rates were associated with an impaired RVEF and a significant predictor of mortality [68]. The same holds true for the left ventricle (LV) where reduced LV circumferential and longitudinal strain rates in precapillary PH are associated with severely impaired RVEF and a higher risk of death [69]. Impaired right atrial (RA) strain and phasic function are a marker of disease severity. Reduced RA strain is associated with decompensated RV function and stiffness [70, 71]. The advancement of fully automated myocardial strain analysis is likely to push its role in future research in the diagnosis and prognosis of PH [72].

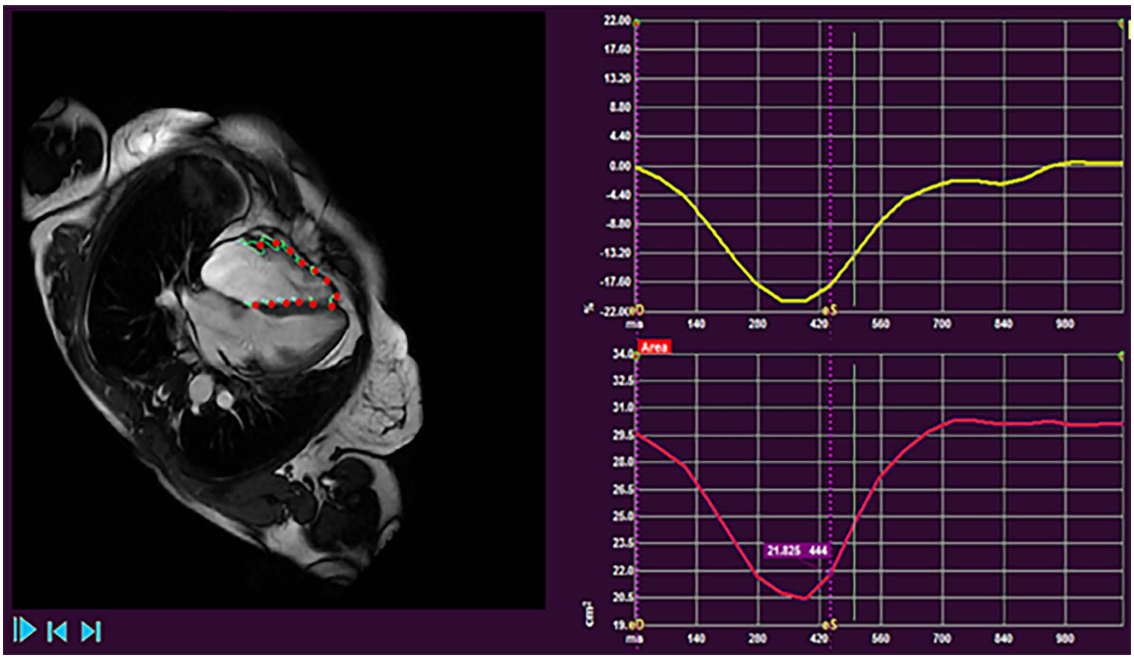


Figure 1.3: Right ventricular myocardial strain analysis in a patient with PH.

Myocardial late gadolinium enhancement

Late gadolinium enhancement (LGE) is a CMR technique to identify the areas of myocardial fibrosis. Gadolinium has paramagnetic properties that shorten the myocardial T_1 times. The T_1 shortening is proportional to the concentration of gadolinium in the extracellular space. Gadolinium enhancement in the normal myocardium clears out early. However, its clearance is restricted in necrotic tissue due to the expansion of the extracellular space and damage to the cellular membranes of the myocytes [73]. LGE is associated with poor outcome and increased mortality in cardiomyopathies [74]. However, a study assessing

LGE in 124 PH patients found that LGE did not predict mortality [75]. This finding confirms the results of two previous studies that suggested no added prognostic information from LGE in PH when added to other CMR parameters [76, 77]. Therefore, LGE in PH and particularly at the RV insertion points or IVS appears to be a consequence of increased mechanical stress and RV remodelling and not a sign of RV decompensation.

Myocardial T_1 and extracellular volume mapping

Native myocardial T_1 and extracellular volume (ECV) mapping are novel biomarkers used in several cardiovascular disorders to aid diagnostic, prognostic and therapeutic decision-making [78]. Myocardial T_1 mapping is a pixel-by-pixel representation of the longitudinal relaxation times (T_1) within a tissue [79]. T_1 values provide surrogate tissue characterisation data that are measured on a standardised scale [80]. Assessing T_1 post-gadolinium can be used to estimate ECV [78, 81]. The ECV is calculated by subtracting the T_1 values of the myocardium and blood pool pre- and post-contrast, corrected for the haematocrit level [78]. Elevated T_1 mapping values and ECV can indicate areas of oedema and fibrosis in the myocardium [82, 83]. Several recent studies have looked into the clinical application of T_1 mapping and ECV in PH [84–89]. T_1-times are elevated in PH and in particular at the RV insertion points and are associated with an increased intraventricular septal angle and LV eccentricity [84, 87]. Increased T_1 values are therefore thought to be related to RV dilatation and the resultant shift of the septum towards the LV. The diagnostic application in PH, however, remains limited. Although T_1 looked promising for differentiating between healthy volunteers and PH, the differences were much smaller in patients without a PH diagnosis in a clinical setting [84]. A raised T_1-value in PH is weakly correlated to RVEF [84, 86, 90]. However, T_1-times did not predict mortality in a large cohort of PH patients [84]. An elevated ECV in PH patients with heart failure and preserved ejection fraction was associated with RV dilatation, stiffness and reduced RV strain and therefore might play a role as a marker for RV remodelling [89]. CMR also plays a role in CTEPH treatment response assessment [91, 92]. Septal myocardial T_1 mapping is elevated in CTEPH patients [85] and reduces after treatment with balloon pulmonary angioplasty [85, 93]. T_1 mapping may, therefore, be utilised in CTEPH therapy monitoring.

Pulmonary MR angiography and perfusion

MR angiography (MRA) has a high spatial and low temporal resolution that allows for the assessment of the pulmonary vasculature. Perfusion MRI, on the other hand, has a low spatial and a high temporal resolution that enables evaluation of the capillary level tissue perfusion which makes it suitable in the clinical assessment of CTEPH [94]. The diagnostic accuracy of dynamic contrast-enhanced perfusion MRI in the diagnosis of CTEPH was shown to be comparable to computed tomography pulmonary angiography (CTPA) and perfusion single-photon emission tomography (SPECT) [95]. Perfusion MRI identified all cases of CTEPH and had comparable sensitivity and specificity to the other modalities. Recent studies have shown that ventilation and perfusion changes in CTEPH can be interrogated using phase-resolved functional lung MRI without the need for contrast agents [96]. Perfusion MRI is likely to play an essential role in the diagnostic pathway in centres that already perform CMR for CTEPH patients.

4D flow

Four-dimensional flow (4D flow) is an emerging MRI technology that offers to circumvent issues with standard ultrasound imaging in PH. 4D flow not only allows 3D visualisation of vascular flow, it also enables an accurate assessment of transvalvular or intra-cavity flow [97]. In the setting of PH, 4D flow has been used to assess the haemodynamic changes in the pulmonary circulation. Abnormal flow patterns in the main pulmonary artery (MPA), namely vortex formation, have been associated with PH [98, 99]. The presence and 'persistence time' of the vortex in the MPA are linearly associated with mPAP [100, 101] and can be used to estimate mPAP. Another physiological vascular parameter characterised with 4D flow is MPA wall shear stress (WSS), which has an impact on vascular remodelling [102]. 4D flow-derived MPA WSS appears to be reduced in patients with PH [102, 103]. Also, in patients with PH who have poor acoustic windows for echocardiography, 4D flow can provide reliable quantification of tricuspid regurgitation [104, 105]. 4D flow could also provide clinically relevant RV diastolic assessment in PH [106, 107]. To summarise, 4D flow MRI can provide several complementary diagnostic information in the assessment of patients presenting with suspected PH. Future studies need to evaluate the incremental role of 4D flow MRI assessment in patients with PH.

1.5 Machine learning

The driver of innovation in cardiac imaging technology has primarily focussed on enhancing imaging acquisition by developing faster scanning techniques, larger fields of view, stronger magnetic fields and new imaging sequences to overcome the challenges of cardiac motion and breathing during acquisition. However, the advancement of machine learning (ML) into cardiac imaging has opened up new directions of innovation, in all aspects of cardiac imaging, including image acquisition, image preprocessing, image analysis and clinical decision making.

ML uses algorithms to recognise patterns in example data to make predictive decisions in unprecedented data [108]. ML can classify the data based on the differentiating patterns it has learnt [1, 109]. Deep learning is a subset of ML, where the algorithm automatically learns multiple features in the image, for example, edges, simple shapes such as circles or lines, the arrangement of the shape in the image and its spatial relationship with other objects. The learnt data are stored in different layers of the algorithm. Each layer of these networks learns features from other layers in a complex way that resembles the neural circuits and are hence called convolutional neural networks (CNN) [110, 111]. The CNN decides which combination of features in the image most resembles the output in the example data to allow it to make decisions in future unseen data. While deep learning is effective in learning the shapes of anatomical structures and has shown to be able to detect abnormal findings on imaging, it has limitation. Humans can only see the input and output of the deep learning process but not the layers in between, an issue sometimes labelled as "black box". The CNN does not label the layers and the connections it has learnt and because of the large number of connections in each CNN, it is not always possible to explain the relationship between the input and outcome, particularly when the output is not meaningful or expected [1].

ML is likely to play an important role in PH [33]. Recent approaches include automated segmentation [112, 113], biventricular 3D model creation [114], computational models and decision tree analysis [115], diagnosis [9] and prognostication [116, 117]. ML has been used to analyse cardiac motion and predict mortality based on reduced ventricular contraction [116]. The ML model was shown to improve outcome prediction compared to

conventional CMR measurements alone. ML was used to identify diagnostic features on CMR classify them into PH or no PH [9]. The discriminating features were mapped onto CMR voxel space and were shown as a visual overlay on the 4 chamber and short-axis images (Figure 1.4). Interestingly, this approach does not require segmentation of the cardiac chambers, allowing for faster processing and reduced segmentation induced error. This study gives a glimpse into the future of PH assessment that allows for rapid and accurate CMR diagnoses. An exciting development would be to utilise ML methods in predicting prognosis and treatment response in PH.

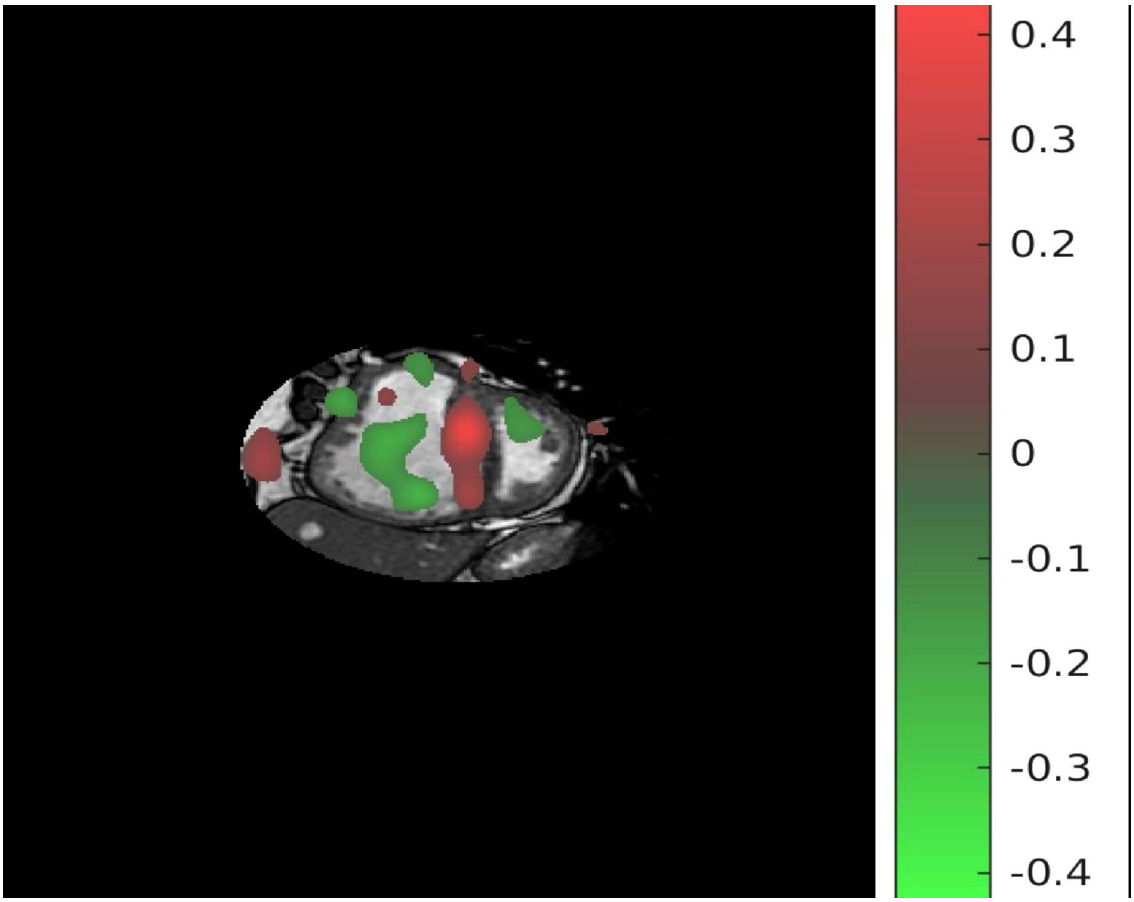


Figure 1.4: Machine learning feature map. Features compatible with PH are in red and non PH features are in green.

1.6 Ongoing research

Repeatability of CMR measurements

The RESPIRE study aims to assess the reproducibility of CMR measurements at follow-up. It will also compare the repeatability of CMR to other endpoints such as walking and

blood tests. This study would help establish the evidence of the usefulness of CMR as a monitoring tool and its sensitivity to change [118].

CMR as clinical trials endpoint

The REPAIR study is the first study to have MRI as a co-primary endpoint [119]. Four ongoing randomised controlled trials, assessing beta-blockers, spironolactone, CXA-10 and dehydroepiandrosterone, have defined CMR parameters as an endpoint to evaluate treatment response [120–123]. A single-arm study of treprostinil in PH has defined the change in RV structure and function, on CMR compared to an echocardiogram, as the primary treatment response outcome [124].

Follow-up CMR assessment

A prospective study is aiming to recruit 180 incident cases of PAH. Participants will have CMR and RHC at baseline and 6- and 24-month follow-up. The aim is to determine poor prognostic markers before decompensation occurs. This research would be valuable in early risk stratification as current studies include patients with more advanced stages of the disease [125].

CTEPH diagnosis and screening

CHANGE-MRI is a large European multicentre study that aims to compare dynamic contrast-enhanced MRI compared to VQ-SPECT in people with suspected CTEPH. This study is anticipated to set the standard for MRI in the diagnostic algorithm for CTEPH [126].

1.7 Conclusions

The last three years have seen several large studies examining the clinical utility of CMR in patients with PH. Evidence confirms the potential for CMR to provide diagnostic and prognostic information that can guide clinical practice. CMR has been utilised in clinical trials to detect the impact of PH therapies and is increasingly proposed as a trial endpoint. Machine learning approaches to improve automation and accuracy of CMR metrics and identify imaging features of PH shows potential and is likely to improve the clinical utility of CMR imaging.

1.8 Author Contributions

S.A. and A.J.S. conceived the idea and need for the review and contributed to the study conception and design. S.A. performed the literature search. S.A. evaluated the included studies and collected relevant data. Material preparation and analysis were performed by S.A.. The manuscript, figures and tables were drafted by S.A.. All authors contributed to the interpretation of data. The first draft of the manuscript was written by S.A. and all authors commented on previous versions of the manuscript. The final draft was written by S.A., A.J.S and P.G., taking into account comments and suggestions from peer reviewers and editors. All authors read and approved the final manuscript. All authors took part in the critical review and drafting of the manuscript and have read and approved the final manuscript.

CHAPTER 2

Cardiac MRI Predicts Clinical Worsening and Mortality in Pulmonary Arterial Hypertension

A Systematic Review and Meta-Analysis

Samer Alabed[1], Yousef Shahin[1], Pankaj Garg[1], Faisal Alandejani[1], Christopher Johns[1], Robert Lewis[2], Robin Condliffe[2], Jim M. Wild[1], David G. Kiely[2] and Andy J. Swift [1]

[1] Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield

[2] Sheffield Pulmonary Vascular Disease Unit, Sheffield Teaching Hospitals

This chapter is based on a publication in *JACC: Cardiovascular Imaging.* 2021 May;14(5):931-942.

Abstract

Objectives This meta-analysis evaluates assessment of pulmonary arterial hypertension (PAH), with a focus on clinical worsening and mortality.

Background Cardiac magnetic resonance (CMR) imaging has prognostic value in the assessment of patients with PAH. However, there are limited data on the prediction of clinical worsening, an important composite end-point used in PAH therapy trials.

Methods The Cochrane Central Register of Controlled Trials, MEDLINE, EMBASE, and Web of Science databases were searched in May 2020. All CMR studies assessing clinical worsening and the prognosis of patients with PAH were included. Pooled hazard ratios of univariate regression analyses for CMR measurements, for prediction of clinical worsening and mortality, were calculated.

Results Twenty-two studies with 1,938 participants were included in the meta-analysis. There were 18 clinical worsening events and 8 deaths per 100 patient-years. The pooled hazard ratios show that every 1% decrease in right ventricular (RV) ejection fraction is associated with a 4.9% increase in the risk of clinical worsening over 22 months of follow-up and a 2.1% increase in the risk of death over 54 months. For every 1 ml/m^2 increase in RV end-systolic volume index or RV end-diastolic volume index, the risk of clinical worsening increases by 1.3% and 1%, respectively, and the risk of mortality increases by 0.9% and 0.6%. Every 1 ml/m^2 decrease in left ventricular (LV) stroke volume index or LV end-diastolic volume index increased the risk of death by 2.5% and 1.8%. LV parameters were not associated with clinical worsening.

Conclusion This review confirms CMR as a powerful prognostic marker in PAH in a large cohort of patients. In addition to confirming previous observations that RV function and RV and LV volumes predict mortality, RV function and volumes also predict clinical worsening. This study provides a strong rationale for considering CMR as a clinically relevant endpoint for trials of PAH therapies.

2.1 Introduction

Pulmonary arterial hypertension (PAH) is characterised by remodelling of the distal pulmonary arteries, leading to an increase in pulmonary vasculature resistance, reduced compliance, and elevated pulmonary artery pressure [12, 13, 30]. Untreated, PAH has high morbidity and mortality that are closely linked to right ventricular dysfunction [30]. A new diagnosis of PAH increases the risk of death at one year fivefold [14]. However, over the last 20 years, treatment advancements have led to an increase in median survival from three to seven years [14, 32, 44], although development of new therapies is still needed.

Recently, clinical studies of PAH therapies have moved from assessing exercise capacity and pulmonary haemodynamics to using composite end-points. One such approach uses the time to clinical worsening. Clinical worsening events include hospitalisation, disease progression, and unsatisfactory longterm clinical response, in addition to mortality [127]. However, given the large number of patients required and the expense of conducting such event-driven studies, cardiac magnetic resonance (CMR) has recently been explored as a primary endpoint to evaluate PAH therapies [119].

CMR is the gold standard method of measuring right ventricular (RV) function, volumes, and mass, and it is an established prognostic and therapy response tool [12, 128, 129]. In 2015, a meta-analysis assessed the prognostic value of CMR measurements in five studies with 332 participants [128]; however, no data were reported on clinical worsening. Since then, multiple new studies assessing clinical worsening in addition to mortality have been published in PAH. The current meta–analysis also includes unpublished supplemental data on CMR metrics from 16 previously published studies, which allowed us to provide new data on the utility of CMR to predict clinical worsening in addition to mortality.

The goal of the current study therefore was to review the evidence for CMR metrics to predict clinical worsening and mortality in patients with PAH.

2.2 Methods

The review was prospectively registered with The International Prospective Register of Systematic Reviews (PROSPERO) on 12/12/19 (ID: CRD42019160296). The Preferred Reporting Items for Systematic reviews and Meta-Analysis guidelines (PRISMA) were followed [130]. Ethical approval was not required for this meta-analysis because it was based on published literature and did not recruit patients.

2.2.1 Criteria for considering studies for this review

Studies of all forms of PAH (including idiopathic (IPAH), heritable, drug- and toxin-induced PAH, and PAH associated with connective tissue disease (CTD); congenital heart disease; HIV infection; and portal hypertension) were considered for inclusion in the meta-analysis. For studies including patients with different forms of pulmonary hypertension (PH), data from these studies were incorporated only if the PAH cohort was separately described or the PAH participants formed at least one-half of the study population. To obtain additional data on patients with PAH, which may have been collected but not published, authors were contacted and supplemental data requested. Case reports or small cases series of <10 participants were excluded. Data collected included any clinically relevant outcomes such as hospitalisation due to heart failure, disease progression, unsatisfactory long-term clinical response, and death. To allow for analysis, only studies that reported Cox regression hazard ratios expressed per unit of measurement were included. One study reporting dichotomised hazard ratios was excluded because raw hazard ratios could not be obtained after contacting the study author.

2.2.2 Search methods for identification of studies

The following databases were systematically searched for relevant studies on 13/05/2020: Cochrane Central Register of Controlled Trials (Central) (Issue 1, May 2020), MEDLINE (ProQuest, 1946 to 6 May 2020), EMBASE (Ovid, 1974 to 2020 Week 20), Web of Science (to 13 May 2020). The reference lists of all relevant articles identified during the full-text screening were scrutinised for relevant studies.

The following search strategy was used:

1. exp ”PULMONARY HYPERTENSION”/
2. exp ”PULMONARY VASCULAR DISEASE”/
3. exp ”PULMONARY HEART DISEASE”/
4. (pulmonary ADJ2 hypertensi*)
5. (1 OR 2 OR 3 OR 4)
6. exp ”CARDIOVASCULAR MAGNETIC RESONANCE”/
7. exp MAGNETIC RESONANCE IMAGING/
8. exp CARDIAC IMAGING/
9. (MRI* OR CMR*)
10. (MR ADJ3 (imag* OR scan*))
11. (prognos* OR predict* OR clinical* OR outcome* OR associa* OR risk* OR death OR mortal* OR surviv* OR follow-up OR course OR progress* OR deteriorat*)
12. exp MORTALITY/
13. exp PROGNOSIS/
14. (11 OR 12 OR 13)
15. (5 AND 10 AND 14)

2.2.3 Data collection and analysis

Selection of studies

One author (S.A.) screened titles and abstracts and retrieved the full texts of all potentially eligible studies. The full texts were reviewed and studies meeting the inclusion criteria were included after discussion with another author (A.J.S.). The selection process was recorded in a PRISMA flow diagram.

Data extraction and management

Data extraction and risk of bias analysis were performed independently by two review authors (S.A. and F.A.), and disagreements were discussed with (A.J.S.). Data were extracted and collated data using a standardised extraction form. The methodological quality of the included studies was assessed using a modified Quality In Prognosis Studies tool (QUIPS) [131]. The corresponding authors of all included studies in the meta-analysis were contacted for hazard ratios (HRs) of unpublished CMR measurements. The unadjusted HRs were thought when bivariate, multivariate or adjusted % predicted HRs

were reported. For the meta-analysis, steps were taken to present data only for patients with PAH; where clarification was required, study authors were contacted directly.

Statistical analysis and data synthesis

HRs with 95% confidence intervals (CIs) of unadjusted univariate event-free survival regression analyses for CMR measurements including right and left ventricular ejection fraction (RVEF and LVEF), RV and LV mass index (RVMI and LVMI), RV and LV end-diastolic volume index (RVEDVI and LVEDVI), RV and LV end-systolic volume index (RVESVI and LVESVI) and RV and LV stroke volume index (RV SVI and LV SVI), were pooled. Published and unpublished data were included in all meta-analyses. Meta-analyses of HRs were conducted using Review Manager 5.4 (The Cochrane Collaboration, 2020) using a random-effect model with 95% CI. Forest plots of the baseline CMR measurements were presented using GraphPad Prism version 8.3 (GraphPad Software, La Jolla CA, USA). All tests were performed at .05 level.

Participant characteristics were presented as mean ± standard deviation (SD). If the median and ranges were reported for demographics and baseline CMR metrics, data were expressed as mean ± SD, using standard approaches [132]. Means and SDs were pooled using the formula provided in Table 7.7.a in the Cochrane Handbook [133]. Between-study heterogeneity was measured using the I^2 statistic. An I^2 >50 was considered as high, and an I^2 >30 as moderate heterogeneity. Meta-regression analyses were performed to investigate age, gender, 6-minute walking test and RHC parameters as study-level covariates on CMR measurements that had a moderate or high statistical heterogeneity. Meta-regression was performed using SPSS Statistics 26, (IBM Corp., Armonk, N.Y., USA). Publication bias was assessed graphically using funnel plots where at least ten studies were included in a meta-analysis.

2.3 Results

2.3.1 Results of the search

The comprehensive systematic literature search identified 10,939 citations (Figure 2.1). Deduplication left 8,119 citations; the majority of studies were excluded from the title and abstract screening because they (i) did not meet the inclusion criteria, primarily due

to the absence of prognostic data, (ii) did not include patients with pulmonary arterial hypertension (PAH) or (iii) were based on echocardiographic or right heart catheter (RHC) metrics and did not report magnetic resonance imaging (MRI) metrics. The full texts of 105 articles were retrieved for more detailed evaluation, of which a further 83 were excluded because they were conference abstracts, MRI findings were not described, included children, pulmonary hypertension (PH) other than PAH or did not perform univariate Cox regression analysis. A total of 22 studies that were included in the meta-analysis [46, 47, 51–53, 68, 70, 77, 117, 134–146].

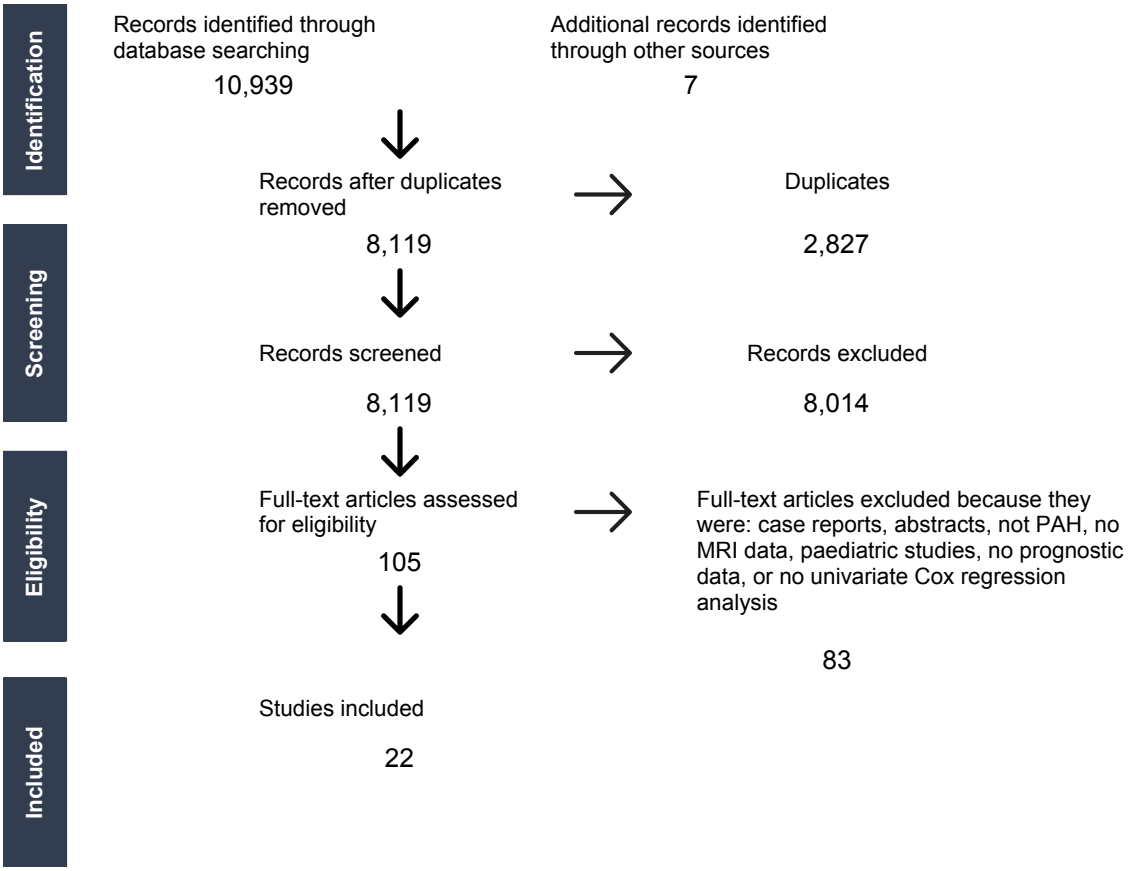


Figure 2.1: PRISMA flow diagram of the literature search.

2.3.2 Description of included studies

Study design

The review includes 14 case series and 8 case-control studies. Prospective recruitment was performed in 12 studies and consecutive inclusion was reported in 10 studies. Apart from Leng 2019 [70], all studies were single tertiary centre studies. The studies were published

between 2007 and 2020, with 16 studies including 1,606 participants published since the previous meta-analysis in 2015. Most studies (18 studies) had a small sample size of <100 patients, with the largest study by Swift 2017 [46] including 576 participants.

Population

The 22 studies were conducted in 10 different countries and included 2,149 participants. A total of 1,938 participants were included in the meta-analysis of whom 97% had PAH and 3% had other types of PH. IPAH comprised 51% and CTD-PAH 26% of the PAH population. Dawes 2018 [117], de Siqueira 2016 [68] and Jose 2019 [138] kindly provided additional data that allowed identification of patients with PAH from a mixed PH cohort.

Participants were aged 52 ± 15 years, with a female predominance (68%) and a pooled average mean pulmonary artery pressure (mPAP) of 49 ± 15 mm Hg, and RV ejection fraction (RVEF) of 37 ± 14%. Details of the included studies are presented in (Table 2.1). The pooled baseline CMR measurements are shown in (Figure 2.2).

2.3.3 Methodological quality of included studies

One-half of the studies had a prospective design, consecutive recruitment of participants, and reported blinding of CMR readers to patient clinical data. The main concern for bias is the small sample size of <100 participants in 18 of the 22 included studies. All studies were performed at PH referral centres and are therefore at risk for referral bias (Figure 2.3).

The detailed results of the quality assessment are described below:

Study participation

Seven studies were rated as a low risk of bias as they had a prospective design and recruited more than 50 consecutive patients. The remainder of the studies were assessed to have some concerns as they were retrospective studies or had a small study size of fewer than 50 patients. In retrospective studies, it is not possible to control the way data is collected and may include missing data, confounding and selective outcome reporting [147]. Jose 2019 was rated to have a high risk of bias as less than 20 participants had pulmonary arterial hypertension (PAH).

Study Author Year	Country	Design	Study period	Size	Sex F%	Age	mPAP	IPAH	CTD	CHD	Other PAH	Other PH	Follow-up (m)	Death n (%)	Clinical events n (%)
Abe 2019	Japan	CC	2008 - 2018	65	88%	56 ± 15	34 ± 11		54		11		42 (13 - 86)	9 (14%)	
Badagliacca 2016	Italy	PCS	2011 - 2013	74	59%	55 ± 13	48 ± 13	74					18 (2 - 33)		31 (42%)
Bredfelt 2018	Sweden	RCS	2003 - 2015	75	71%	57 ± 19	45 ± 11	33	33			9	28	29 (39%)	7 (9%)
Brewis 2016	UK	CC	2004 - 2014	140	66%	55 ± 16	48 ± 13	75	53	1	11		69	61 (44%)	
Dawes 2018 *	UK	CC	2004 - 2013	256	44%	65 ± 17	43 ± 16	57			31	168	48 (24 - 68)	34 (39%)	
de Siqueira 2016 *	USA	CC	2003 - 2013	93	74%	52 ± 12	40 ± 15	23	25		22	23	24 (6 - 52)		25 (36%)
Freed 2012	USA	PCS	2009 - 2010	58	74%	53 ± 14	49 ± 16	24			20	14	10 ± 6	6 (10%)	13 (22%)
Gan 2007	Holland	CC	2001 - 2005	70	79%	50 ± 15	53 ±14	49	16		5		48	18 (26%)	
Grapsa 2020	UK	PCS	N.R	30	80%	47 ± 5	N.R	30					24 (17- 24)	8 (26%)	
Jose 2019 *	USA	RCS	2013 - 2019	38	68%	51 ± 17	45 ± 15	18				20	20 (11 - 35)		4 (11%)
Kang 2013	S Korea	PCS	2009 - 2010	30	74%	45 ± 13	51 ± 23	19	2	7	2		17 ± 7	1 (3%)	6 (20%)
Knight 2015	UK	CC	2012 - 2013	40	75%	50 ± 5	46 ± 13	12	20			8	20 ± 8	1 (3%)	8 (20%)
Leng 2019	Singapore	CC	2015 - 2018	80	79%	37 ± 15	56 ± 22	21	10	40	9		24 (2-57)	6 (8%)	8 (10%)
Li 2017	China	PCS	2010 - 2013	41	71%	29 ± 9	61 ± 16	41					27 (21 - 41)	7 (17%)	10 (24%)
Mouratoglou 2018	Greece	PCS	N.R	36	78%	51 ± 14	N.R	12	7	9	2	6	20 (4-37)	0	14 (39%)
Sato 2015	Japan	PCS	2009 - 2013	68	76%	55 ± 22	37 ± 11	10	17		4	37	24 (9-34)	10 (15%)	6 (9%)
Simpson 2019	USA	PCS	2007 - 2014	64	91%	57 ± 11	N.R	22	42				50 (29 - 66)	30 (46%)	
Swift 2017	UK	RCS	2008 - 2015	576	54%	57 ± 16	48 ± 13	260	195	63	58		42 (17 - 142)	221 (38%)	
Van de Verdonk 2011	Holland	PCS	2002 - 2007	110	76%	53 ± 15	49 ± 16	73	20		17		59 (30 - 74)	30 (27%)	2 (2%)
van Wolferen 2007	Holland	PCS	1999 - 2005	64	73%	43 ± 13	56 ± 13	64					32 ± 16	19 (30%)	
Wang 2020	China	CC	2013 - 2018	100	70%	37 ± 14	62 ± 22	33	8	58	1		15 (7 - 27)	9 (9%)	21 (21%)
Yamada 2012	Japan	RCS	2003 - 2010	41	71%	39 ± 14	51 ± 14	41					44 ± 25		32 (78%)

Table 2.1: Characteristics of included prognostic studies.

Follow-up presented as mean +/− standard deviation or median (range). CC, Case-control; RCS, retrospective case series; PCS, prospective case series; mPAP, mean pulmonary artery pressure; RVEF, right ventricle ejection fraction; IPAH, idiopathic pulmonary arterial hypertension; CTD, connective tissue disease; CHD, congenital heart disease; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; n, number; N.R, not reported, ∗ only PAH patients were included in the analysis.

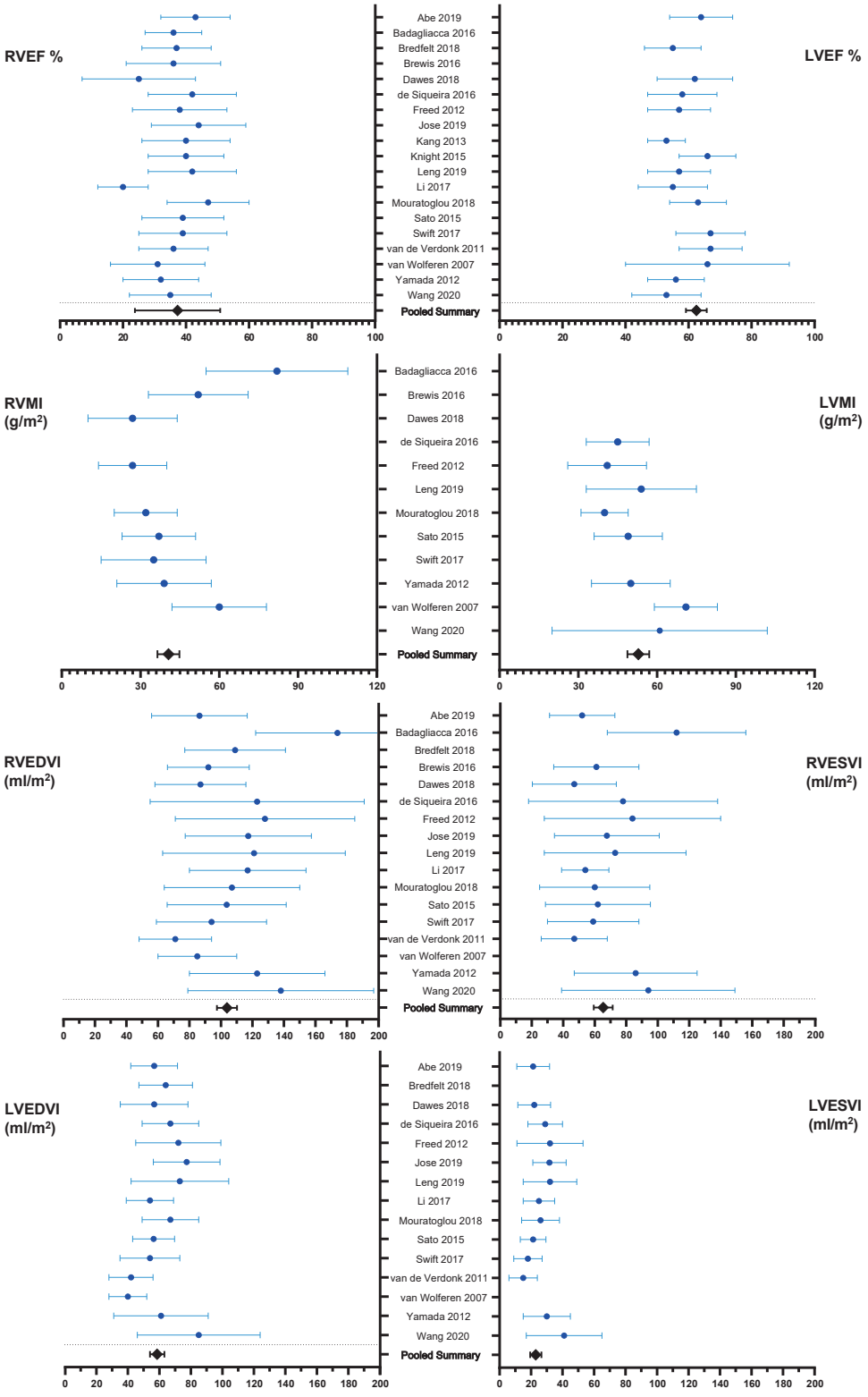


Figure 2.2: Pooled baseline CMR characteristics.

The included studies had homogenous mean baseline CMR measurements as shown by the overlapping confidence intervals, with relatively more heterogeneity in RV mass and volumes. The overall pooled mean CMR measurements show moderately impaired RV function and volumes at baseline and indicate a relatively advanced disease.
RVEF, right ventricle ejection fraction; RV, right ventricle; LV, left ventricle; EDVI, end-diastolic volume index; ESVI, end-systolic volume index; MI, mass index; SVI, stroke volume index

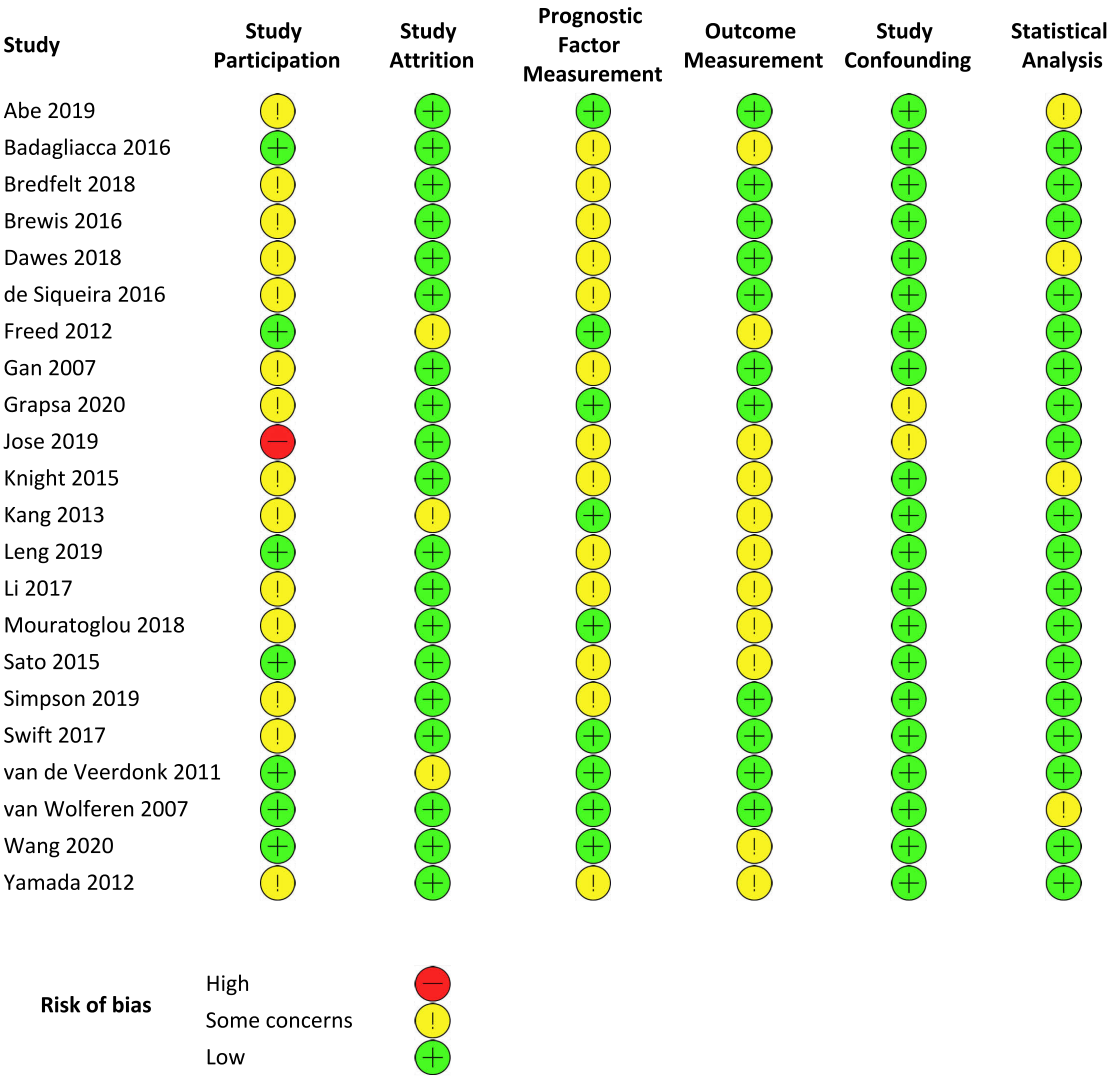


Figure 2.3: Risk of bias summary for each included study.

Study attrition

Twenty-two studies reported no loss to follow-up. Seven studies reported a loss of follow-up and explained the reason for it. Loss of follow-up of more than 10% was judged to represent some concerns for bias (Freed 2012; Kang 2013; van de Veerdonk 2011). No study had more than 20% attrition which was judged as a high risk of bias.

Prognostic factor measurement

Nine studies were judged to have a low risk of bias as they reported blinding of the cardiac MRI (CMR) assessor to clinical and outcome data. In these studies, intra-observer and interobserver variability were also analysed by having a random sample of CMRs read by a second CMR assessor. Thirteen studies did not report blinding of the CMR assessor and were judged to have some concerns for bias.

Outcome measurement

Twelve studies were judged to have a low risk of bias for outcome measurement if they either reported mortality only as the study end-point or reported blinding of the outcome assessor if they included other clinical outcomes in addition to death. Ten studies were rated to have some concern for bias as they included clinical worsening outcomes in addition but not blind the outcome assessor to patient data.

Study confounding

All included studies reported important patient characteristics and clinical measurements including age, sex, disease subtype, RHC findings and 6-min walk distance (6MWD). Nineteen studies also assessed the effects of confounding factors on the prognostic factors using multivariate regression analysis and were rated as having a low risk of bias. One study did not perform a multivariate analysis and was rated to have some concerns for confounding bias.

Statistical analysis

The statistical methods were adequately explained in all studies. In the studies included in the meta-analysis, Abe 2019 was rated as having some concerns for bias as they excluded patients from the univariate and multivariate analysis if they did not have a clinical event. All studies included in the meta-analysis presented univariate Cox proportional hazards

ratios (HR) for CMR measurements. The results of Dawes 2018, Knight 2015 and van Wolferen 2007 had extremely wide confidence intervals and very large effect sizes which are at odds with the remainder of the included studies. We rated this discrepancy to be some concern for bias in the statistical analysis of these studies.

2.3.4 Meta-analyses of CMR indices

Clinical worsening was analysed separately to mortality in a subgroup analysis. In 10 studies, providing data exclusively on mortality, 459 deaths (36%) in 1,282 participants occurred over a mean follow-up of 54 ± 5 months (8 deaths per 100 patient-years). The hazard ratios of the meta-analysis are presented in Table 2.2. A drop of 1% in RVEF increased the risk of death by 2.1%. A decrease of 1 ml/m^2 in left ventricular (LV) stroke volume index or LV end-diastolic volume index (LVEDVI) increased the risk of death by 2.5% and 1.8% respectively. An increase in RV volumes, right ventricular end-systolic volume index (RVESVI) or right ventricular end-diastolic volume index (RVEDVI), by 1 ml/m^2 increased the risk of mortality by 0.9% and 0.6%, respectively. The forest plots for RV and LV function and mass are shown in Figure 2.4 and forests plots for RV and LV volume measurements in Figure 2.5.

Table 2.2: Results of meta-analyses of hazard ratios for CMR measurements

CMR measurement	Overall meta-analysis		Mortality outcome		Clinical worsening	
	HR (95% CI)	studies (n)	HR (95% CI)	studies (n)	HR (95% CI)	studies (n)
RVEF	0.965 (0.954-0.976)	20 (1804)	0.979 (0.969-0.990)	8 (1148)	0.951 (0.939-0.964)	12 (656)
RVEDVI	1.007 (1.005-1.010)	18 (1744)	1.006 (1.003-1.008)	7 (1118)	1.010 (1.006-1.013)	11 (626)
RVESVI	1.010 (1.008-1.013)	17 (1676)	1.009 (1.005-1.012)	7 (1118)	1.013 (1.008-1.018)	10 (558)
RVSVI	0.989 (0.978-1.001)	13 (1328)	0.984 (0.965-1.004)	5 (944)	0.992 (0.979-1.004)	8 (384)
LVEF	0.992 (0.984-1.000)	15 (1561)	0.994 (0.986-1.003)	7 (1118)	0.980 (0.963-0.998)	8 (443)
LVEDVI	0.985 (0.974-0.995)	15 (1561)	0.982 (0.968-0.996)	7 (1118)	0.986 (0.969-1.004)	8 (443)
LVESVI	0.991 (0.979-1.003)	14 (1421)	0.985 (0.967-1.003)	6 (978)	0.997 (0.979-1.014)	8 (443)
LVSVI	0.976 (0.960-0.993)	11 (1344)	0.975 (0.956-0.995)	7 (1118)	0.976 (0.940-1.012)	4 (226)
RVMI	1.008 (1.001-1.016)	10 (1220)	1.006 (1.000-1.012)	5 (943)	1.018 (0.994-1.041)	5 (277)
LVMI	1.009 (0.997-1.020)	11 (1357)	1.005 (0.995-1.016)	6 (1030)	1.022 (0.991-1.053)	5 (327)

HR, hazard ratio; CI, confidence intervals, RVEF, right ventricle ejection fraction; RV, right ventricle; LV, left ventricle; EDVI, end-diastolic volume index; ESVI, end-systolic volume index; MI, mass index; SVI, stroke volume index

In 12 studies, providing data on clinical worsening, 218 (33%) events occurred in 656 participants over a mean follow-up of 22 ± 4 months (18 clinical worsening events per 100 patient-years). The composite outcome of clinical worsening included hospitalisation for

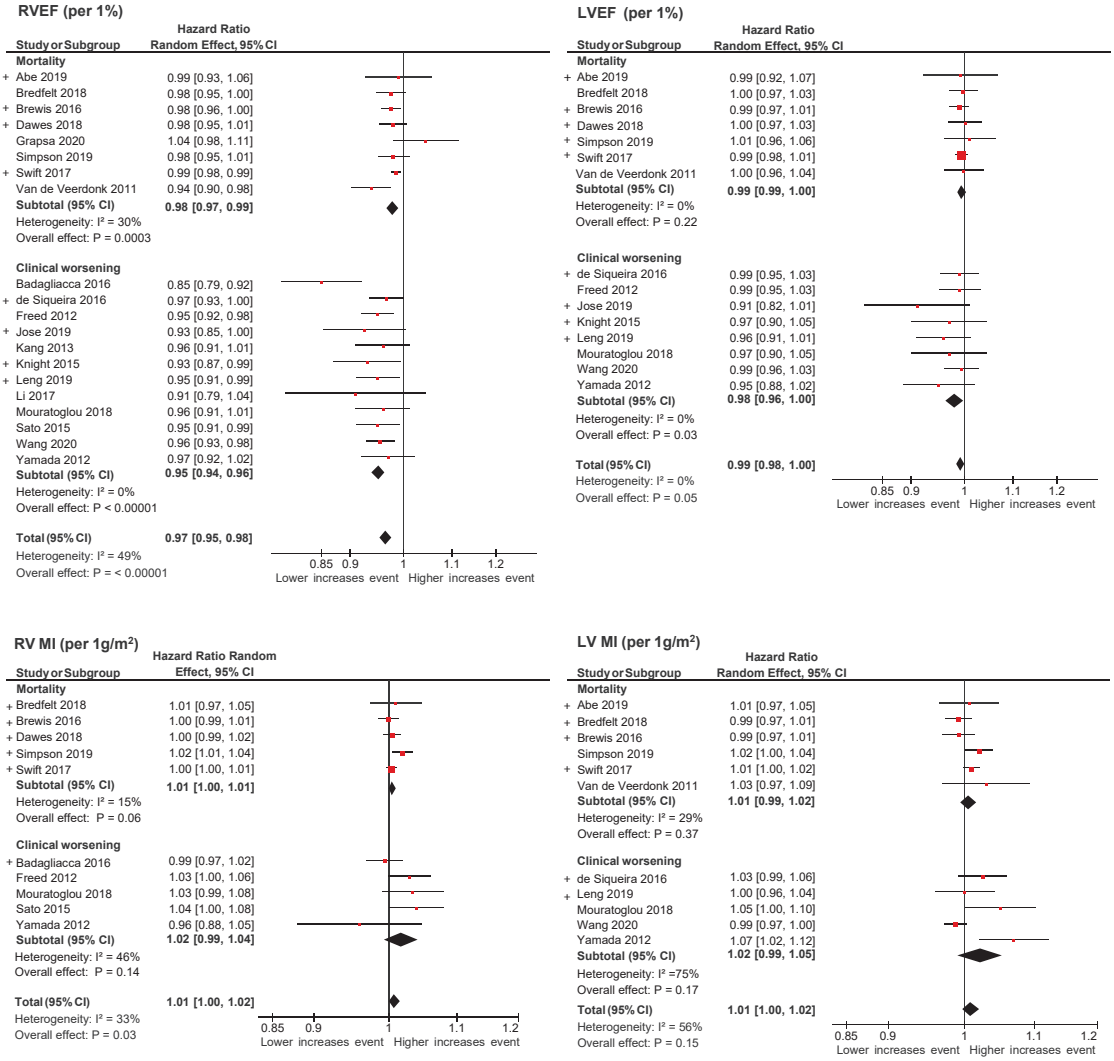


Figure 2.4: Meta-analyses of RV and LV function and mass

The meta-analyses of RV and LV function and mass showed that RVEF and RVMI are significant prognostic marker. RVEF could predict clinical worsening separate from mortality, while RVMI is a non-specific prognostic marker.
For abbreviation list see legend for Figure 2.2.
Unpublished data is indicated by (+).

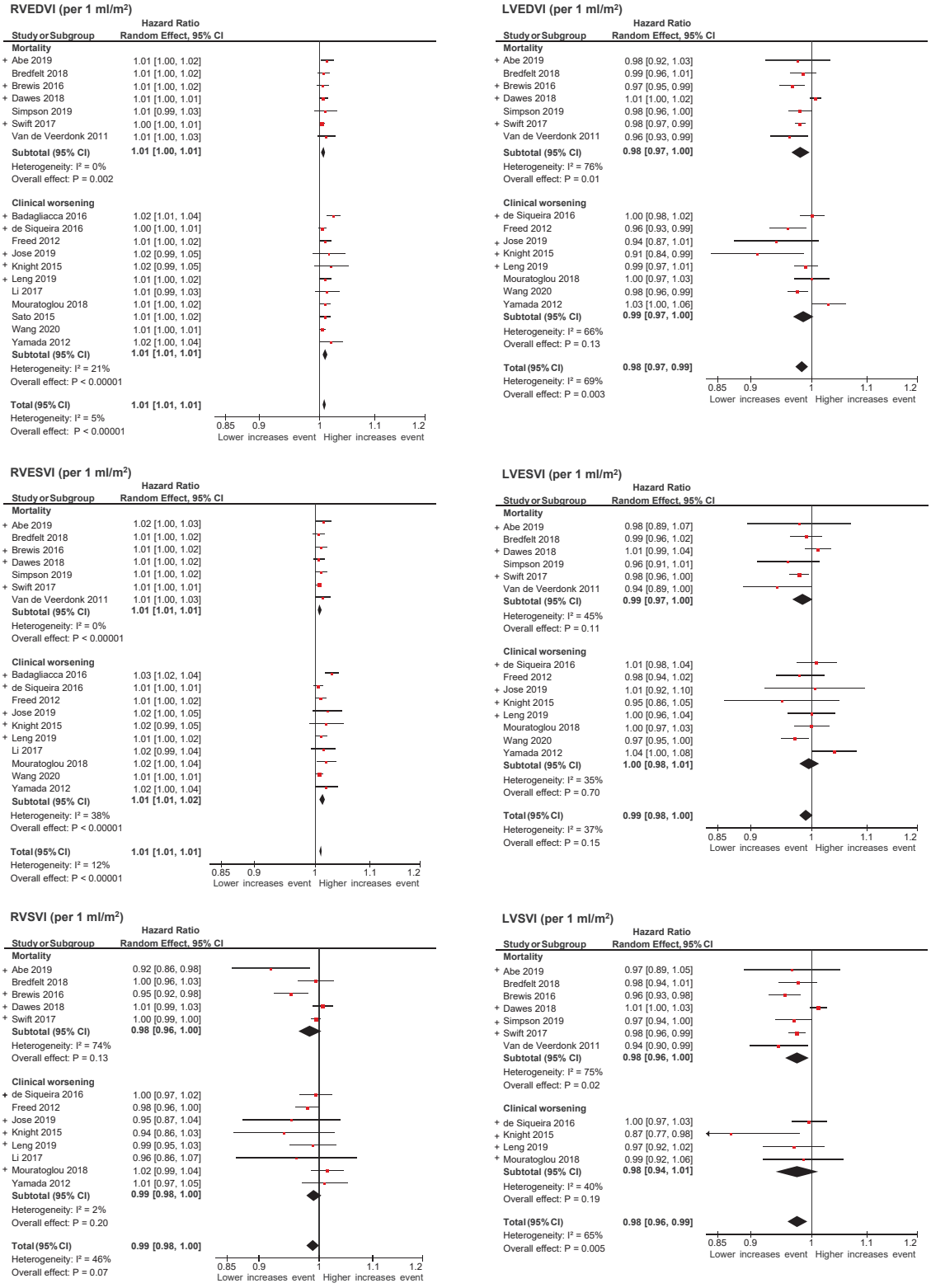


Figure 2.5: Meta-analyses of RV and LV volume measurements

RV and LV volumes are significant prognostic markers. A decrease in RV volumes can predict mortality and clinical worsening, while an increase in LV volumes indicates an increased risk for death only.
For abbreviation list see legend for Figure 2.2.
Unpublished data is indicated by (+).

heart failure (42%), escalation to prostacyclin treatment (18%), deterioration in World Health Organization (WHO) functional class (3%), a reduction in exercise capacity (2%), need for lung transplantation (2%), nonspecified aforementioned nonfatal event (14%) and all-cause death (19%). RV but not LV volumetric and functional measurements predicted clinical worsening. A 1% deterioration in RVEF was associated with a 4.9% increase in the risk of clinical worsening whilst a 1 ml/m^2 increase in RVESVI or RVEDVI was associated with an increase of clinical worsening of 1.3% and 1%, respectively.

Additional meta-analyses

Right atrial measurements

Right atrial (RA) volume is not a significant prognostic marker (HR 1.01, 95% CI 1.00 to 1.02; $P = 0.09$; participants = 201; studies = 3). RA area index is a significant prognostic marker (HR 1.08, 95% CI 1.04 to 1.12; $P < 0.001$; participants = 106; studies = 2).

Pulmonary artery measurements

Surrogate markers for pulmonary artery (PA) stiffness such as relative area change (RAC) and PA distensibility were reported in two studies only [46, 136]. The meta-analysis showed that PA distensibility was prognostically significant but PA RAC did not reach statistical significance. Although PA RAC was significant in each study individually, their pooled result was imprecise and had wide confidence intervals.The PA RAC meta-analysis result was HR 0.92, 95% CI 0.85 to 1.00; $P = 0.05$; participants = 646; studies = 2.

Stroke volume / RV end-systolic volume

In CTD-PAH, PA RAC and load-independent measurements such as RV-PA coupling measurements were significantly prognostic [46, 51, 148]. Four studies, including 845 participants, reported the stroke volume divided by RV end-systolic volume to estimate RV elastance (Ees) divided by PA elastance (Ea). The pooled CMR Ees / Ea ratio is a significant prognostic marker (HR 0.47, 95% CI 0.33 to 0.68; $P < 0.001$).

Heterogeneity

There was high statistical heterogeneity in the overall result of LV mass index, LVSVI and LVEDVI, and moderate heterogeneity in LV endsystolic volume index, RVEF, and RV stroke volume index. A meta-regression model of the logHR of these variables showed no evidence of a linear relationship with age, male sex, 6-min walking test, or right heart

catheterisation parameters (cardiac index, mean right atrial pressure and pulmonary vascular resistance) (Table 2.3). There was not enough data to perform a meta-regression for functional class, disease stage or treatment status. There may, however, be sources of heterogeneity that could not be assessed in a meta-regression analysis where not enough data were available. There were differences in the types of clinical worsening events used as endpoints, length of follow-up, and the subgroups of PAH studied. Other causes of heterogeneity may include variation in baseline CMR measurements (Figure 2.2), disease severity and treatment status. There is also geographical variation; eleven studies were from European centres, four from the USA and seven from Japan, South Korea, China and Singapore.

Table 2.3: Meta-regression of CMR measurements with moderate to high statistical heterogeneity

CMR measurement	Covariate coefficient (P value)					
	Age	Male gender	6MWD	CI	mRAP	PVR
RVEF	0.40 (0.08)	-0.11 (0.64)	-0.45 (0.10)	0.13 (0.64)	0.44 (0.09)	0.52 (0.10)
RVEDVI	-0.04 (0.88)	-0.18 (0.46)	0.01 (0.98)	-0.17 (0.56)	-0.29 (0.29)	-0.42 (0.20)
RVSVI	-0.02 (0.95)	0.23 (0.44)	insufficient data	0.39 (0.27)	0.29 (0.41)	insufficient data
LVEDVI	-0.24 (0.40)	-0.13 (0.63)	0.06 (0.86)	0.07 (0.82)	-0.25 (0.43)	insufficient data
LVESVI	-0.206 (0.50)	-0.32 (0.26)	-0.39 (0.27)	-0.20 (0.58)	-0.41 (0.24)	insufficient data
LVMI	0.28 (0.39)	-0.49 (0.11)	insufficient data	0.35 (0.36)	0.27 (0.48)	insufficient data

6MWD, 6-minute walking distance; mRAP, mean right atrial pressure; CI, cardiac index; PVR, pulmonary vascular resistance (Wood units)

Publication bias

Most studies included in the meta-analysis were contacted for unpublished data to reduce the risk of publication bias. The study authors of 13 included studies kindly replied to our requests for additional data where they published results for a mixed PH cohort (Dawes 2018 [117], de Siqueira 2016 [68], Jose 2019 [138]), bivariate or multivariate hazard ratios (HRs) (Brewis 2016), adjusted % predict HRs (Swift 2017), non-indexed volumetric measurements (Badagliacca 2016) or reported a subset of CMR indices (Abe 2019 [51], Bredfelt 2018 [135], Knight 2015 [140], Simpson 2019 [52], Leng 2019 [70], Mouratoglou 2018 [142], Van de Verdonk 2011 [47]). The authors provided us with PAH results only, univariate and non-adjusted HRs, indexed volumetric measurements and the results for additional CMR metrics. Five studies published prior to 2015 reporting only a few CMR

metrics did not respond to our requests for additional information. The results of Van Wolferen 2007 ([145]) had very large effect sizes and standard errors, following discussion with the co-authors we understand this is due to scaling of the CMR measurements to the standard deviation rather than the unit of measurement. We therefore decided not to pool the results of Van Wolferen 2007 with the rest of the studies due to the different unit of scaling used.

Publication bias was assessed graphically using funnel plots where at least ten studies were included in a meta-analysis (Figure 2.6). The funnel plots of RVEF, LVEF, RVMI and LVEDVI showed minor asymmetry which may indicate that a small study with extreme effect sizes was not published.

2.4 Discussion

To the best of our knowledge, this study is the largest meta-analysis of CMR imaging in patients with PAH and the first to report on clinical worsening in addition to mortality. We have confirmed that CMR imaging is a powerful prognostic marker in a large cohort of patients from multiple institutions, across several continents and using different imaging platforms. In addition, we have shown that CMR imaging predicts clinical worsening in patients with PAH. Our findings highlight the clinical utility of CMR imaging and support further evaluation of this modality as a clinically meaningful trial endpoint for the assessment of new therapies for PAH (Table 2.4).

Clinical worsening as a composite endpoint has been shown to predict mortality [149] and has established itself as a primary efficacy endpoint in trials of PAH therapies [127, 150]. Although heart failure and all-cause mortality are included in all PAH trials using a composite clinical worsening endpoint, these trials vary in their inclusion and definition of progression markers, such as change in exercise tolerance or functional capacity [151], and they may use different thresholds to define a meaningful change [152]. Nonetheless, study designs using time to clinical worsening have been increasingly adopted to evaluate PAH therapies. However, such studies require large numbers of participants and a prolonged period of follow-up, usually lasting for several years. As a consequence and given recent events such as the coronavirus disease 2019 pandemic, there has been a focus on considering

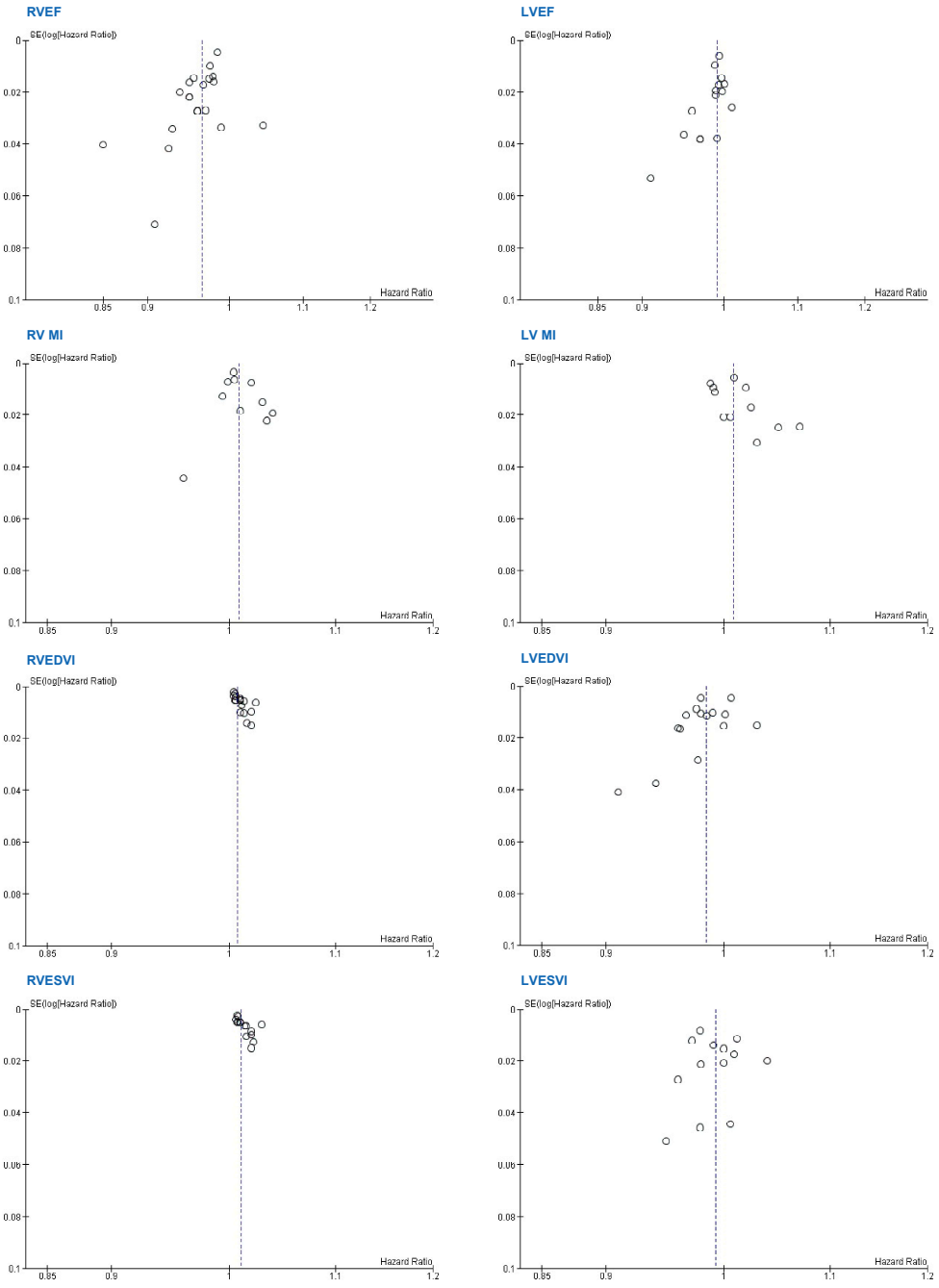


Figure 2.6: Funnel plots of the meta-analyses of CMR measurements

Table 2.4: Summary of findings table

Review question	What are the CMR predictors for clinical worsening and mortality in patients with PAH?		
Population	1,938 participants, including 68% females, aged 52 ± 15 years. Participants had more advanced disease and intermediate to high risk for one-year mortality.		
Follow-up	22 ± 4 months for clinical worsening and 54 ± 5 months for mortality		
Setting	Tertiary pulmonary hypertension referral centres		
Studies	Case series and case-control studies		
Quality of evidence	Some concerns for bias due to small sample sizes, retrospective design, lack of blinding in most studies and non-consecutive inclusion in half of the studies.		
Results	**Increment**	**Clinical worsening (over 22 months)**	**Mortality risk (over 54 months)**
RVEF	per 1% decrease	4.9% increase	2.1% increase
RVESVI	per 1 ml/m^2 increase	1.3% increase	0.9% increase
RVEDVI	per 1 ml/m^2 increase	1% increase	0.6% increase
LVSVI	per 1 ml/m^2 decrease	Not significant	2.5% increase
LVEDVI	per 1 ml/m^2 decrease	Not significant	1.8% increase

CMR, Cardiac MRI; PAH, pulmonary arterial hypertension; RVEF, right ventricle ejection fraction; RV, right ventricle; LV, left ventricle; EDVI, end-diastolic volume index; ESVI, end-systolic volume index; SVI, stroke volume index

clinical trial endpoints, which allow the impact of candidate therapies to be assessed over a shorter period and using an endpoint that correlates with clinically meaningful events.

In this meta-analysis, we have shown for the first time in a large cohort of patients that CMR-derived RV volumetric and functional metrics but not LV measurements predict clinical worsening. This information should be helpful to regulatory authorities who are keen to ensure that proposed trial endpoints have clinical relevance. In addition, this meta-analysis confirms the prognostic value of CMR metrics in a substantial cohort of patients, which has allowed an assessment of the impact of change on specific metrics concerning clinical worsening, including mortality. A 1% decrease in RVEF is associated with a 4.9% increase in the risk of clinical worsening and a 2.1% increase in the risk of death. In addition a 1 ml/m^2 increase in indexed RV volumes is associated with a 0.6% to 0.9% increase in the risk of mortality and 1% to 1.3% increase in the risk of clinical worsening. Although this incremental change in RV volumes is smaller than the 1.8% to 2.5% associated with a 1 ml/m^2 decrease in LV volume, the overall risk of mortality is more linked to RV volume, previously highlighted in large cohort studies [46, 59]. Specifically, the increase in RV volume due to dilation in response to an increase in afterload is substantially larger than the change in LV volume, occurring as a consequence of ventricular interaction [48]. Particularly, in advanced disease in PAH when uncoupling of the RV and its load occurs [153].

This meta-analysis has shown that an increased RV mass has prognostic value but does not predict clinical worsening. In PAH, an increase in RV mass and RV hypertrophy is likely to represent an appropriate adaptive response to an increase in afterload [154]. In a CMR study in which patients were monitored over five years, RV wall thickness was not associated with increased mortality in patients who were judged to be clinically stable [48]. Moreover, a disproportionate increase in right ventricular mass index (RVMI) compared with RVEDV indicates concentric hypertrophy and is associated with a favourable outcome in IPAH [134]. Eccentric hypertrophy with a disproportionate increase in RVEDV compared to RVMI is considered a maladaptive response to increased afterload and is associated with a poor outcome [134, 154]. In IPAH, therefore, caution should be exercised when using mass

measurements in isolation because they give incomplete information on RV adaptation. Further study of the relationship between RV mass and volume would be helpful. In CTD-PAH, in which the natural history of the disease is different, RVMI and ventricular mass index (VMI) appear to have greater prognostic value than RV function or volumes [77, 117, 138, 148]. A 10% increase in RVMI and VMI was associated with an increased risk of death of 10% - 15% [52]. A VMI ≥ 0.7 was associated with 35% mortality at one year and 67% mortality at two years [148]. RV hypertrophy in CTD-PAH may be an early prognostic marker for mortality rather than just an adaptive response to PH [52]. This finding emphasises the importance of considering the clinical context when using tools to assess prognosis.

Several additional CMR measurements have been shown in small studies to have prognostic value mentioned in the "Additional meta-analyses". These analyses include right atrial (RA) volume and area, pulmonary artery relative area change (RAC) and distensibility and the ratio of stroke volume / RV end-systolic volume (CMR Ees / Ea). Although the European Society of Cardiologists and European Respiratory Society guidelines include RA size and the presence of pericardial effusion on echocardiogram as prognostic factors in PAH, there is only limited data on these metrics as a prognostic marker using CMR [13]. RA volume on CMR has been evaluated in two studies only. The meta-analysis of their results did not reach statistical significance; however, the direction of effect suggests that an increased RA volume is associated with a poor outcome. Further studies assessing CMR RA volume and its relation to clinical worsening and prognosis in PAH prognosis are needed. Two additional RA CMR markers that have been reported to be prognostically significant but could not be studied in the meta-analysis as only one study reported each metric; these were RA emptying fraction and RA strain [70, 155]. RA emptying fraction is the difference in the RA max - RA min volume divided by RA max volume and RA strain was calculated as the distance between the posterior RA wall and the atrioventricular junction on the 4-chamber view. It is worth noting that each study assessing RA area and volume used different measuring methods. RA was measured either on short-axis stacks, 4-chamber views or axial cine images covering the whole heart. Measuring the RA on short-axis images can be misleading and the most accurate way of

capturing the entire atria would be to assess the axial images. Despite these difference in the RA area and volume measurements between studies, they all agreed that RA size is a marker of poor prognosis in PAH.

Additional CMR prognostic markers that were not assessed in the meta-analysis include myocardial strain analysis, myocardial T_1 mapping and late gadolinium enhancement. The small number of studies reporting these markers or the absence of Cox regression analysis prevented a meaningful pooling of their results. Strain analysis using feature tracking seems to be a promising prognostic marker [68–70]; however, it needs to be evaluated further in a more extensive survival study. Late gadolinium enhancement [75–77] and T_1-mapping [84] have an unclear additive prognostic value in PAH.

The meta-analysis is based on a population likely to have disease at the more severe end of the spectrum. The results may therefore not be generalisable to patients with more modest disease, in which age and comorbidity may have more of an impact on prognosis. A recently published, large well-designed study has shown that CMR could be used to establish thresholds for mortality risk in PAH [59]. This study showed that CMR metrics can be used to improve risk stratification when incorporated into the French Registry approach or REVEAL (Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management) risk scores [156, 157]. In the study by Lewis et al. [59], an RVEF <37%, an RVESVI of >54 ml/m^2 and LVEDVI of <52 ml/m^2 were associated with a high risk of mortality. In this meta-analysis, the pooled RVEF was 37%, RVESVI 63 ml/m^2 and LVEDVI 57 ml/m^2. All included studies used the current guideline criteria of an mPAP threshold of $\geq$ 25 mmHg. A new threshold of $>$ 20 mm Hg with a pulmonary vascular resistance $\geq$ 3 Wood units has recently been proposed as a haemodynamic definition for PAH, being two standard deviations above the normal threshold [20]. There remains a lack of evidence therefore regarding the prognostic value of CMR in patients who have modest PAH and those with mPAP $>$ 20mm Hg, in whom other factors may be more of a driver to clinical worsening and mortality.

2.4.1 Limitations

An extensive systematic literature search was performed and a pre-published protocol was followed. However, a potential limitation of this study is that inclusion was assessed by a single investigator. Any doubt regarding study selection was discussed with another investigator, however. This meta-analysis contains previously unpublished data for participants included in previously published studies. However, this approach has allowed improved data completeness and additional analysis. The included studies including supplemental data are indicated by (+) in Figure 2.4 and Figure 2.5. Patients in this study included a cohort with a predominantly intermediate and high risk of one-year mortality and likely represent a cohort with more severe disease. Although the results of the meta-analysis suggest that CMR imaging, as performed in expert centres, strongly associates with outcomes, some caution is warranted in its application in less-experienced centres given the limited existence of multicentre studies. In some instances, heterogeneity is high, and greater caution in interpretation is therefore indicated. Finally, only limited data provided on the potential of CMR metrics such as myocardial strain analysis, RA size, PA wall stiffness and four-dimensional flow parameters and the application of artificial intelligence approaches to large imaging data sets, which may add clinical value.

2.5 Conclusion

Clinical worsening is an important composite endpoint used in therapy trials in PAH. This meta-analysis is the first to study the role of CMR in the prediction of clinical worsening in PAH. In a meta-analysis of a large cohort with PAH, we showed that CMR predicts clinical worsening in addition to confirming its prognostic value. In a metaregression, we have also shown that CMR predicts clinical worsening and mortality independent of age, sex, pulmonary hemodynamics, and walking distance. This study provides further data supporting the clinical utility of CMR in patients with PAH. The findings of this meta-analysis provide a strong rationale for future research to consider CMR as a clinically relevant endpoint for therapy trials in PAH.

2.6 Acknowledgements

We are very grateful for the precious support and collaboration of the following people for kindly providing unpublished data: Dr Nobuya Abe and Dr Masaru Kato (Abe 2019), Prof Roberto Badagliacca and Prof Carmine Dario Vizza (Badagliacca 2016), Dr Melanie Brewis (Brewis 2018), Dr Ellen Ostenfeld (Bredfelt 2018), Dr Tim Dawes and Dr Declan O'Regan (Dawes 2018), Dr Maria Eduarda Menezes de Siqueira and Prof Adriano Mendes Caixeta (de Siqueira 2016), Dr Dan Knight and Prof Vivek Muthurangu (Knight 2015), Dr Liam Zhong (Leng 2019), Dr Sophia Mouratoglou and Dr George Giannakoulas (Mouratoglou 2018), Dr Arun Jose and Dr Jean Elwing (Jose 2019), Dr Catherine Simpson and Prof Paul Hassoun (Simpson 2019), Dr Andy Swift (Swift 2017) and Dr Mariëlle van de Veerdonk and Prof Anton Vonk Noordegraaf (Van de Veerdonk 2011). We thank Dr Maria Sukhanenko for her help with extracting data from a Russian study.

2.7 Author Contributions

S.A. and A.J.S. conceived the idea and need for the systematic review and contributed to the study conception and design. Protocol registration in PROSPERO was performed by S.A.. S.A. created the search strategy and performed the literature search. S.A. performed screening and eligibility assessments. S.A. evaluated the included studies and collected relevant data. Material preparation and analysis were performed by S.A.. The manuscript, figures and tables were drafted by S.A.. All authors contributed to the interpretation of data. The first draft of the manuscript was written by S.A. and all authors commented on previous versions of the manuscript. The final draft was written by S.A., A.J.S and D.G.K., taking into account comments and suggestions from peer reviewers and editors. All authors read and approved the final manuscript. All authors took part in the critical review and drafting of the manuscript and have read and approved the final manuscript.

CHAPTER 3

Quality of Reporting in AI Cardiac MRI Segmentation Studies

A Systematic Review and Recommendations for Future Studies

Samer Alabed[1,2], Ahmed Maiter[1,2], Mahan Salehi[1], Aqeeb Mahmood[3], Sonali Daniel[3], Sam Jenkins[3], Marcus Goodlad[1], Michael Sharkey[1,2], Michail Mamalakis[4],Vera Rakocevic[3], Krit Dwivedi[1,2], Hosamadin Assadi[5], James M. Wild[1], Declan P. O'Regan[6], Haiping Lu[4], Rob van der Geest[7], Pankaj Garg[5], Andrew J. Swift[1,2]]

[1] Infection, Immunity & Cardiovascular Disease, University of Sheffield

[2] Department of Clinical Radiology, Sheffield Teaching Hospitals

[3] Medical School, University of Sheffield

[4] Department of Computer Science, University of Sheffield

[5] University of East Anglia, Norwich Medical School

[6] MRC London Institute of Medical Sciences, Imperial College London

[7] Leiden University Medical Center, Leiden, the Netherlands

This chapter is based on a publication in *Frontiers in Cardiovascular Medicine.* 2022 July;9:956811.

Abstract

Objectives This systematic review aimed to evaluate the quality of reporting in Artificial Intelligence (AI) studies of cardiac MRI (CMR) segmentation.

Background There has been a rapid increase in the number of AI studies of CMR segmentation aiming to automate image analysis. However, advancement and clinical translation in this field depend on researchers presenting their work in a transparent and reproducible manner.

Methods MEDLINE and EMBASE were searched for AI CMR segmentation studies in April 2022. Any fully automated AI method for segmentation of cardiac chambers, myocardium or scar on CMR was considered for inclusion. For each study, compliance with the Checklist for Artificial Intelligence in Medical Imaging (CLAIM) was assessed. The CLAIM criteria were grouped into study, dataset, model and performance description domains.

Results 209 studies published between 2012 and 2022 were included in the analysis. Studies were mainly published in technical journals (58%), with the majority (57%) published since 2019. Studies were from 37 different countries, with most from China (26%), the United States (18%) and the United Kingdom (11%). Short axis CMR images were most frequently used (70%), with the left ventricle the most commonly segmented cardiac structure (49%). Median compliance of studies with CLAIM was 67% (IQR $59-73\%$). Median compliance was highest for the model description domain (100%, IQR $80-100\%$) and lower for the study (71%, IQR $63-86\%$), dataset (63%, IQR $50-67\%$) and performance (60%, IQR $50-70\%$) description domains.

Conclusion This systematic review highlights important gaps in the literature of CMR studies using AI. We identified key items missing - notably poor description of patients included in the training and validation of AI models and inadequate model failure analysis - that limit the transparency, reproducibility and hence validity of published AI studies. This review may support closer adherence to established frameworks for reporting standards and presents recommendations for improving the quality of reporting in this field.

3.1 Introduction

Cardiac MRI (CMR) is the gold standard for non-invasive assessment of cardiac structures. Quantitative measurement of cardiac volumes can be achieved with CMR and relies on accurate segmentation of structures on CMR images. Manual segmentation is routinely performed by cardiac imaging experts but suffers from a number of drawbacks. In addition to being laborious and time-intensive, manual segmentation is operator-dependent, potentially impacting interobserver agreement. As the demand for cardiac imaging continues to grow and outpaces the supply of trained readers, there is an increasing need for automation [1, 158].

Artificial intelligence (AI) is changing medical imaging through the automation of complex and repetitive tasks, including the segmentation of anatomical structures [159]. Machine learning is a subfield of AI that is commonly used for image analysis and processing in medical applications. Machine learning algorithms learn by experience, typically in a supervised manner: the algorithm is trained on labelled data, such as a set of manually segmented CMR images, where the manual segmentation provides the reference standard or ground truth. The algorithm identifies discriminative features and patterns in this image data, which are incorporated to generate a model that can perform the task - such as segmentation of the cardiac chambers - on new unlabelled data without the need for explicit programming. Machine learning itself encompasses a diverse range of techniques, including deep learning, which can be applied to the segmentation of structures in imaging [160].

A growing number of studies have reported the use of AI methods for segmentation in CMR. The manner in which these studies are reported is important. Transparent reporting of methods and results facilitates reproducibility and allows proper evaluation of validity. Equally, a consistent standard of reporting aids comparison between studies and may improve accessibility of the literature, which may be of particular benefit in a rapidly expanding field such as AI. The need for consistency in reporting medical research

is well recognised and reflected in various guidelines and checklists for different study types. The Checklist for Artificial Intelligence in Medical Imaging (CLAIM), [161] has adopted the validated and widely used Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines and incorporated domains specific to AI studies, including detailed descriptions of data sources, model design and performance evaluation. This systematic review aimed to evaluate the quality of reporting of studies involving AI CMR segmentation by assessing compliance with CLAIM.

3.2 Methods

The study protocol was registered with The International Prospective Register of Systematic Reviews (PROSPERO; registry number CRD42022279214). The study was undertaken and is presented in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines [162]. No ethical approval was required.

3.2.1 Inclusion and exclusion criteria

Studies reporting the use of AI for segmentation of structures in CMR were considered for inclusion. Studies were deemed eligible if they reported: (1) any type of fully automated AI method (including machine learning, deep learning and neural networks), (2) segmentation of cardiac chambers, myocardium or scar tissue and (3) use of adult human CMR images, regardless of acquisition methods (such as use of intravenous contrast), parameters, post-processing methods and software. Exclusion criteria were as follows: absence of a newly developed segmentation model (e.g. studies assessing existing methods), use of semi-automated AI methods (where the segmentation process required manual input), multiorgan segmentation, combined segmentation of multiple imaging modalities (e.g. CMR and CT), segmentation of cardiac vessels (e.g. aorta, pulmonary artery, coronary arteries) or pericardial tissue, use of non-human or ex-vivo images, and conference publications. Non-English language publications were excluded. Figure 3.1 shows an example of automatic biventricular [7] (Figure 3.1A) and four-chamber [163] (Figure 3.1B) segmentation on CMR.

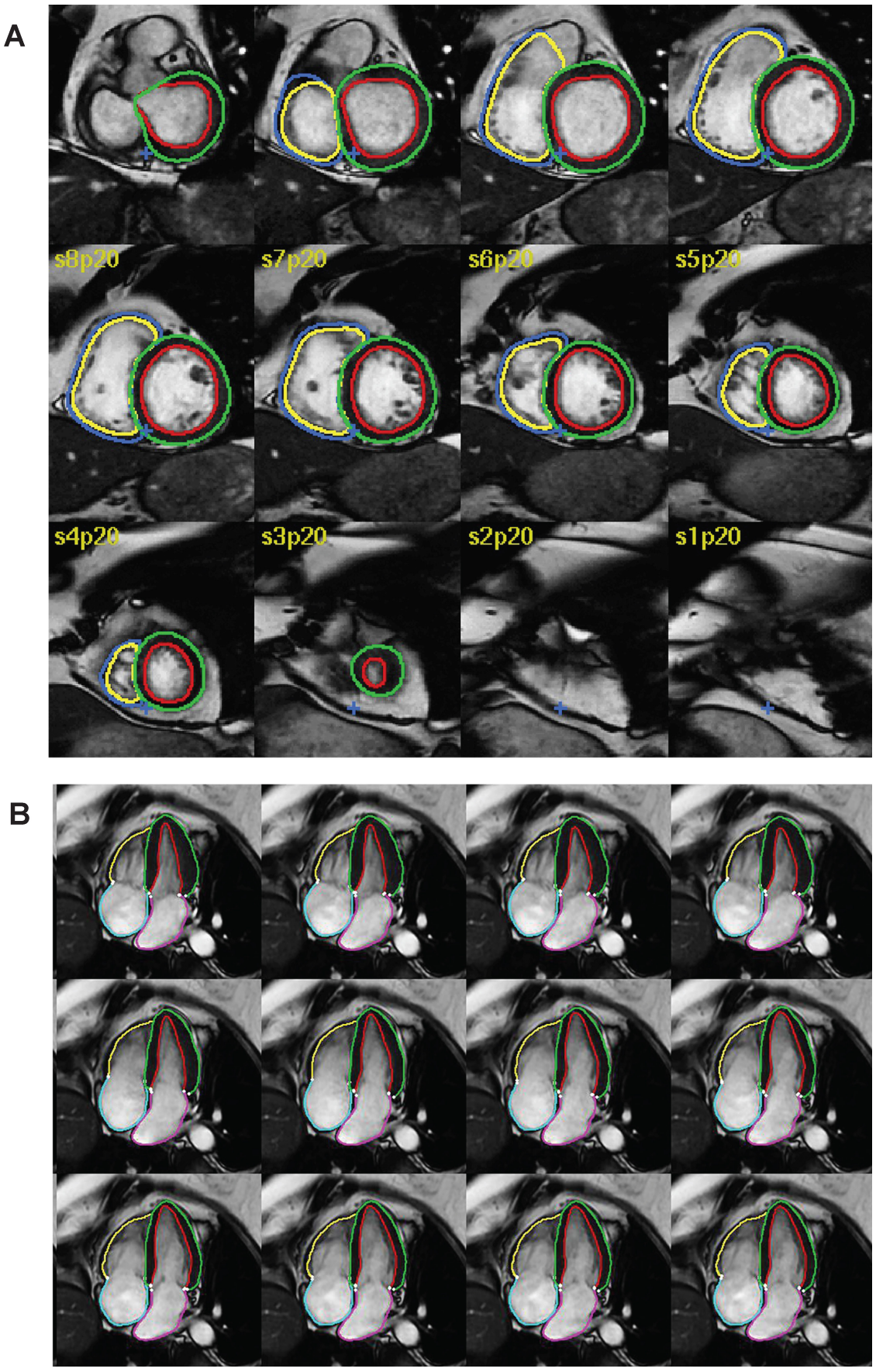


Figure 3.1: Examples of AI cardiac MRI segmentation. **A**: biventricular and **B**: four-chamber segmentation. The coloured contours in green and red show the left ventricular epi- and endocardium, respectively. The contours in dark blue and yellow show the right ventricular epi- and endo- cardium, respectively. The pink and turquoise contours outline the left and right atria, respectively.

3.2.2 Search method

The MEDLINE and EMBASE databases were searched for relevant studies on April 20 2022.

The following search strategy was used:

1. automat*
2. exp ARTIFICIAL INTELLIGENCE/
3. exp MACHINE LEARNING/
4. exp DEEP LEARNING/
5. exp DEEP NEURAL NETWORK/
6. exp CONVOLUTION ALGORITHM/
7. ((deep or supervised or unsupervised or machine) and learning)
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. exp MAGNETIC RESONANCE IMAGING/
10. exp CARDIOVASCULAR MAGNETIC RESONANCE/
11. (Cine or MRI or MRA or (Magnetic and resonance))
12. 9 or 10 or 11
13. (heart or cardi* or myocard* or coronar* or ventric* or LV or RV or atri*)
14. (segment* or conto* or annotat* or label*)
15. exp IMAGE SEGMENTATION/
16. exp SEGMENTATION ALGORITHM/
17. 14 or 15 or 16
18. 8 and 12 and 13 and 17

3.2.3 Study selection

Figure 2 indicates the flow of study identification and inclusion. Duplicate studies were removed following the initial database search. The titles and abstracts of the remaining studies were screened for relevance. The full texts of all potentially relevant studies were retrieved and assessed for eligibility against the inclusion and exclusion criteria. Conference abstracts and studies lacking sufficient information for evaluation were excluded at this point. Screening was performed independently by (S.A.) and by (S.D., A.M.2., M.S.2.)

and full texts were assessed for eligibility by (S.A., A.M.1 and M.S.), with (S.A.) acting as an arbitrator.

3.2.4 Data extraction

Three authors extracted data from the included studies (S.A., A.M1., M.S.1.) according to a standardised checklist. Half of the included studies were also evaluated independently by an additional five authors (S.D., A.M.2., S.J., M.G., H.A.) for the purpose of quality control. All discrepancies were resolved with discussion, with S.A. acting as an arbitrator, and the final extracted data confirmed. Descriptive information about each study was recorded, including publication details (type, source, country, year), data used (type of data set, type of CMR image, segmented structures) and AI model (validation and performance evaluation methods). The studies were assessed for compliance against the 42 criteria of CLAIM, which were grouped into four domains: study description (9 criteria), dataset description (17 criteria), model description (6 criteria) and model performance (10 criteria). For each criterion, compliance was marked as yes, no or not applicable (N/A). Studies deemed N/A were excluded when evaluating the proportion of studies compliant with CLAIM criteria. For studies using solely public datasets, the following criteria were marked as N/A, as they can be considered implicit in the use of publicly available data sources: retrospective or prospective study, source of ground truth annotations, annotation tools, de-identification methods and inter- and intra-rater variability. Additionally, the following criteria were marked as N/A for all studies: rationale for choosing the reference standard (as manual expert contouring is the standard in the field) and registration number and name of registry. Descriptive data and the number of studies compliant with CLAIM criteria are presented as proportional values (%).

3.3 Results

3.3.1 Search results

The database search yielded 2855 hits from which the title and abstract screening identified 364 relevant studies of which 155 were excluded because they were conference reports, non-CMR, animal studies, extra-cardiac, semi-automated or non-AI segmentation. The subsequent full-text assessment deemed 209 eligible for inclusion in the analysis (Figure 3.2).

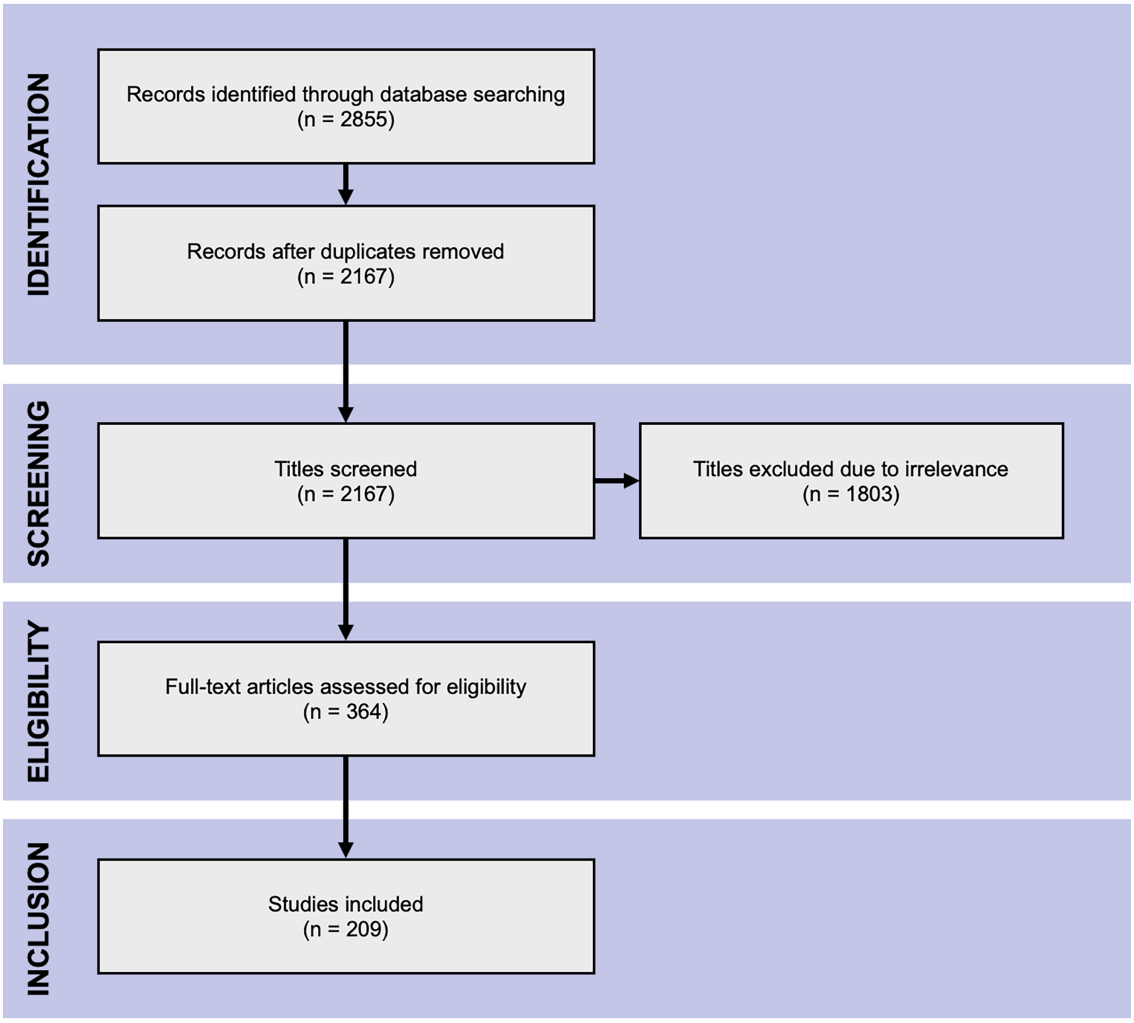


Figure 3.2: The Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart of literature search.

3.3.2 Included studies

Descriptive information for all of the 209 included studies are provided in Supplementary Table 1. Selected metrics are highlighted in Figure 3.3. The majority of studies (57%) were published since 2019 (Figure 3.3A). Most studies were published in technical journals (58%), with a minority published in medical (31%) or hybrid (11%) journals. The studies were undertaken in 37 different countries (Figure 3.3B), with just over half coming from China (26%), the USA (18%) and the UK (11%).

Publicly available datasets were used in 49% of studies, and single or multicentre non-public datasets used in 61%, 17% of studies used multiple or combined datasets (including multiple public datasets and a combination of public and non-public datasets). A minority of studies (6%) did not report their data source (Figure 3.3C). Of the public datasets used, the majority (86%) had been made available through Medical Image Computing and Computer-Assisted Intervention (MICCAI) challenges or the Cardiac Atlas Project [164] (Figure 3.3D). Most studies reported the number of cases used (95%), with a range of 3 to 12984 and a median of 78. Short axis CMR images were most frequently used (70%), while 14% of studies did not report the specific type of CMR image used for segmentation (Figure 3.3E). The left ventricle was the most commonly segmented structure, either alone or in combination (49%, Figure 3.3F). Segmentation of multiple structures was reported in 23% of studies.

Model validation was mostly reported using internal holdout methods (78%), such as cross-validation. A minority reported testing on external and mainly public datasets (22%, Figure 3.3G). The Dice similarity coefficient (DSC) was used to assess model performance in 79% of studies, either alone or in combination with other metrics such as the Hausdorff distance or the Jaccard index (Figure 3.3H). Few studies (10%) provided working links to publicly available code, with a further 1% indicating that code was available on request.

3.3.3 Compliance with CLAIM

Results for compliance with the domains and selected individual criteria of CLAIM are summarised in Figure 3.4. The complete results are presented in Table 3.1. The median

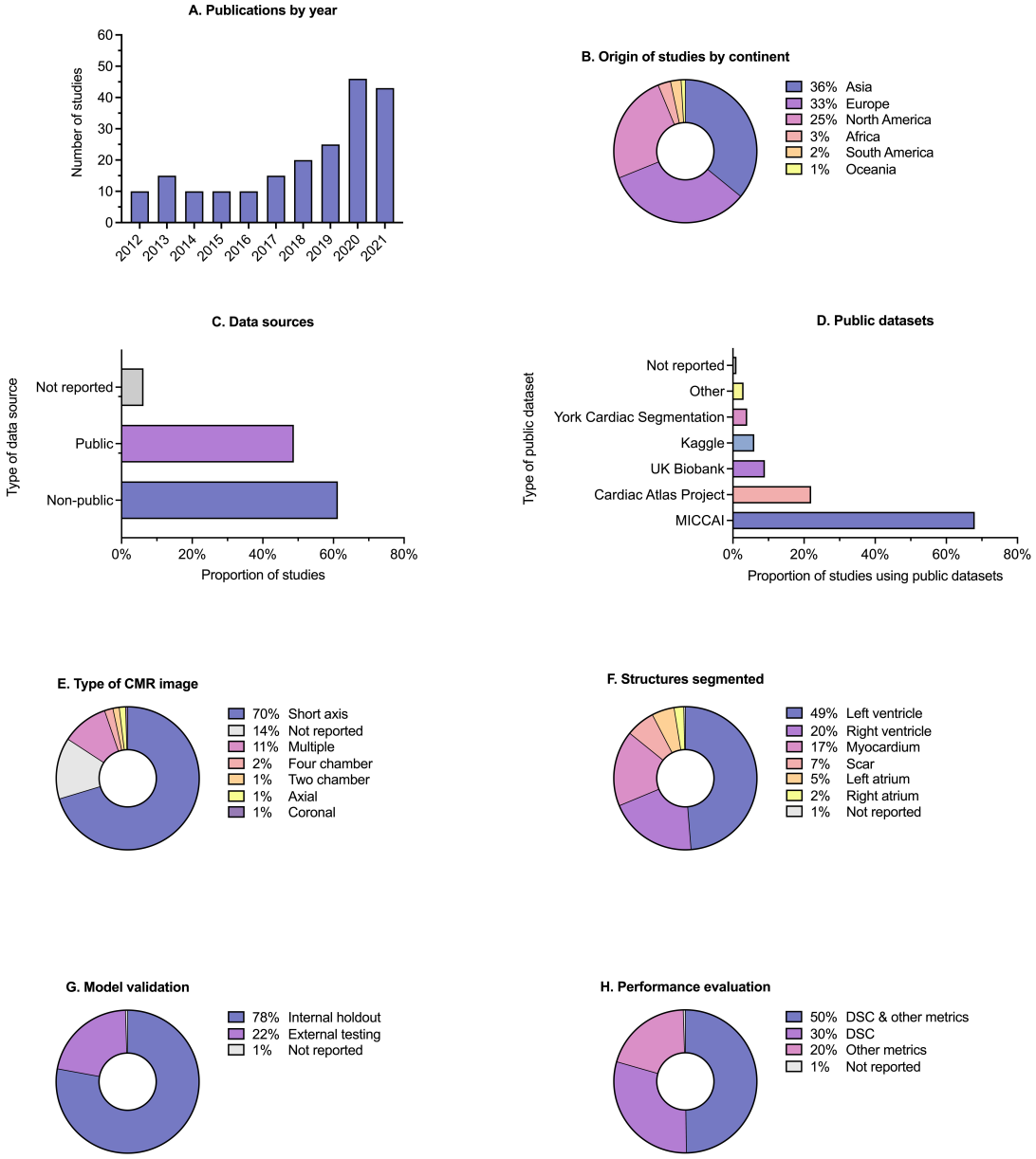


Figure 3.3: Descriptive information for the 209 included studies. **A**: publication dates; five studies (2.4%) were included from early 2022 and are not indicated here. **B**: location of origin of studies. **C**: data sources; the proportion of studies which used public and non-public datasets is shown, with some studies having used multiple or combined datasets. **D**: public datasets used by studies, where relevant. **E**: type of CMR images used. **F**: cardiac structures segmented; some studies performed segmentation on multiple structures. **G**: method of model validation. **H**: method of model performance evaluation.

compliance of all studies with all 42 criteria of CLAIM was 67% (IQR $59 - 73\%$). Notable results excluding non-applicable criteria are as follows.

Table 3.1: Compliance with CLAIM checklist.

	No.	Criteria	Domain	Yes	No
Title & abstract					
Title	1	Identification as a study of AI methodology, specifying the category of technology used (e.g., deep learning)	Study description	90.9%	9.1%
Abstract	2	Structured summary of study design, methods, results, and conclusions	Study description	52.6%	47.4%
Introduction					
Introduction	3	Scientific and clinical background, including the intended use and clinical role of the AI approach	Study description	92.3%	7.7%
	4	Study objectives and hypotheses	Study description	94.3%	5.7%
Methods					
Study design	5	Prospective or retrospective study	Study description	36.0%	64.0%
	6	Study goal, such as model creation, exploratory study, feasibility study, non-inferiority trial	Study description	95.2%	4.8%
Data sources	7	Data sources	Dataset description	93.8%	6.2%
	8	Eligibility criteria: how, where, and when potentially eligible participants or studies were identified (e.g., symptoms, results from previous tests, inclusion in registry, patient-care setting, location, dates)	Dataset description	74.2%	25.8%
	9	Data pre-processing steps	Dataset description	93.8%	5.7%
	10	Selection of data subsets, if applicable	Dataset description	92.8%	6.7%
	11	Definitions of data elements, with references to Common Data Elements	Dataset description	99.5%	0.5%
	12	De-identification methods	Dataset description	11.2%	88.8%
	13	How missing data were handled	Dataset description	8.6%	91.4%
Ground truth reference standard	14	Definition of ground truth reference standard, in sufficient detail to allow replication	Dataset description	67.6%	32.4%
	15	Rationale for choosing the reference standard (if alternatives exist)	Dataset description	N/A	N/A
	16	Source of ground-truth annotations; qualifications and preparation of annotators	Dataset description	54.8%	45.2%
	17	Annotation tools	Dataset description	30.6%	69.4%
	18	Measurement of inter- and intrarater variability; methods to mitigate variability and/or resolve discrepancies	Dataset description	41.9%	58.1%
Data Partitions	19	Intended sample size and how it was determined	Dataset description	4.3%	95.7%

Compliance with CLAIM checklist. (cont.)

	20	How data were assigned to partitions; specify proportions	Dataset description	89.4%	10.6%
	21	Level at which partitions are disjoint (e.g., image, study, patient, institution)	Dataset description	87.0%	13.0%
Model	22	Detailed description of model, including inputs, outputs, all intermediate layers and connections	Model description	94.7%	5.3%
	23	Software libraries, frameworks, and packages	Model description	74.2%	25.4%
	24	Initialization of model parameters (e.g., randomization, transfer learning)	Model description	91.7%	8.3%
Training	25	Details of training approach, including data augmentation, hyperparameters, number of models trained	Model description	78.3%	21.7%
	26	Method of selecting the final model	Model description	91.6%	8.4%
	27	Ensembling techniques, if applicable	Model description	50.0%	50.0%
Evaluation	28	Metrics of model performance	Model performance	99.5%	0.5%
	29	Statistical measures of significance and uncertainty (e.g., confidence intervals)	Model performance	77.5%	22.5%
	30	Robustness or sensitivity analysis	Model performance	60.8%	39.2%
	31	Methods for explainability or interpretability (e.g., saliency maps), and how they were validated	Model performance	64.1%	35.9%
	32	Validation or testing on external data	Model performance	21.5%	78.5%
Results					
Data	33	Flow of participants or cases, using a diagram to indicate inclusion and exclusion	Dataset description	10.0%	90.0%
	34	Demographic and clinical characteristics of cases in each partition	Dataset description	18.2%	81.8%
Model performance	35	Performance metrics for optimal model(s) on all data partitions	Model performance	88.9%	11.1%
	36	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	Model performance	20.7%	79.3%
	37	Failure analysis of incorrectly classified cases	Model performance	32.1%	67.9%
Discussion					
Discussion	38	Study limitations, including potential bias, statistical uncertainty, and generalizability	Model performance	76.1%	23.9%
	39	Implications for practice, including the intended use and/or clinical role	Model performance	75.6%	24.4%
Other information					
Other information	40	Registration number and name of registry	Study description	N/A	N/A
	41	Where the full study protocol can be accessed	Study description	0.0%	100.0%
	42	Sources of funding and other support; role of funders	Study description	82.3%	17.7%

N/A = not applicable. Checklist adapted from Mongan et al. 2020 [161].

- Study description

  Median compliance with the study description domain was 71% (IQR 63 – 86%). Almost all studies clearly indicated the use of AI methods (91%) and their objectives (94%). Where non-public datasets were used, only a minority of studies (36%) indicated whether these were prospective or retrospective. No studies provided access to a full study protocol. Sources of funding were declared in 82% of studies.

- Dataset description

  Median compliance with the dataset description domain was 63% (IQR 50 – 67%), the lowest of the four domains. The source of the dataset was reported in most studies (94%). While most studies provided eligibility criteria for included cases (74%), few studies reported their demographic and clinical characteristics (18%) or indicated the flow of these cases (10%) in sufficient detail. Details regarding the calculation of the intended sample size (4%) and how missing data were handled (9%) were also infrequently reported. The definition of the ground truth reference standard was provided in 68% of studies. Where non-public datasets were used, the source of ground truth annotations and annotation tools were stated in 55% and 31% of studies respectively, with inter- and intra-rater variability reported in 42%. The majority of studies reported data preprocessing steps (94%), definitions of data elements (99.5%), how data were assigned to partitions (89%) and the level at which partitions were disjoint (87%).

- Model description

  Median compliance with the model description domain was 100% (IQR 80 – 100%), the highest of the four domains. The majority of studies provided details about the model used (95%), initialisation of model parameters (92%), training approach (78%) and method of selecting the final model (92%). The software libraries, frameworks and packages used were reported in 74%.

- Model performance

  Median compliance with the performance description domain was 60% (IQR 50−70%). A minority of studies reported testing on external data (22%) Almost all studies provided metrics of model performance (99.5%). Most studies provided statistical measures of significance and uncertainty when reporting results (78%). Many

studies provided forms of robustness or sensitivity analysis (61%) and methods for explainability and interpretability (64%). A minority of studies reported failure analysis for incorrectly classified cases (32%). Most studies discussed their limitations (76%) and implications for practice (76%).

3.4 Discussion

Poor reporting is a major source of research waste [165, 166] and ultimately may hinder advancement of AI research in the medical field. This systematic review evaluated the quality of reporting in AI studies involving automatic segmentation of structures on cardiac MRI. 209 studies were included from 2012 to early 2022. Each study was assessed for compliance with CLAIM, a checklist that attempts to provide a 'best practice' framework for the reporting and publication of AI research in medical imaging [161]. We identified major gaps in reporting and make a number of recommendations in order for this to be addressed (Table 3.1).

Table 3.1: Main recommendations for AI study reporting are based on the gaps in the literature identified in this systematic review.

	Recommendation	Importance
General	Utilise a reporting framework (e.g., CLAIM).	Comparability of studies.
	Use of consistent and descriptive terminology.	Accessibility and comparability of studies.
Data sources	Describe the source of data, including patients' eligibility criteria, their numbers and demographic and clinical characteristics.	Contextualising model performance and generalisability.
	Clarify the number of scans and the flow of both patients and scans into different datasets (e.g. training, validation, and testing).	Understanding model performance and generalisability.
	Use publicly available datasets.	Comparability of models against a common benchmark.
Model training and evaluation	Describe the neural network, software packages and libraries in sufficient detail.	Study reproducibility.
	Define how the reference contours were generated, the experience of the annotator and annotation tools used.	Understanding model performance and generalisability.
	Explain the method of model training and performance.	Understanding model performance and generalisability.
	Test the model performance on external data with different characteristics to the training data.	Study and model reliability. Understanding model generalisability. Implementation in clinical practice.
	Perform failure analysis and report the limitations of the model.	Understanding model performance and generalisability.
	Publication of open-source code.	Understanding model performance and generalisability.

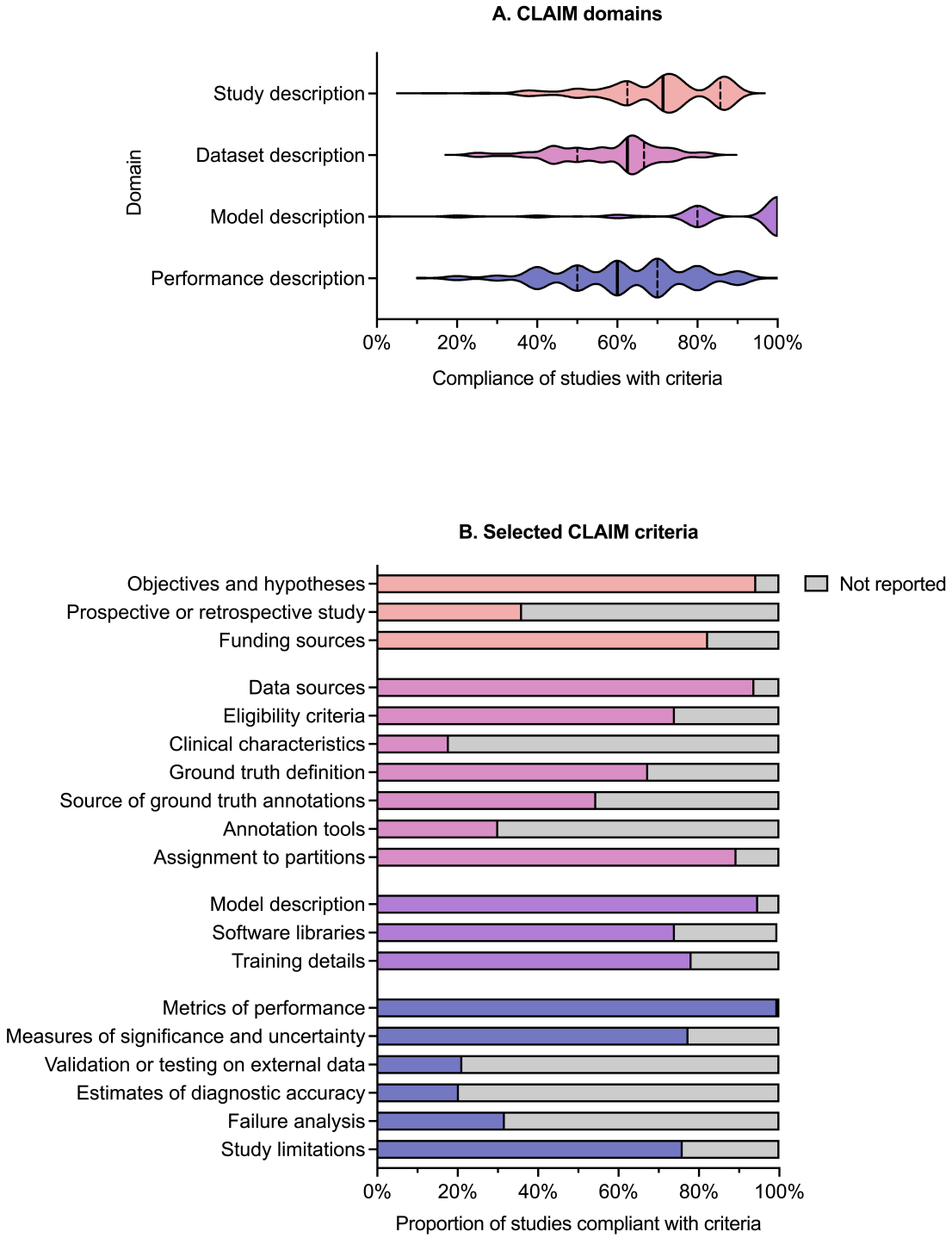


Figure 3.4: **A**: violin plot showing compliance of the 209 included studies with the CLAIM criteria, grouped into domains of study, dataset, model and performance description. Median (solid line) and 1st and 3rd quartile (dashed lines) values are indicated. **B**: proportion of studies compliant with selected CLAIM criteria, grouped by domain (the titles of the individual criteria have been shortened for ease of reading).

Accurate and sufficiently detailed descriptions of study materials and methods are of particular importance for AI studies in medical imaging to allow the assessment of reproducibility and reliability of results. Overall compliance with CLAIM was highest for the model description domain, with most studies providing a description of the model and details of training approaches. However, this was lowest for the dataset description domain, which indicated variable reporting of the data sources used to train and evaluate models.

A good understanding of data sources is a prerequisite for evaluating the validity of AI models. Although most studies identified their data sources, this was a significant omission in the studies that did not and one which greatly limits their interpretability. Public datasets were used in almost half of the studies, with the majority of these made available through segmentation challenges hosted by MICCAI (Table 3.2).

Table 3.2: Public datasets.

Dataset	Year	No. of patients	Cardiac chamber	Source
York [167]	2008	33	LV	Hospital for Sick Children, Canada
LVSC [168]	2009	45	LV	Sunnybrook Health Sciences Centre, Canada.
LVSC II [169]	2011	200	LV	Multicentre
LVIC [170]	2012	30	LV	Multicentre
RVSC [171]	2012	48	RV	Rouen University Hospital, France
cDEMRIS [172]	2012	60	LA	Multicentre
LASC [173]	2013	30	LA	King's College London, UK
SLAWT [174]	2016	10	LA	Single centre not specified
HVSMR-I [175]	2016	20	Whole heart	Boston Children's Hospital, USA
MM-WHS [176]	2017	60	Whole heart	Single centre not specified
ACDC [177]	2017	150	LV, RV	University Hospital of Dijon, France
LASC II [178]	2018	154	LA	The University of Utah, USA
LVQuan'18 [179]	2018	175	LV	London HealthCare, Ontario Canada
LVQuan'19 [180]	2019	85	LV	Not specified.
MS-CMRS [181]	2019	45	LV, RV	Not specified
OCMR [182]	2020	286	Whole heart	Multicentre
EMIDEC [183]	2020	150	LV	University Hospital of Dijon, France
M&MS [184]	2020	350	Whole heart	Multicentre
HVSMR-II [185]	2021	90	Whole heart	Boston Children's Hospital, USA
M&Ms-2 [184]	2021	360	Whole heart	Multicentre, Spain
LAScarQS [186]	2022	194	LA (LGE)	Multicentre

ACDC; Automatic Cardiac Diagnosis Challenge, cDEMRIS; Cardiac Delayed Enhancement Segmentation Challenge, EMIDEC; automatic Evaluation of Myocardial Infarction from Delayed-Enhancement Cardiac MRI, HVSMR; Whole-Heart and Great Vessel Segmentation; LA; Left Atrium, LAScarQS, Left Atrial and Scar Quantification & Segmentation Challenge, LASC; Left Atrial Segmentation Challenge, LGE; Late Gadolinium Enhancement, LV; Left Ventricle, LVIC; Left Ventricle Infarct Challenge, LVQuan; Left Ventricle Full Quantification, LVSC; Left Ventricle Segmentation Challenge, MM-WHS; Multi-Modality Whole Heart Segmentation, M&MS; Multi-Disease, Multi-View & Multi-Center Right Ventricular Segmentation in Cardiac MRI, MS-CMRS; Multi-Sequence Cardiac MR Segmentation, RV; Right Ventricle, RVSC; Right Ventricle Segmentation Challenge, SLAWT; Segmentation of Left Atrial Wall for Thickness

Public datasets contain previously de-identified and expertly contoured images, making them attractive to researchers. The proportion of studies using datasets from MICCAI challenges underlines its role as a driver for advancing the field. Importantly, the use of public datasets facilitates reproducibility and aids comparison between segmentation methods. However, public sources are not without their limitations. Public datasets consist of entirely retrospective data, which may place constraints on study design and model training. They are often small in size with limited demographic and clinical diversity, and therefore have inherent selection bias. Systematic biases affecting patient demographics are of serious concern in the application of AI methods to clinical practice. For example, a previous analysis of AI-based segmentation in CMR using a large-scale database found systematic bias for both sex and race [187] and similar biases have been reported for AI in radiographic imaging [188]. The use of diverse datasets when training, validating and testing models is essential for generalisability and translation to clinical practice. A model trained on a dataset from one population does not guarantee equal performance on another. Multiple data sets, such as both retrospective and prospective, could be used in combination to improve the generalisability of AI models being trained. Even accounting for the use of public datasets, we found that few studies reported the intended sample size (which influences statistical power and reliability of results) or the demographic and clinical characteristics of the cases in each partition, (which indicates selection bias, confounders and generalisability). Providing summary information about the age and sex of cases is important, but may be insufficient in isolation. We noted that studies often lacked details about the proportions of cases with different pathologies, and the demographics for these groups. Furthermore, studies should not assume that readers are familiar with public datasets, and if these are used then detailed demographics and clinical characteristics should still be reported. The performance and validity of any model depend on the data on which it is trained and the data sources, including the rationale behind their choice and the intended sample size, should be clearly indicated. Study methodology must be reported in sufficient detail to enable accurate reproduction of results. Notably, the definition of the ground truth reference standard, the source of ground truth annotations and the annotation tools used were absent in a substantial number of studies. Understanding the structures included in the ground truth contours and the expertise of the annotator is

essential in evaluating the training process and ultimately contextualising the model's performance. The proportions of studies that provided sufficiently detailed descriptions of the ground truth and its source were lower than expected for the field. For example, judging from the figures present in the included studies, ventricular trabeculations were usually included in the blood pool contours, although few studies described this process. Similarly, many studies failed to report the specific type of image used for ground truth annotation and model training and testing. While this could be inferred from figures, it remains essential information for understanding models and their generalisability. Finally, only a handful of studies indicated how missing data were handled and no studies indicated where a full study protocol could be accessed.

Detailed description of model training and performance is expected in this field. Testing model performance on external data was performed in less than a quarter of all studies. Model generalisability can only be fully evaluated when performance is assessed in demographic and clinical populations different from the original training cohort. The reported external datasets were small and captured only limited variations in imaging appearances. This represents a major hurdle to overcome before AI models can be implemented in clinical practice. We also noted subjectively that many publications used the terms 'validation' and 'test' interchangeably, or failed to distinguish these methods clearly. Regarding the use of data in AI studies, a validation set is used to optimise hyperparameters and performance between training epochs, while a testing set is used to assess the performance of the final model. The lack of consistent terminology in studies can limit the interpretability of their models and blur the distinction between internal holdout and external testing methods. Additionally, few studies reported failure analysis of incorrectly classified cases, suggesting that most did not explore the reasons for model underperformance. Furthermore, the vast majority of studies did not discuss the limitations of their methods, limiting their transparency. Open publishing of source code is a contentious topic in AI research and was only provided in one in ten of all studies. The public availability of code aids transparency, assists peer review and facilitates the development of new models, but bears important implications for ownership and rights.

The use of reporting frameworks, such as CLAIM, can be beneficial. For example, they may help to inform study design and highlight areas that may require rectification prior to dissemination of results. Frameworks assist standardisation in reporting, improving comparability and interpretability by the wider scientific community. Study accessibility is also an important consideration in advancing the field. Regardless of journal type, AI studies in medical imaging need to cater for a broad potential readership, from clinicians to computer scientists. More standardised reporting and the use of consistent and accessible terminology are important in this regard.

3.4.1 Limitations

We acknowledge limitations in this systematic review. Firstly, this review focused solely on AI segmentation in CMR studies. However, these findings are likely to apply to AI studies in other cardiac imaging modalities, such as echocardiogram, CT coronary angiography or nuclear myocardial perfusion studies. Furthermore, given that AI studies in chest imaging have shown similar shortcomings in reporting quality [189], our findings may be more broadly relevant to AI studies in medical imaging. Secondly, while our systematic search aimed to identify all published AI CMR segmentation studies, the body of unpublished, pre-print or technical conference literature is vast. A Github or `arxiv.org` search reveals numerous segmentation attempts of varying levels of reporting quality and beyond the scope of this review to capture. Thirdly, even despite the use of structured tools such as CLAIM, there remains an element of subjectivity in determining report quality, such as the amount of information required for a study to be deemed reproducible.

3.5 Conclusion

This systematic review highlights the variability in reporting and identifies gaps in the existing literature of studies using AI segmentation of CMR images. We identified several key items that are missing in publications - most strikingly poor description of patients included in the training and validation of AI models and inadequate model failure analysis - which may limit study transparency, reproducibility and validity. This review supports closer adherence to established frameworks for reporting standards, such as CLAIM. In light of these findings, we have presented a number of recommendations for improving the

quality of reporting of AI studies in both CMR and the wider field of cardiac imaging.

3.6 Author Contributions

S.A. and A.J.S. conceived the idea and need for the systematic review and contributed to the study conception and design. Protocol registration in PROSPERO was performed by (S.A. and S.D.). S.A. created the search strategy and performed the literature search. S.A. and (S.D., M.S.1., M.S.2., A.M.2.) performed screening and eligibility assessments independently. S.A., S.D., A.M.1., A.M.2., M.S.1., M.S.2., S.J., M.G, V.R., H.A. evaluated the included studies and collected relevant data. Material preparation and analysis were performed by S.A., A.M.1., M.S., A.M.2. and S.J.. The manuscript, figures and tables were drafted by S.A., A.M.1., S.J. and M.S.1.. All authors contributed to the interpretation of data. The first draft of the manuscript was written by S.A., A.M.1., A.M.2., V.R. and S.D. and all authors commented on previous versions of the manuscript. The final draft was written by S.A. and A.M.1., taking into account comments and suggestions from experts in the field A.J.S., D.P.O.R., H.L., R.V.D.G., M.M. and P.G.. All authors read and approved the final manuscript. All authors took part in the critical review and drafting of the manuscript and have read and approved the final manuscript.

CHAPTER 4

Development and Validation of Artificial Intelligence Cardiac MRI Measurements: Relationship to Heart Catheterization and Mortality Prediction

Samer Alabed[1,2], Faisal Alandejani[1], Krit Dwivedi[1,2], Kavita Karunasaagarar[1,2], Michael Sharkey[1,2], Michail Mamalakis[3], Patrick de Koning[4], Pankaj Garg[4], Attila Tóth[6], Yousef Shahin[1,2], Christopher Johns[1,2], David Capener[1], Steven Wood[3], Peter Metherall[3], Alexander Rothman[1], Robin Condliffe[1,7], Neil Hamilton[7], James M. Wild[1], Declan P. O'Regan[8], Haiping Lu[9] , David G. Kiely[1,7], Rob van der Geest[4], Andrew J. Swift[1,2]

[1] Infection, Immunity & Cardiovascular Disease, University of Sheffield

[2] Department of Clinical Radiology, Sheffield Teaching Hospitals

[3] Department of Computer Science, University of Sheffield

[4] Leiden University Medical Center, Leiden, the Netherlands

[5] Norwich Medical School, University of East Anglia

[6] Semmelweis University Heart and Vascular Center, Budapest, Hungary

[7] Sheffield Pulmonary Vascular Disease Unit, Sheffield Teaching Hospitals

[8] MRC London Institute of Medical Sciences, Imperial College London

This chapter is based on a publication in *Radiology.* 2022 Oct;305(1):68-79.

Abstract

Objectives To develop and evaluate a deep learning tool for quantitative evaluation of cardiac MRI functional studies and assess its use for prognosis in patients suspected of having pulmonary hypertension.

Background Cardiac MRI measurements have diagnostic and prognostic value in the evaluation of cardiopulmonary disease. Artificial intelligence approaches to automate cardiac MRI segmentation are emerging but require clinical testing.

Methods A retrospective multicentre and multivendor data set was used to develop a deep learning–based cardiac MRI contouring model using a cohort of patients suspected of having cardiopulmonary disease from multiple pathologic causes. Correlation with same-day right heart catheterisation (RHC) and scan-rescan repeatability was assessed in prospectively recruited participants. Prognostic impact was assessed using Cox proportional hazard regression analysis of 3487 patients from the ASPIRE (Assessing the Severity of Pulmonary Hypertension In a Pulmonary Hypertension Referral Centre) registry, including a subset of 920 patients with pulmonary arterial hypertension. The generalisability of the automatic assessment was evaluated in 40 multivendor studies from 32 centres.

Results The training data set included 539 patients (mean age, 54 years ± 20 [SD]; 315 women). Automatic cardiac MRI measurements were better correlated with RHC parameters than were manual measurements, including left ventricular stroke volume (r = 0.72 vs 0.68; $P = .03$). Interstudy repeatability of cardiac MRI measurements was high for all automatic measurements (intraclass correlation coefficient range, $0.79 - 0.99$) and similarly repeatable to manual measurements (all paired t test $P > .05$). Automated right ventricle and left ventricle cardiac MRI measurements were associated with mortality in patients suspected of having pulmonary hypertension.

Conclusion An automatic cardiac MRI measurement approach was developed and tested in a large cohort of patients, including a broad spectrum of right

ventricular and left ventricular conditions, with internal and external testing. Fully automatic cardiac MRI assessment correlated strongly with invasive haemodynamics, had prognostic value, were highly repeatable, and showed excellent generalisability. (Clinical trial registration no. NCT03841344)

Summary

An Artificial intelligence cardiac MRI segmentation model was developed to automate cardiac MRI measurements and subsequently tested against invasive right heart catheterisation parameters and prediction of patient mortality.

Key Results

- A retrospective training data set of 539 patients with left and right heart disease was used to train an artificial intelligence (AI) model for cardiac MRI measurements.
- Same-day cardiac MRI and right heart catheterisation demonstrated strong correlation that was higher with AI measurements than with manual measurements for left ventricular stroke volume ($r = 0.74$ vs. 0.68; $P = .03$; $n = 178$).
- AI-measured right ventricular end-systolic volume, ejection fraction, and mass all predicted mortality in patients with pulmonary arterial hypertension (hazard ratios, $1.40, 0.76$, and 1.15, respectively; $P = .001$; $n = 920$).

4.1 Introduction

Cardiac MRI is the reference standard for measuring cardiac chambers and has an important role in the diagnosis and prognosis of cardiovascular disease. Manual measurements are obtained by tracing the cardiac chambers in end-diastole and end-systole, a time–consuming process that requires a specialized workforce. Efforts to automate cardiac MRI measurements have evolved over recent years [159] and have achieved comparable results to manual assessments in assessing the left ventricle (LV) [190]. However, greater internal and external testing and clinical benchmarks for automatic cardiac MRI quantification are required.

Automatic assessment in the right ventricle (RV) is a challenge because of the variation in the shape, thickness, and complex anatomy, particularly at the base and outflow tract [159, 171]. Additionally, the RV shape can undergo extreme morphologic changes in conditions such as pulmonary hypertension [30]. Automating RV assessments has the potential to improve reproducibility of RV analysis. To date, artificial intelligence (AI) biventricular segmentation studies are based on small single–centre and single–vendor data sets, include limited numbers of patients with conditions affecting the RV, and do not assess RV mass [159].

The aim of our study was to develop and comprehensively evaluate an automated deep learning quantitative analysis of LV and RV cardiac MRI measurements. We also sought to assess the hypothesis that an AI cardiac MRI biventricular analysis correlates with invasive haemodynamics, predicts mortality in pulmonary hypertension, and is repeatable and generalisable.

4.2 Materials and Methods

4.2.1 Study Sample

Our study involved a retrospective training data set and a testing data set (Figure 4.1). The training data set included 611 studies performed at two university teaching hospitals (Sheffield Teaching Hospitals, Sheffield; and Heart and Vascular centre of Semmelweis University, Budapest) in 539 participants with various cardiac abnormalities. The cardiac

abnormalities included LV diseases such as ischaemic and non-ischaemic cardiomyopathies, valvular heart disease, systolic and diastolic dysfunction and RV disease such as pulmonary hypertension. The Budapest data set included 192 patients with 211 MRI studies randomly chosen in patients referred for investigation of suspected or confirmed LV disease. Nineteen patients had repeat imaging. Overall, the data set included 180 MRI studies with LV disease and 32 normal MRI studies. The Budapest studies were performed with a Philips MRI system and was used to train the initial LV and RV segmentation model. For the Sheffield data set, consecutive patients suspected of having pulmonary hypertension who underwent cardiac MRI were identified from the Assessing the Severity of Pulmonary Hypertension In a Pulmonary Hypertension Referral Centre (known as ASPIRE) registry between 2007 and 2021 [191]. The ASPIRE registry includes all patients with suspected pulmonary hypertension (PH) referred to The Sheffield Pulmonary Vascular Disease Unit (SPVDU) at the Royal Hallamshire Hospital (Sheffield, UK). The ASPIRE cardiac MRI subcohort is a well described patient population that has been analysed in multiple studies; list of studies shown at the end of this chapter.

Patients with incomplete, unavailable, or unretrievable short–axis stack were excluded. An off–line human–in–the–loop approach (HITL) was used, wherein the initial segmentation model trained with the MRI scans acquired at Budapest was tested in the Sheffield data set and a random sample of cases that had suboptimal or failed segmentations were included for further training (Figure 4.2, Figure 4.3). The first round of training included 220 MRI studies and the second round included 180 studies that were again identified from suboptimal segmentations resulting from the refined segmentation model. The first, second, and final training rounds were performed with Philips, GE, and Siemens MRI systems, respectively.

For the clinical testing, four data sets were included: a prospective same–day repeatability cohort (n = 46); a prospective same–day right heart catheterisation (RHC) cohort (n = 179); an external test cohort (n = 40) from 32 centres across England, Wales, and Scotland; and MRI studies not included in the training from the ASPIRE registry used for the assessment of mortality prediction (n = 3782).

Participants were recruited prospectively for our reproducibility analysis as part of the Repeatability and Sensitivity to Change of Noninvasive End Points in Pulmonary Arterial Hypertension, or RESPIRE, study [192] (ClinicalTrials.gov Identifier: NCT03841344). Ethical approval for the study was granted by the local ethics committee and institutional review board (ASPIRE, reference $c06/Q2308/8$; REC $17/YH/0016$; and RESPIRE, REC 15/YH/0269). All prospectively recruited participants gave written informed consent. All data were strictly anonymized before analysis. We followed the Checklist for Artificial Intelligence in Medical Imaging (known as CLAIM) for reporting AI studies [161].

4.2.2 Imaging Procedures

MRI Protocol

Cardiac MRI was performed with 1.5−T MRI systems from three vendors (Signa HDx, GE Healthcare; Avanto, Siemens Solutions; and Achieva, Philips Healthcare). Multisection short–axis cine images were obtained by using a standard cardiac–gated balanced steady–state free precession sequence of 8−mm section thickness and 20 phases per cardiac cycle (Signa HDx; GE Healthcare), 6−mm section thickness and 25 phases per cardiac cycle (Avanto; Siemens Solutions), and 8−mm section thickness and 25 phases per cycle (Achieva; Philips Healthcare). The parameters (repetition time msec/echo time msec) were 3.7/1.6 (Signa HDx; GE Healthcare), 38.92/1.13 (Avanto; Siemens Solutions), and 2.72/1.36 (Achieva; Philips Healthcare). Two–dimensional phase–contrast sequences were acquired perpendicular to the long axis of the aortic lumen by using through–plane velocity encoding. All phase–contrast sequences were performed with GE MRI systems with the following imaging parameters: 5.6/2.7; section thickness, 10mm; 20 phases; and velocity encoding, 150 cm per second in the section direction. For the scan-rescan acquisitions, patients had their first scan in the morning and the repeat scan in the afternoon.

Image Analysis

Manual segmentations of biventricular epicardial and endocardial contours on short–axis stack images for the training and testing data sets were performed by seven observers (A.T., D.C., K.K., A.J.S., S.S., F.A.A., and S.A., with $19, 17, 13, 11, 4, 3$, and 3 years of specialist cardiac MRI experience, respectively). All manual contours were reviewed by one author (A.J.S., a level 3 accredited cardiac MRI radiologist). Trabeculations were included

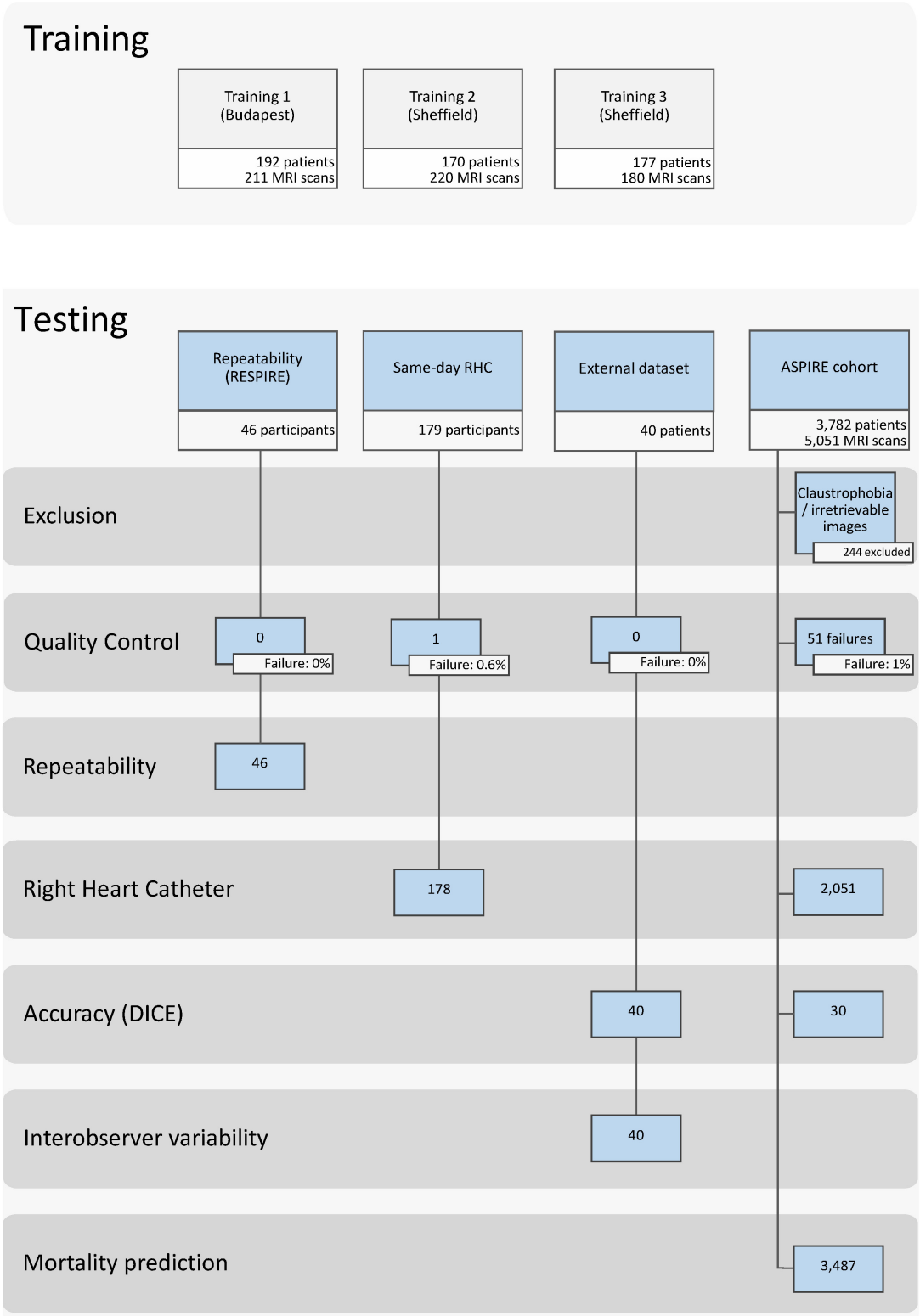


Figure 4.1: Study participant flow chart for the training and testing cohorts. ASPIRE = Assessing the Severity of Pulmonary Hypertension In a Pulmonary Hypertension Referral Centre, RESPIRE = Repeatability and Sensitivity to Change of Noninvasive End Points in Pulmonary Arterial Hypertension, RHC 5 right heart catheter.

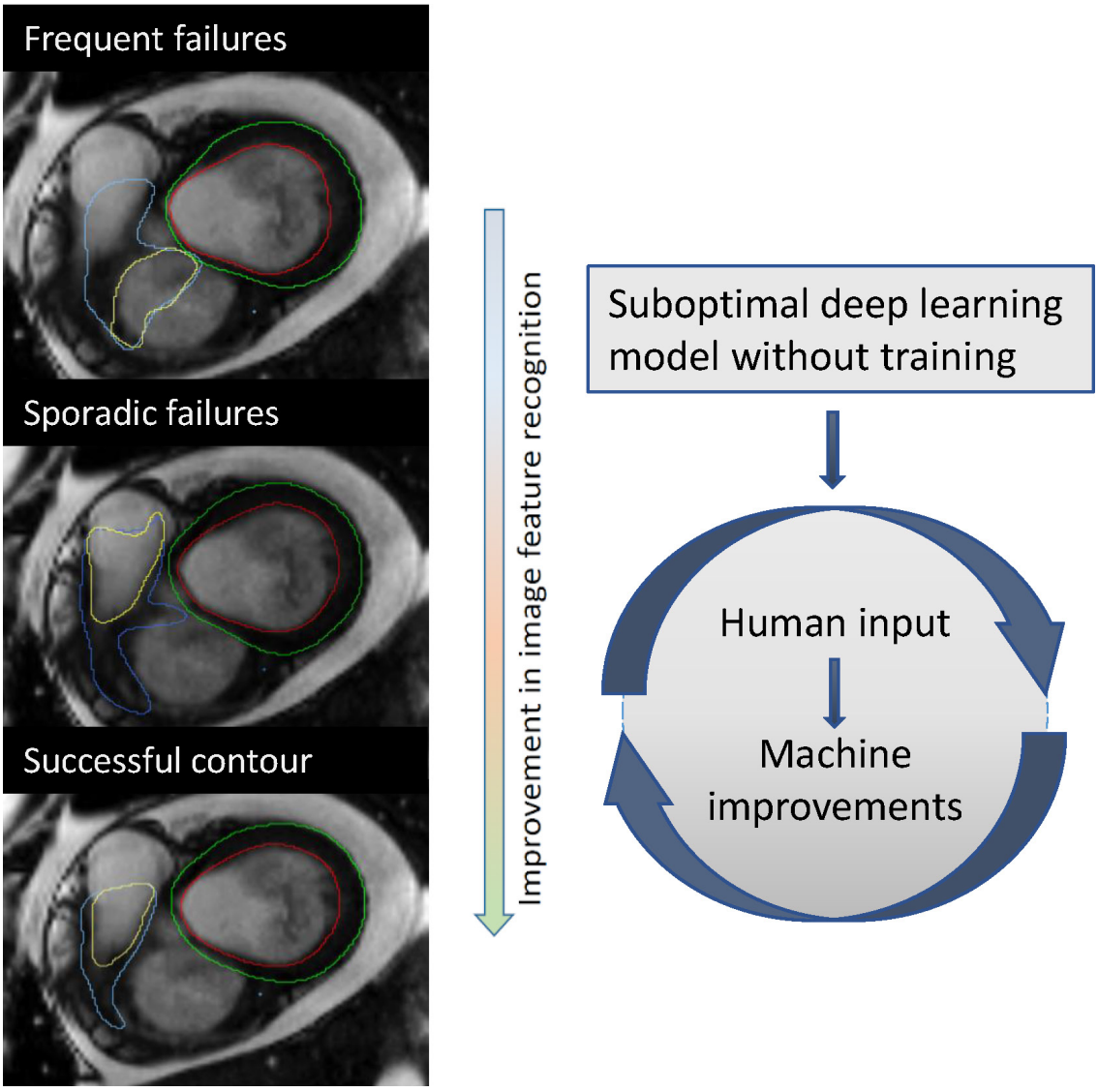


Figure 4.2: Example of improvement following additional training. This example demonstrates improvement of the right ventricular base after additional training. The first model missed the right ventricular outflow tract and included the right atrium instead (top image: yellow annotation showing right ventricular endocardial border), whereas the final model correctly included the right ventricular outflow tract and excluded the right atrium (bottom image).

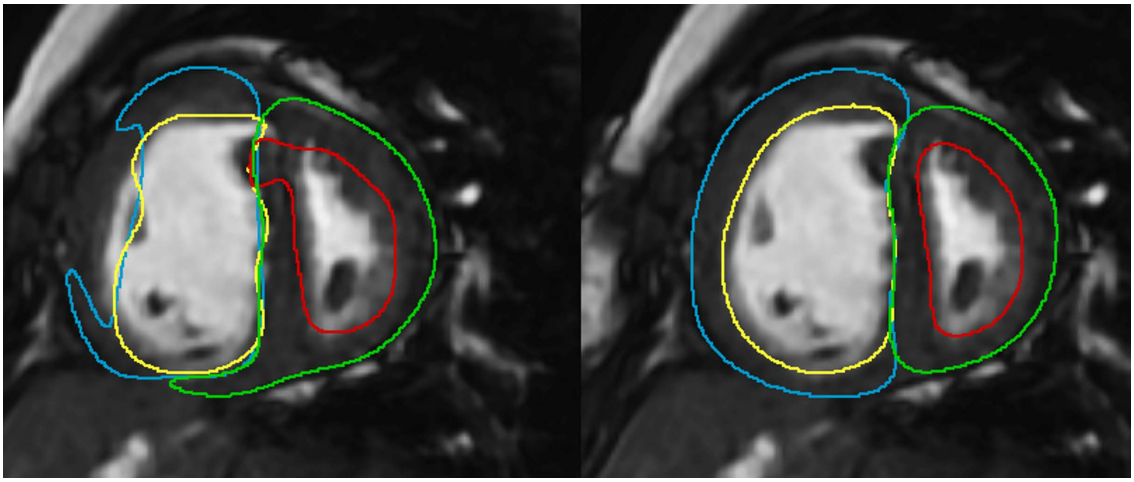


Figure 4.3: Improvement of automated Segmentation - with additional RV training (pre-Sheffield training on the left and post-training on the right)

in the blood pool, and the outflow tract was included for both the RV and LV. In the ASPIRE cohort, trabeculations were excluded from the blood pool (performed by D.C.). Manual segmentation for the Dice accuracy analysis was performed independently (by K.K. and A.J.S.), and for the external cohort testing (by S.A. and A.J.S). The scan–rescan segmentations were performed by a senior cardiac MRI radiographer with 3 years of experience who was not involved in the model training. All manual contouring was performed blinded to the clinical data and RHC results. Software was used for manual contouring (MASS, research version 2020; Leiden University Medical centre). A visual quality review for all segmentation was performed together by two authors (S.A. and A.J.S.) to identify the failure rate of the final segmentation model. Failed segmentations were those that resulted in visually unacceptable contours and would lead to incorrect measurement. Contours with minimal errors that were deemed to not affect cardiac MRI measurements were labelled suboptimal segmentations.

Image Storage

All Sheffield CMR scans were stored in POLARIS XNAT , which is a secure online platform hosted at the servers of the Academic Unit of Radiology at the University of Sheffield. Polaris XNAT stores pseudonymised medical imaging data and is only accessible from on-campus or via a virtual private network (VPN) or through a secure externally-accessible XNAT server that can be shared with collaborators. Every added CMR scan is anonymised and uploaded in Digital Imaging and Communications in Medicine (DICOM) format. Each

patient in the XNAT database has a unique identifier that is linked to a Sheffield Teaching Hospitals (STH) identifier number. The document linking both identifiers are stored on an encrypted and password protected departmental computer and on the University of Sheffield password protected and encrypted network storage system.

AI Model Development

The convolutional neural network (CNN) used for the experiments had a UNET–like architecture [193] that in turn, is based on a deep learning framework for medical image analysis called NiftyNet [194]. The model implementation was similar to ResUNet∗ with 16 convolutional layers, including residual learning units, and was implemented by using Python (version 3.6.9; Python Software Foundation) and TensorFlow (version 1.12) [195]. Input images were resampled to a fixed pixel spacing of 1 mm and cropped to a 256×256 matrix size and zero–filled when required. For training, the Adam optimizer method was used, the learning rate was selected as 0.001, and cross–entropy was used as the loss function. Data augmentation was performed by creating new training samples by randomly rotating, flipping, shifting, and modifying the image intensities of the original images. Each training batch included a random selection of 20 images. The fixed number of epochs was set at 30, with all images used once during every epoch. The raw output of the CNN is a per–section 2D binary mask for every anatomical structure. Postprocessing of the raw CNN output is performed using the VTK Visualization Toolkit [196]. First, taking into account the known section positioning, connected component analysis is used to extract the largest three–dimensional ($3D$) component for every structure. The marching cubes algorithm is then applied to convert the extracted object into a $3D$ surface mesh. Irregularities in the surface mesh are reduced by applying a mesh smoothing operation. Finally, per–section contours are generated by computing the intersection of every image section with the $3D$ mesh. No correction for section misalignment was performed. Single–centre, mixed LV pathology data of a cohort of 211 patients from Budapest were used to train an initial biventricular segmentation network. In our study, cohort LV and RV endocardial and epicardial contours were defined in the end–diastolic and end–systolic phases by a cardiac radiologist with > 19 years of experience (A.T.). As this model is based on cardiac MRI

∗ https://github.com/dmolony3/ResUNet

data from a single MRI system (Philips), the initial AI model was further developed in an offline HITL approach with the aim of creating a CNN free from vendor, centre and patient pathology bias. Two steps of interactive training (200 scans per step) were completed in a 347−patient multipathology cohort with 400 multivendor (250 GE/150 Siemens) studies (Figure 4.2). Automatic contours were checked after each iteration of training, and subsequent training was targeted on refining suboptimal or failed segmentations. To optimally benefit from the available imaging data, in the incremental model improvement process, the segmented images of all cardiac phases were included in the training set. This resulted in a large number (113, 192) of images in the training dataset. All experiments were executed on a standard PC with an Intel Core $i7$ CPU with 64 GB of internal RAM memory equipped with an Nvidia GTX 1080 TI GPU with 12 GB of memory. The time needed to train the final model was 33 hours, while the time needed for automated segmentation of all cardiac phases and sections of a single scan, including extraction of quantitative results, was less than 30 seconds.

Statistical Analysis

Continuous variables are presented as proportions, means ± SDs, or medians with interquartile ranges for data with asymmetric distributions. Variable standardization was performed to allow comparison of the different continuous variables on the same scale by subtracting the mean for each variable and dividing it by its SD. Cardiac MRI volumetric measurements were indexed for body surface area. Measurements were corrected for age and sex by calculating the percentage predicted values per published reference data [197, 198]. The interstudy repeatability was assessed with interclass correlation coefficient (ICC) and Bland–Altman analysis to compare the scan–rescan variation in the automated and manual cardiac MRI measurements. The paired t test was calculated to compare the differences in scan–rescan measurements between AI and manual assessment. Spearman correlation coefficient was used to compare the LV stroke volume to the stroke volume derived from RHC and the aortic forward flow volume at the LV outflow tract measured by phase–contrast imaging. RHC stroke volume was derived by dividing cardiac output by heart rate. RV ejection fraction and ventricular mass index were correlated to RHC pulmonary vascular resistance and mean pulmonary artery pressure. Ventricular mass

index was calculated as the RV end–diastolic mass–to–LV end–diastolic mass ratio (RV mass–to–LV mass ratio). The z test using the method by Steiger [199] was performed to test for differences between manual and AI correlations with RHC and phase–contrast imaging. Uni– and multivariable Cox proportional hazard regression hazard ratios were calculated for both the age– and sex–adjusted cardiac MRI parameters. Collinearity was tested by using Spearman correlation test. A correlation of r greater than 0.8 was considered to be closely related. All patients were followed up until the all–cause mortality or administrative censoring date (June 20, 2021). No patient was lost to follow–up. The interstudy repeatability was assessed with ICC and Bland–Altman analysis to compare the scan–rescan variation in the automated and manual cardiac MRI measurements. The paired t test was calculated to compare the differences in the scan–rescan measurements between AI and manual assessment. The agreement between the AI and manual cardiac MRI measurements in the external test data set was analysed with ICC and Bland–Altman plots. The accuracy of the AI contours relative to the manual contours was estimated by calculating the Dice similarity coefficient in 30 studies from an internal test data set randomly chosen from the cohort and in the 40 studies from the external test data set. The Dice score measured the ratio of overlap and distance between the manual and automatically segmented areas; a higher value indicated better accuracy of the contouring model relative to the manual segmentation. Statistical analyses were performed by using the Pingouin (version 0.5) [200] and Lifelines (version 0.26) [201] Python libraries, and graphs were produced by using the Matplotlib library (version 3.5) [202]. A P value of .05 or less indicated statistical significance.

4.3 Results

4.3.1 Study Sample Characteristics

The total study included 4289 patients and 5630 cardiac MRI studies, after excluding 244 patients because of either incomplete or unretrievable imaging (Figure 4.4). The median age in the training data set was 58 years (IQR, 34 years), 66 years (IQR, 21 years) in the ASPIRE cohort, 67 years (IQR, 19 years) in the patients with same–day RHC, and 48 years (IQR, 22 years) in the scan–rescan patients. The ratios of women were as follows:

training data set, 315 of 539 (58%); ASPIRE cohort, 2158 of 3487 (61%); patients with same–day RHC, 131 of 178 (56%); and scan–rescan patients, 35 of 46 (78%) (Table 4.1 and Table 4.2).

Table 4.1: Baseline characteristics of the training set.

Parameter	***Training 1*** ***N = 192***	***Training 2*** ***N = 170***	***Training 3*** ***N = 177***
Centre	Budapest	Sheffield University	Sheffield Hospitals
MRI Vendor	Philips	GE	Siemens and GE
No. MRI scans	211	220	180
Age *(years)*	32 (23-55)	66 (51-77)	67 (54-75)
Sex *(female)*	83 (43%)	120 (70%)	112 (63%)
BSA *(m^2)*	1.89 ± 0.28	1.83 ± 0.21	1.83 ± 0.23
Diagnosis			
Left Heart Disease	138 (72%)	32 (19%)	21 (12%)
Lung Disease	NA	26 (15%)	13 (7%)
PAH	NA	48 (28%)	89 (50%)
CTEPH	NA	34 (20%)	10 (6%)
Other PH	NA	1 (0%)	7 (4%)
Other (not PH)	54 (28%)	29 (17%)	37 (21%)
Cardiac MRI			
RVEF *(%)*	55 ± 11	41 ± 12	42 ± 14
RVESVi *(ml/m^2)*	45 ± 24	65 ± 32	66 ± 38
RVEDVi *(ml/m^2)*	95 ± 31	107 ± 40	109 ± 45
RVEDMi *(g/m^2)*	24 ± 7	25 ± 9	26 ± 10
LVEF *(%)*	53 ± 13	52 ± 10	53 ± 11
LVESVi *(ml/m^2)*	49 ± 32	38 ± 17	37 ± 23
LVEDVi *(ml/m^2)*	99 ± 32	80 ± 26	76 ± 30
LVSVi *(ml/m^2)*	50 ± 11	42 ± 15	39 ± 13
VMI *(ratio)*	0.35 ± 0.12	0.51 ± 0.21	0.46 ± 0.19

Note.—Unless otherwise indicated, data are numbers of patients; data in parentheses are percentages. Mean data are ± standard deviation.
Abbreviations in Table Table 4.1 and Table 4.2: BSA = body surface area, CO = cardiac output, CTEPH = chronic thromboembolic pulmonary hypertension, EDMi = end-diastolic mass index, EDVi = end-diastolic volume index, EF = ejection fraction, ESVi = end-systolic volume index, ISWT = incremental shuttle walking test, LV = left ventricle, mPAP = mean pulmonary artery pressure, NA, not applicable; PAH = pulmonary arterial hypertension, PAWP = pulmonary arterial wedge pressure, PH = pulmonary hypertension, PVR = pulmonary vascular resistance, RHC = right heart catheterisation, RV = right ventricle, SvO2 = mixed venous oxygen saturation, VMI = ventricular mass index, WHO = World Health Organization.

4.3.2 Quality Control

An example of the AI segmentation of the short-axis stack throughout the cardiac cycle is shown in Figure 4.5. The overall failure rate of the automatic segmentation was 1.0% (53 of 5316), almost exclusively caused by congenital heart diseases such as a ventricular septal defect (Figure 4.6A) or artifacts and technical issues affecting image quality. In 91

Table 4.2: Baseline characteristics of the testing set

	ASPIRE ***n = 3487***	***Same-day RHC*** ***n = 178***	***Repeatability*** ***n = 46***
Age *(years)*	66 (53-74)	67 (56-75)	48 (40-62)
Sex *(female)*	2158 (61%)	131 (56%)	35 (76%)
BSA *(m^2)*	1.88 ± 0.24	1.92 ± 0.29	1.89 ± 0.20
Diagnosis			
Left Heart Disease*	741 (21%)	28 (12%)	NA
Lung Disease*	480 (13%)	29 (12%)	NA
PAH	920 (26%)	49 (21%)	36 (77%)
CTEPH	623 (19%)	77 (33%)	NA
Other PH	88 (3%)	7 (3%)	NA
Other (not PH)	635 (18%)	49 (21%)	10 (33%)
WHO functional class			
I	42 (1%)	5 (2%)	NA
II	403 (11%)	30 (13%)	2 (6%)
III	2501 (71%)	182 (77%)	30 (83%)
IV	314 (9%)	15 (6%)	4 (11%)
ISWT distance *(m)*	225 (193)	249 (212)	518 (321)
RHC			
mPAP *(mmHg)*	42 (21)	36 (22)	54 (16)
PVR *(dyns.s.cm^{-5})*	460 (562)	429 (587)	NA
PAWP *(mmHg)*	12 (6)	10 (6)	NA
SV *(ml)*	62 (31)	61 (34)	64 (34)
CO *(L/min)*	4.95 (1.89)	4.70 (1.59)	4.51 (1.66)
SvO_2 *(%)*	66 (12)	68 (12)	64 (14)
Cardiac MRI			
RVEF *(%)*	40 ± 13	41 ± 13	43 ± 9
RVESVi *(ml/m^2)*	64 ± 35	62 ± 46	60 ± 27
RVEDVi *(ml/m^2)*	104 ± 41	102 ± 70	104 ± 34
RVEDMi *(g/m^2)*	25 ± 9	25 ± 14	26 ± 7
LVEF *(%)*	53 ± 10	54 ± 9	59 ± 7
LVESVi *(ml/m^2)*	35 ± 14	36 ± 20	32 ± 10
LVEDVi *(ml/m^2)*	74 ± 21	77 ± 40	77 ± 16
LVSVi *(ml/m^2)*	39 ± 12	41 ± 22	45 ± 10
LVEDMi *(g/m^2)*	50 ± 13	51 ± 25	45 ± 8
VMI *(ratio)*	0.51 ± 0.17	0.50 ± 0.16	0.58 ± 0.19

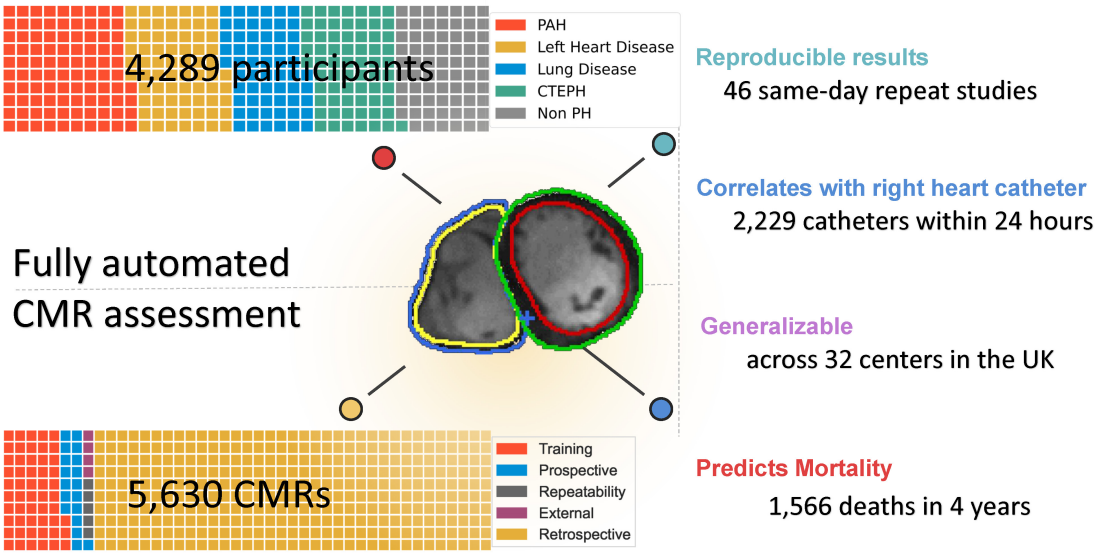


Figure 4.4: Summary of datasets and main results.

of 5316 studies (1.7%), there were segmentation errors mainly affecting the heart apex (Figure 4.6).

4.3.3 Correlations with Invasive haemodynamics and Phase Contrast Flow

The mean for cardiac MRI–estimated LV stroke volume were 78 mL ± 24 (SD) and 79 mL ± 26 for AI and manual assessments, respectively. The RHC–derived LV stroke volume was 66 mL ± 23 and the phase–contrast mean aortic forward flow volume was 68 mL ± 21. The correlation between RHC and cardiac MRI LV stroke volume (Figure 4.7) was higher for AI than for manual measurements ($r = 0.74$ vs 0.68, respectively; $P = .03$) (Figure 4.7A, Table 4.3).

Both AI and manually derived LV stroke volume showed similar correlation with the aortic forward flow volume ($r = 0.73$ and 0.70, respectively; $P = .29$; $n = 118$), although variability is evident between the methods of stroke volume calculation, which may in part be due to technical factors, intracardiac shunts or valvular abnormalities such as mitral or aortic regurgitation (Figure 4.7B). The AI–measured ventricular mass index (RV mass–to–LV mass ratio) had a higher correlation with pulmonary vascular resistance (Figure 4.7D) and mean pulmonary artery pressure (Figure 4.7F) than the manual measurements ($r = 0.64$ vs 0.44 [$P < .001$] and 0.56 vs 0.37 [$P < .001$], respectively).

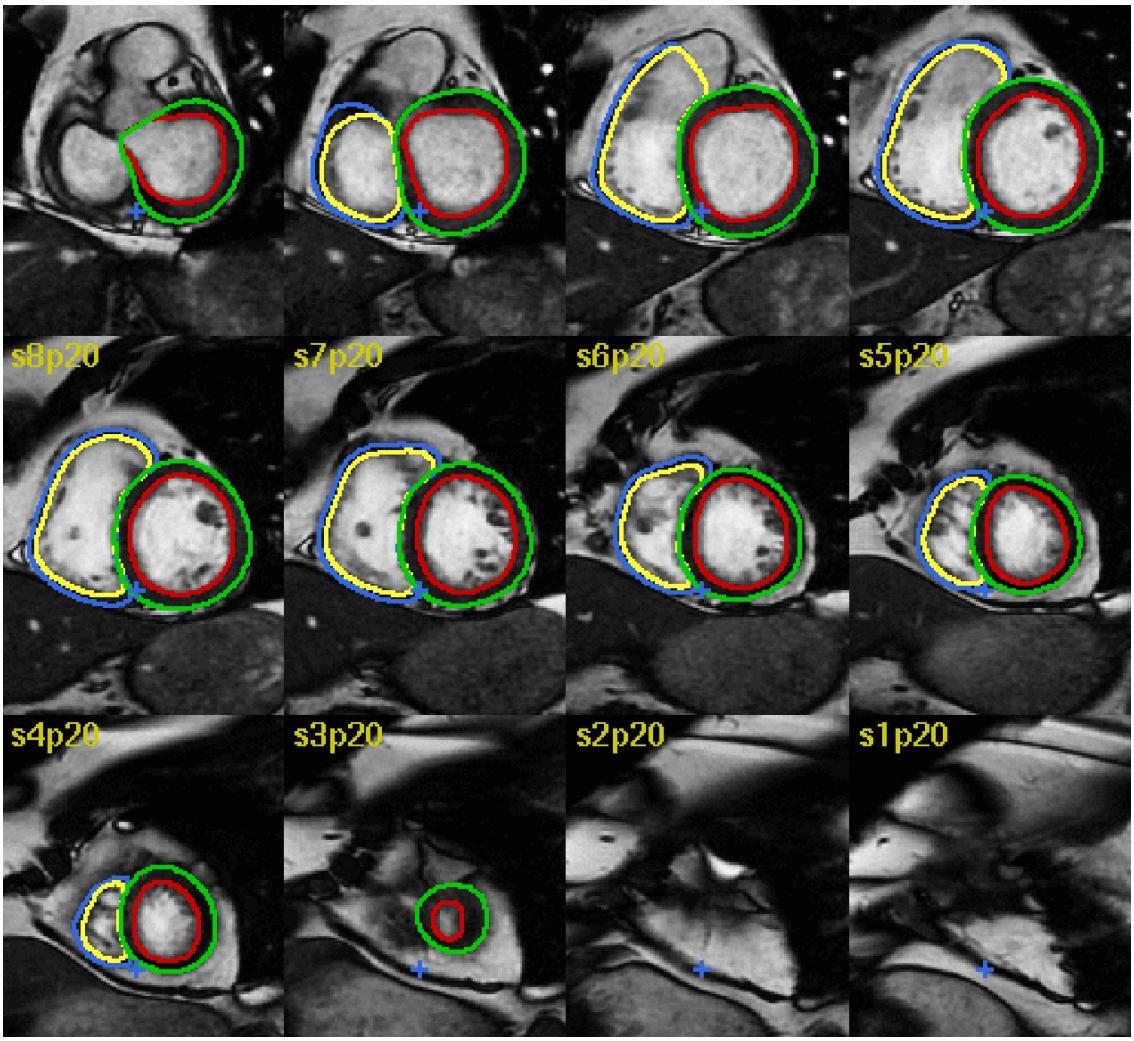


Figure 4.5: Example of an automatically segmented cardiac MRI with RV dysfunction and RV dilatation. The AI model was particularly accurate at segmenting the RV base, which is the area typically most challenging for manual assessors.

Table 4.3: The relationship between cardiac MRI, right heart catheterisation parameters and phase encoding aortic forward flow volume in the same-day RHC cohort

RHC	*Cardiac MRI*	*AI (r Value)*	*Human (r Value)*	*P Value*	*No.*
SV *(ml)*	**LVSV** *(ml)*	0.74	0.68	**.03**	178
PVR *(dyns.s.cm^{-5})*	**VMI**	0.62	0.41	**<.001**	178
PVR *(dyns.s.cm^{-5})*	**RVEF** *(%)*	-0.70	-0.69	.75	178
mPAP *(mmhg)*	**VMI**	0.56	0.37	**<.001**	178
mPAP *(mmhg)*	**RVEF** *(%)*	-0.66	-0.67	.76	178
Aortic flow *(ml)*	**LVSV** *(ml)*	0.73	0.70	.29	117

Note.—Human evaluation was performed by a senior cardiac MRI radiographer with 17 years of experience.
AI = artificial intelligence, EF = ejection fraction, LVSV = left ventricle stroke volume, mPAP = mean pulmonary artery pressure, NA = not applicable, PVR = pulmonary vascular resistance, RHC = right heart catheterisation, RV = right ventricle, VMI = ventricular mass index.

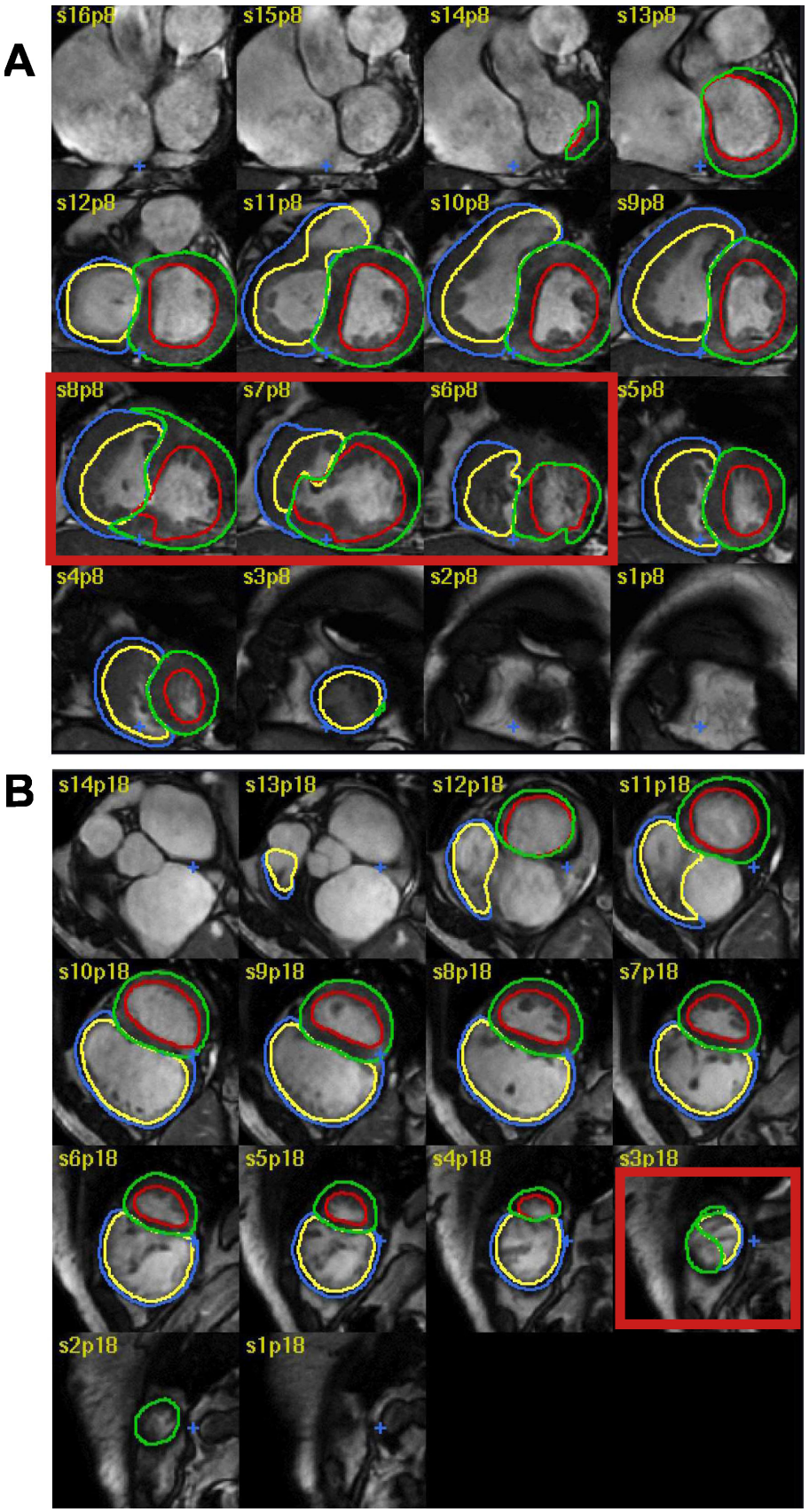


Figure 4.6: Examples of failed and suboptimal artificial intelligence (AI) segmentations. (A) Major failure because of congenital heart disease causing the left ventricular (LV) contours to extend into the right ventricle (RV; red box). (B) Minor failure at the apex where the RV was incorrectly labelled as LV (red box). The red, green, blue, and yellow circles indicate the LV endocardial, LV epicardial, RV endocardial, and RV epicardial contours, respectively.

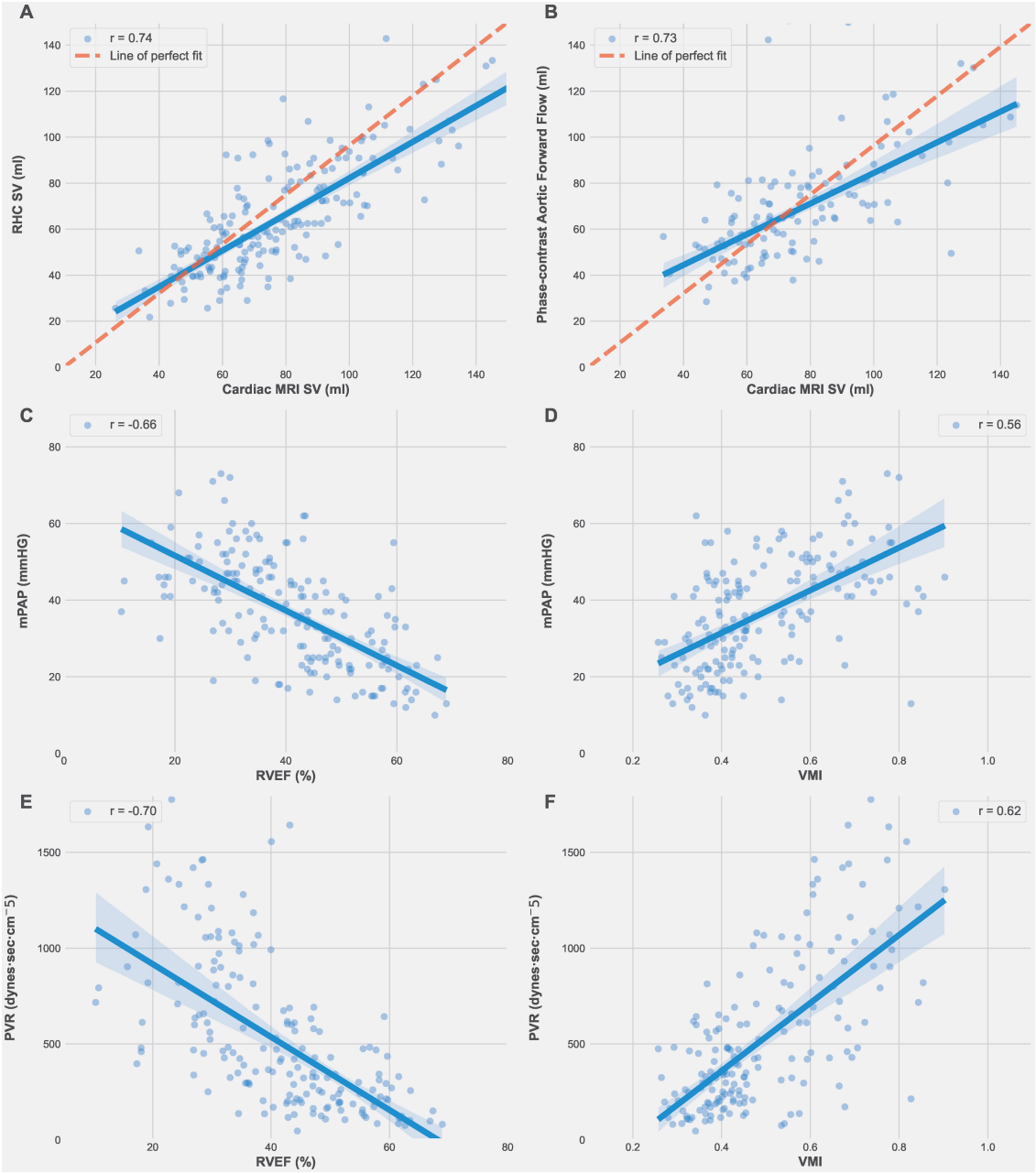


Figure 4.7: Graphs show the relationship between automatic cardiac MRI measurements, right heart catheterisation (RHC) and phase-contrast aortic flow. Automatic cardiac MRI measurements were compared to (A) RHC stroke volume (SV) and (B) phase-contrast aortic flow in 178 patients of the same-day RHC cohort. (C) Mean pulmonary artery pressure (mPAP) was compared with right ventricle ejection fraction (RVEF) and (D) ventricular mass index (VMI; RV mass–to–). (E) Pulmonary vascular resistance (PVR) was compared to RVEF and (F) VMI.

There were good correlations between RV ejection fraction and mean pulmonary artery pressure (Figure 4.7C) and pulmonary vascular resistance (Figure 4.7E), with no evidence of a difference between AI and manual readings. Pulmonary vascular resistance and mean pulmonary artery pressure correlated similarly with AI and manual RV ejection fraction ($P = .75$ and .76, respectively). The correlation between AI–based cardiac MRI measurements and RHC was confirmed in 2051 patients in the ASPIRE cohort (Table 4.4, Figure 4.8).

Table 4.4: The Relationship Between Automatic Cardiac MRI Measurements and Right Heart Catheterisation In the ASPIRE Registry Cohort

RHC	***Cardiac MRI***	**AI (r Value)**	***P Value***	**No.**
SV *(ml)*	**LVSV** *(ml)*	0.72	**<.001**	2044
PVR *(dyns.s.cm^{-5})*	**VMI**	0.70	**<.001**	1941
PVR *(dyns.s.cm^{-5})*	**RVEF** *(%)*	-0.65	**<.001**	1941
mPAP *(mmhg)*	**VMI**	0.65	**<.001**	2051
mPAP *(mmhg)*	**RVEF** *(%)*	-0.60	**<.001**	2051

Note.— Right heart catheterisation was performed on the same day of the cardiac MRI (in 80% of patients) or within 24 hours. RHC, right heart catheterisation; mPAP, mean pulmonary artery pressure; PVR, pulmonary vascular resistance; SV, stroke volume; LVSV, left ventricle stroke volume; RVEF, right ventricle ejection fraction; VMI, ventricular mass index

4.3.4 Mortality Prediction

Automatic cardiac MRI measurements were assessed in 3487 patients from the ASPIRE registry. The study population included patients with multiple pathologic diseases, predominantly pulmonary arterial hypertension (920 of 3487; 26%), left heart disease (741 of 3487; 21%), lung diseases (480 of 3487; 13%), chronic thromboembolic pulmonary hypertension (623 of 3487; 19%), and without pulmonary hypertension (635 of 3487; 18%). During the mean follow–up period (3.8 years) 1604 of 3487 (46%) patients died. Other than RV stroke volume, all cardiac MRI parameters predicted mortality (Table 4.5). RV parameters including RV mass were prognostic markers in the subgroup with pulmonary arterial hypertension ($n = 920$) (Table 4.5). RV ejection fraction remained a significant prognostic marker in a multivariable analysis including age, World Health Organization function class, incremental shuttle walking test, RHC parameters (mean pulmonary artery pressure, pulmonary arterial wedge pressure, cardiac output, and mixed venous oxygen saturation) and cardiac MRI variables (age– and sex–corrected RV and LV ejection fraction and mass index) (Table 4.5).

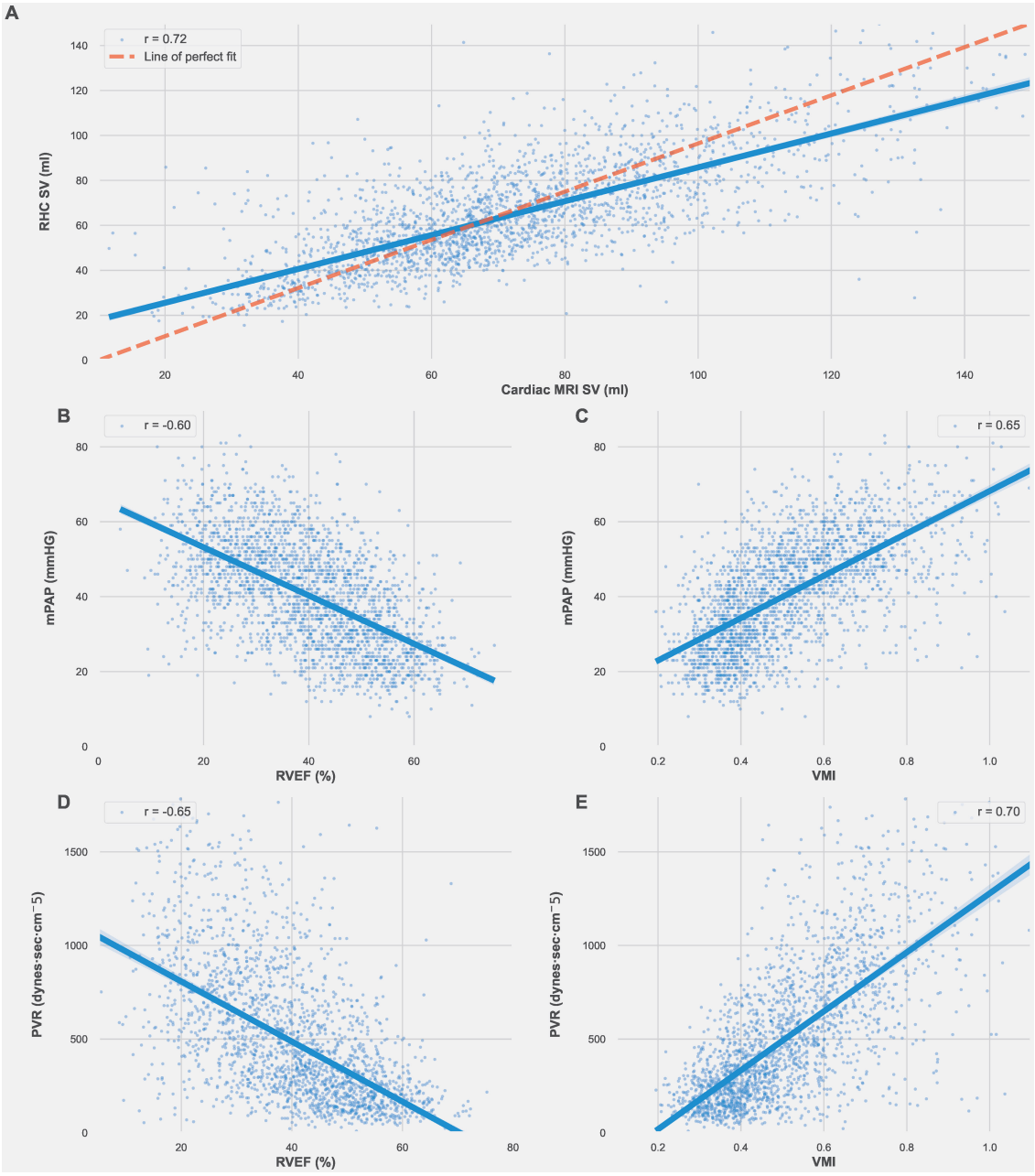


Figure 4.8: The relationship between automatic cardiac MRI measurements and right heart catheterisation in the ASPIRE cohort. Right heart catheterisation was performed 2,051 patients within the same day of cardiac MRI (80% of patients) or within 24 hours.

Table 4.5: Uni– and multivariable Cox regression hazard ratios for AI cardiac MRI measurements

	ASPIRE Cohort *(n= 3487)*			***PAH Subgroup*** *(n=920)*		
Parameter	**Hazard Ratio**		***P Value***	**Hazard Ratio**		***P Value***
Univariable Cox regression						
Age *(years)*	1.04	(1.03, 1.04)	**<.001**	1.04	(1.04, 1.05)	**<.001**
Sex	1.38	(1.26, 1.51)	**<.001**	1.16	(0.96, 1.41)	**.13**
WHO functional class	1.60	(1.52, 1.68)	**<.001**	1.35	(1.23, 1.47)	**<.001**
ISWT distance *(m)*	0.46	(0.42, 0.51)	**<.001**	0.49	(0.42, 0.57)	**<.001**
mPAP *(mmHg)*	1.25	(1.18, 1.32)	**<.001**	0.87	(0.77, 0.98)	**.02**
PVR *(dyns.s.cm^{-5})*	1.23	(1.17, 1.30)	**<.001**	0.99	(0.88, 1.12)	.89
PAWP *(mmHg)*	1.13	(1.06, 1.19)	**<.001**	0.97	(0.85, 1.10)	.62
SV *(ml)*	0.75	(0.71, 0.80)	**<.001**	0.82	(0.72, 0.94)	**<.001**
CO *(L/min)*	0.75	(0.70, 0.81)	**<.001**	0.84	(0.74, 0.96)	**.01**
SvO2 *(%)*	0.75	(0.71, 0.79)	**<.001**	0.82	(0.73, 0.93)	**<.001**
RVEF *(%predict)*	0.656	(0.63, 0.69)	**<.001**	0.76	(0.69, 0.84)	**<.001**
RVESVi *(%predict)*	1.493	(1.44, 1.55)	**<.001**	1.40	(1.28, 1.52)	**<.001**
RVEDVi *(%predict)*	1.360	(1.31, 1.40)	**<.001**	1.19	(1.09, 1.29)	**<.001**
RVSVi *(%predict)*	0.963	(0.92, 1.01)	.131	0.95	(0.87, 1.03)	.199
RVEDMi *(%predict)*	1.393	(1.34, 1.45)	**<.001**	1.15	(1.06, 1.26)	**.001**
LVEF *(%predict)*	0.795	(0.76, 0.83)	**<.001**	0.94	(0.85, 1.04)	.231
LVEDVi *(%predict)*	0.943	(0.90, 0.99)	**.019**	1.00	(0.91, 1.10)	.564
LVESVi *(%predict)*	1.113	(1.07, 1.16)	**<.001**	1.07	(0.98, 1.18)	.115
LVSVi *(%predict)*	0.811	(0.77, 0.85)	**<.001**	0.95	(0.86, 1.05)	.881
LVEDMi *(%predict)*	1.080	(1.03, 1.13)	**.001**	1.06	(0.96, 1.17)	.238
VMI *(%predict)*	1.364	(1.31, 1.43)	**<.001**	1.12	(1.03, 1.24)	**.010**
Multivariable Cox regression						
Age *(years)*	1.04	(1.03, 1.04)	**<.001**	1.04	(1.04, 1.06)	**<.001**
WHO functional class	1.50	(1.21, 1.85)	**<.001**	0.87	(0.54, 1.41)	.58
ISWT distance *(m)*	0.58	(0.51, 0.67)	**<.001**	0.62	(0.49, 0.79)	**<.001**
mPAP *(mmHg)*	1.07	(0.95, 1.21)	.27	0.80	(0.62, 1.04)	.09
PAWP *(mmHg)*	0.97	(0.89, 1.06)	.54	1.02	(0.86, 1.20)	.85
CO *(L/min)*	0.99	(0.90, 1.09)	.87	1.01	(0.80, 1.29)	.91
SvO_2 *(%)*	0.94	(0.85, 1.04)	.24	0.90	(0.72, 1.12)	.35
RVEF *(%predict)*	0.80	(0.70, 0.92)	**<.001**	0.70	(0.53, 0.92)	**.01**
RVEDMi *(%predict)*	1.03	(0.90, 1.18)	.65	0.97	(0.71, 1.31)	.83
LVEF *(%predict)*	0.99	(0.90, 1.09)	.83	1.11	(0.92, 1.35)	.27
LVEDMi *(%predict)*	1.21	(1.04, 1.41)	**.01**	1.10	(0.77, 1.58)	.59

Note.—Data in parentheses are 95% CIs. The total number of deaths in the ASPIRE cohort and the subgroup of the ASPIRE cohort with pulmonary artery hypertension were 1604 of 3487 (46%) and 459 of 920 (50%), respectively. The combinations of end-diastolic, endsystolic, and stroke volume correlated highly with each other and with the ejection fraction for both the right ventricular and left ventricular measurements. Therefore, only ejection fraction was included in the multivariable Cox regression analysis. Cardiac MRI parameters were analysed as the percentage of predicted values for an age- and sex-matched healthy population.

CO = cardiac output, EF = ejection fraction, EDMi = end-diastolic mass index, EDVi = end-diastolic volume index, ESVi = end-systolic volume index, ISWT = incremental shuttle walking test, LV = left ventricle, mPAP = mean pulmonary artery pressure, PAH = pulmonary arterial hypertension, PAWP = pulmonary arterial wedge pressure, PVR = pulmonary vascular resistance, RV = right ventricle, SvO2 = mixed venous oxygen saturation, VMI = ventricular mass index, WHO = World Health Organization.

The uni– and multivariable Cox regression results for the available manual cardiac MRI measurements are provided in (Table 4.6).

Table 4.6: Uni– and multivariable Cox regression hazard ratios for manual cardiac MRI measurements

	ASPIRE Cohort (*n= 2633*)			*PAH Subgroup* (*n=746*)		
Parameter	**Hazard Ratio**		***P Value***	**Hazard Ratio**		***P Value***
Univariable Cox regression						
RVEF *(%predict)*	0.63	(0.59, 0.67)	**<.001**	0.70	(0.62, 0.79)	**<.001**
RVESVi *(%predict)*	1.35	(1.31, 1.39)	**<.001**	1.43	(1.28, 1.59)	**<.001**
RVEDVi *(%predict)*	1.20	(1.16, 1.23)	**<.001**	1.18	(1.08, 1.28)	**<.001**
RVSVi *(%predict)*	0.94	(0.87, 1.00)	.07	0.95	(0.84, 1.07)	.42
RVEDMi *(%predict)*	1.11	(1.06, 1.15)	**<.001**	1.05	(0.90, 1.23)	.52
LVEF *(%predict)*	0.84	(0.80, 0.89)	**<.001**	0.90	(0.80, 1.01)	.07
LVEDVi *(%predict)*	0.85	(0.78, 0.91)	**.019**	0.92	(0.79, 1.07)	.29
LVESVi *(%predict)*	1.04	(0.98, 1.11)	.16	1.02	(0.89, 1.18)	.74
LVSVi *(%predict)*	0.76	(0.71, 0.82)	**<.001**	0.87	(0.75, 1.01)	.08
LVEDMi *(%predict)*	1.03	(0.98, 1.10)	.26	1.28	(1.00, 1.64)	.05
VMI *(%predict)*	1.19	(1.12, 1.26)	**<.001**	1.01	(0.89, 1.14)	.91
Multivariable Cox regression						
Age *(years)*	1.04	(1.03, 1.05)	**<.001**	1.04	(1.02, 1.06)	**<.001**
WHO functional class	1.50	(1.17, 1.92)	**<.001**	0.79	(0.41, 1.52)	.58
ISWT distance *(m)*	0.56	(0.47, 0.66)	**<.001**	0.65	(0.48, 0.89)	**<.001**
mPAP *(mmHg)*	1.12	(0.99, 1.27)	.08	0.81	(0.61, 1.08)	.08
PAWP *(mmHg)*	0.95	(0.86, 1.06)	.35	0.90	(0.74, 1.10)	.35
CO *(L/min)*	0.96	(0.85, 1.08)	.49	0.90	(0.66, 1.23)	.49
SvO_2 *(%)*	0.94	(0.83, 1.06)	.33	0.77	(0.56, 1.06)	.33
RVEF *(%predict)*	0.75	(0.65, 0.85)	**<.001**	0.83	(0.64, 1.07)	**.01**
RVEDMi *(%predict)*	0.81	(0.67, 0.98)	**.03**	0.82	(0.57, 1.19)	**.03**
LVEF *(%predict)*	1.07	(0.96, 1.18)	.21	1.16	(0.92, 1.48)	.21
LVEDMi *(%predict)*	1.40	(1.15, 1.70)	**<.001**	1.80	(1.10, 2.96)	.59

Note.—Manual cardiac MRI measurements were available for 2633/3487 (76%) patients of the ASPIRE cohort and 746/920 (81%) patients of the PAH subgroup. The total number of deaths in the ASPIRE cohort with manual measurements available were 1044/2633 (40%) and 281/746 (38%) in the PAH subgroup. The results are not directly comparable to the automated segmentation results because manual segmentations were not available for a large number of patients and because of the differences in handling of the trabeculation. Trabeculations were excluded from the blood pool in the ASPIRE manual segmentation. The combinations of end-diastolic, end-systolic and stroke volume correlated highly with each other and with the ejection fraction for both the RV and LV measurements, therefore only ejection fraction was included in the multivariable Cox regression analysis. All cardiac MRI parameters were analysed as the percentage of predicted values for an age- and sex-matched normal population.

4.3.5 Repeatability Assessment

The interstudy repeatability of cardiac MRI measurements was high for both AI and manual measurements. The automatic LV and RV volumetric and mass measurements ICC were 0.92 and 0.99, respectively. The ICC for LV and RV ejection fraction was

0.80 and 0.90, respectively (Table 4.7). The differences in the scan–rescan measurements were not different between AI and manual (t test $P = .73$ for RV ejection fraction and .8 for LV ejection fraction) (Table 4.8). Bland–Altman plots showed strong agreement between manual and automatic measurements, with small mean absolute differences ranging between 0 mL and 4 mL in the scan–rescan measurements (Figure 4.9). Examples of MRI scans with higher differences are shown in Figure 4.10.

Table 4.7: Interstudy repeatability for automatic and manual cardiac MRI parameters.

Cardiac MRI	***AI***		***Manual***	
RVESV *(ml)*	0.99	(0.98, 0.99)	0.98	(0.96, 0.99)
RVEDV *(ml)*	0.98	(0.97, 0.99)	0.97	(0.95, 0.98)
RVSV *(ml)*	0.92	(0.85, 0.96)	0.84	(0.70, 0.91)
RVEF *(%)*	0.89	(0.80, 0.94)	0.78	(0.60, 0.88)
RVEDM *(g)*	0.98	(0.96, 0.99)	0.90	(0.81, 0.94)
LVESV *(ml)*	0.96	(0.92, 0.98)	0.96	(0.92, 0.98)
LVEDV *(ml)*	0.98	(0.97, 0.99)	0.96	(0.93, 0.98)
LVSV *(ml)*	0.95	(0.91, 0.97)	0.93	(0.88, 0.96)
LVEF *(%)*	0.79	(0.61, 0.88)	0.88	(0.78, 0.93)
LVEDM *(g)*	0.99	(0.98, 0.99)	0.94	(0.89, 0.97)

Note.—Data are interclass correlation coefficients; data in parentheses are 95% CIs. Interstudy repeatability assessment for the automatic and manual cardiac MRI measurement was performed in 46 participants in the RESPIRE cohort who had same-day repeat scans.
EDM = end-diastolic mass, EDV = end-diastolic volume, EF = ejection fraction, ESV = end-systolic volume, LV = left ventricle, RV = right ventricle.

4.3.6 External Testing

The external dataset included 40 patients from 32 different centres in the UK (Figure 4.11). The sample included 20 Siemens scans (Avanto 13, Aera 6, Symphony 1), 15 Philips scans (Achieva 11, Ingenia 4) and 5 GE scans (HDxt 3, Excite 1, Artist 1). Bland–Altman plots showed strong agreement between manual and automatic measurements, with small mean absolute differences ranging between 0 mL and 4 mL in the scan–rescan measurements (Figure 4.9). Examples of MRI scans with higher differences are shown in (Table 4.10). The ICC was 0.93 for LV ejection fraction and 0.94 for RV ejection fraction, and the ICCs for LV and RV mass were 0.95 and 0.92, respectively. Bland–Altman plots showed small absolute mean differences (Figure 4.12 and Figure 4.13).

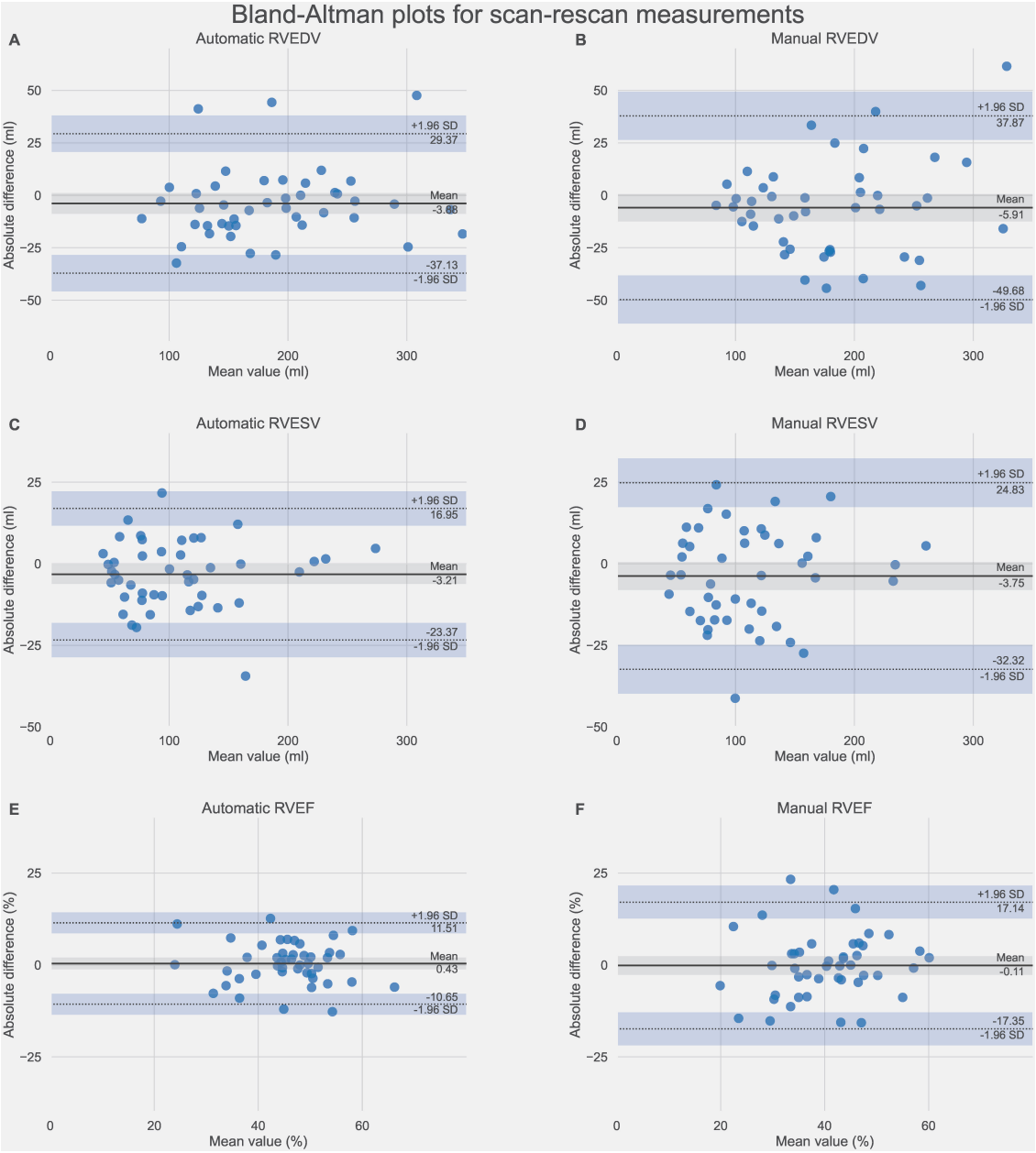


Figure 4.9: Bland–Altman plots of scan–rescan repeatability for the automatic compared to the manual right ventricular parameters. Same day scan–rescan cardiac MRIs were performed in 46 participants to compare the repeatability of the (A, C, D) automatic and (B, D, F) manual measurements. RVEDV = right ventricular end-diastolic volume, RVEF = right ventricular ejection fraction, RVESV = right ventricular end-systolic volume.

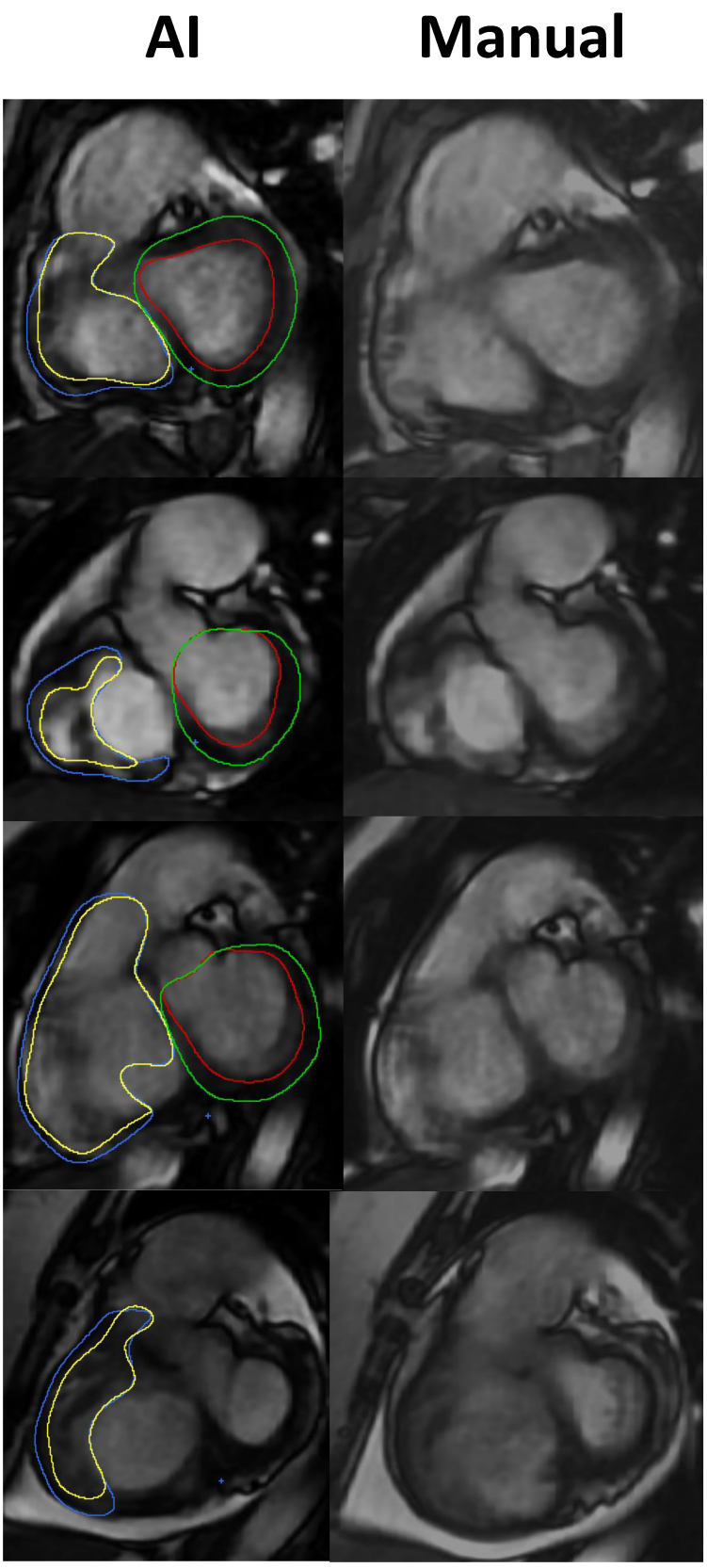


Figure 4.10: Examples of scan-rescan images highlighting the differences in the automatic vs manual segmentation The examples show cases where the manual reader has incompletely contoured the base explaining some of the volume difference between AI and manual measurements.

Table 4.8: Comparison between the automatic and manual cardiac MRI measurements in the scan-rescan repeatability cohort

Cardiac MRI Parameter	AI 1	AI 2	AI1 - AI2	Man 1	Man 2	Man 1 - Man 2	*P Value*
***RVEDV** (ml)*	193 ± 69	189 ± 72	4	187 ± 66	181 ± 70	6	.63
***RVESV** (ml)*	108 ± 53	105 ± 53	3	113 ± 50	109 ± 52	4	.84
***RVSV** (ml)*	85 ± 26	84 ± 27	1	74 ± 26	71 ± 29	3	.7
***RVEF** (%)*	45 ± 9	46 ± 9	1	41 ± 10	40 ± 11	1	.73
***RVEDM** (g)*	47 ± 15	46 ± 15	1	29 ± 12	29 ± 11	0	.52
***RVCO** (l/min)*	6 ± 2	6 ± 2	0	5 ± 2	5 ± 2	0	.79
***LVEDV** (ml)*	157 ± 43	156 ± 45	1	142 ± 41	140 ± 45	2	.9
***LVESV** (ml)*	66 ± 26	67 ± 24	1	52 ± 23	51 ± 24	1	.8
***LVSV** (ml)*	91 ± 21	89 ± 25	2	90 ± 24	89 ± 27	1	.94
***LVEF** (%)*	59 ± 7	58 ± 7	1	65 ± 9	64± 9	1	.8
***LVEDM** (g)*	90 ± 21	90 ± 23	0	70 ± 20	72 ± 24	2	.19
***LVCO** (l/min)*	6 ± 1	6 ± 2	0	6 ± 2	6 ± 2	0	.57

Note.—Table shows the mean and standard deviation for the AI and Manual cardiac MRI measurements with the absolute mean difference ± SD between the first and second MRI study for the AI and manual assessment in 46 patients of the RESPIRE cohort. The paired t test compares the scan-rescan variation between the automatic and manual measurements and shows no difference in the repeatability of the AI model and manual assessor.
Man, manual; RV, right ventricle; LV, left ventricle; ESV, end-systolic volume; EDV, end-diastolic volume; ESV, end-systolic volume; SV, stroke volume; EF, ejection fraction; EDM, end-diastolic mass; CO, cardiac output

Table 4.9: DICE scores in the internal and external test cohorts.

DICE score	Internal Test (n=30)	External Test (n=40)
LV ENDO ED	0.96	0.95
LV ENDO ES	0.93	0.89
LV EPI ED	0.96	0.95
LV EPI ES	0.95	0.94
RV ENDO ED	0.95	0.91
RV ENDO ES	0.93	0.88
RV EPI ED	0.95	0.92
RV EPI ES	0.93	0.89

The internal test DICE was randomly chosen from the ASPIRE cohort (n= 30). The DICE score measures the ratio of overlap and distance between the manual and automatically segmented areas. A higher DICE value indicates a better accuracy of the contouring model and stronger agreement with the manual segmentation. The DICE analysis showed excellent agreement in the AI and manual LV and RV epi- and endocardial end-systolic and end-diastolic contours in both the internal and external test cohorts.
RV, right ventricle; LV, left ventricle; ED, end-diastolic; ES, end-systolic; ENDO, endocardial contours; EPI, epicardial contours.

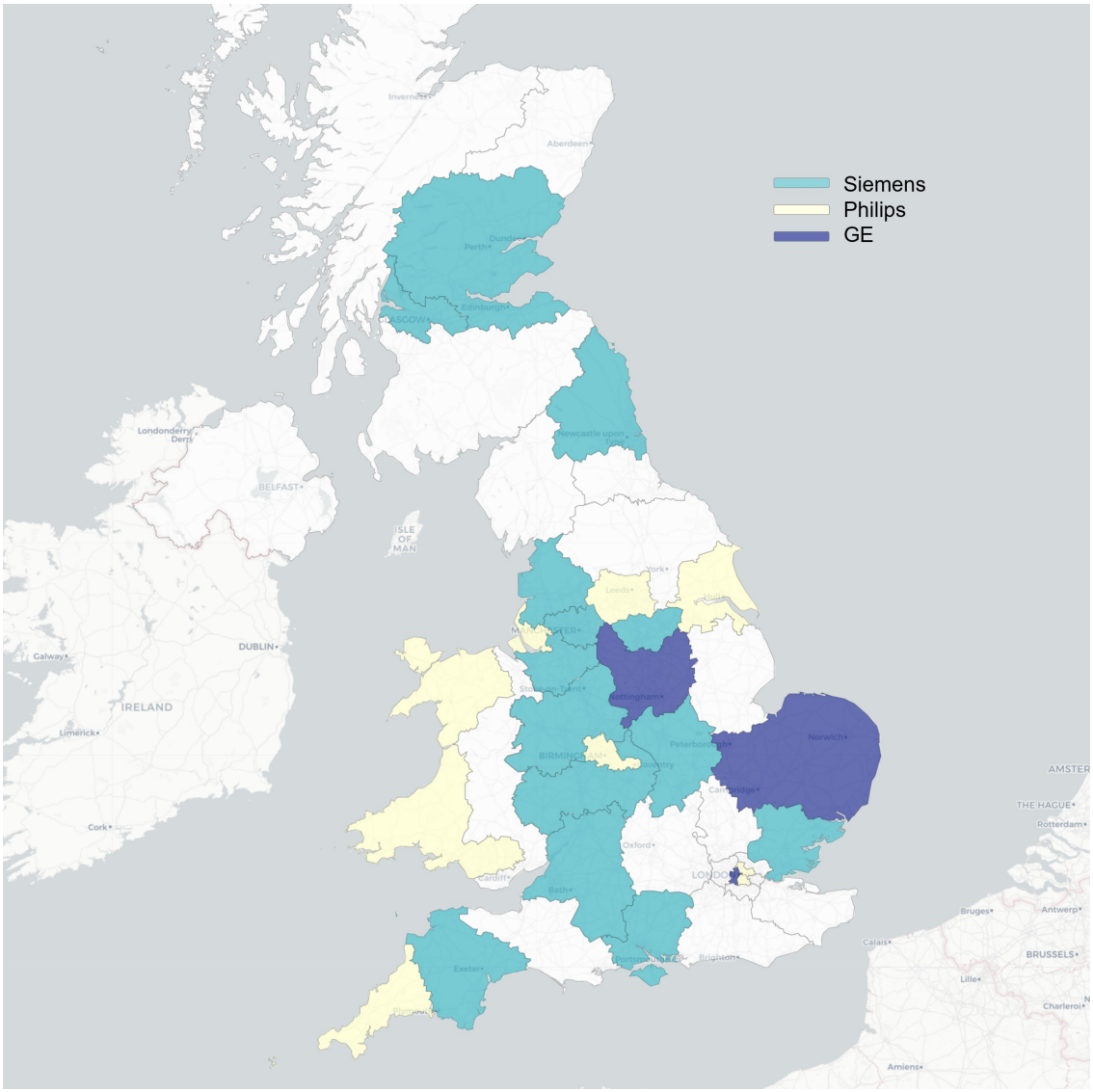


Figure 4.11: External dataset map. The external dataset was imaged at different centers in the UK using different MRI vendors and systems.

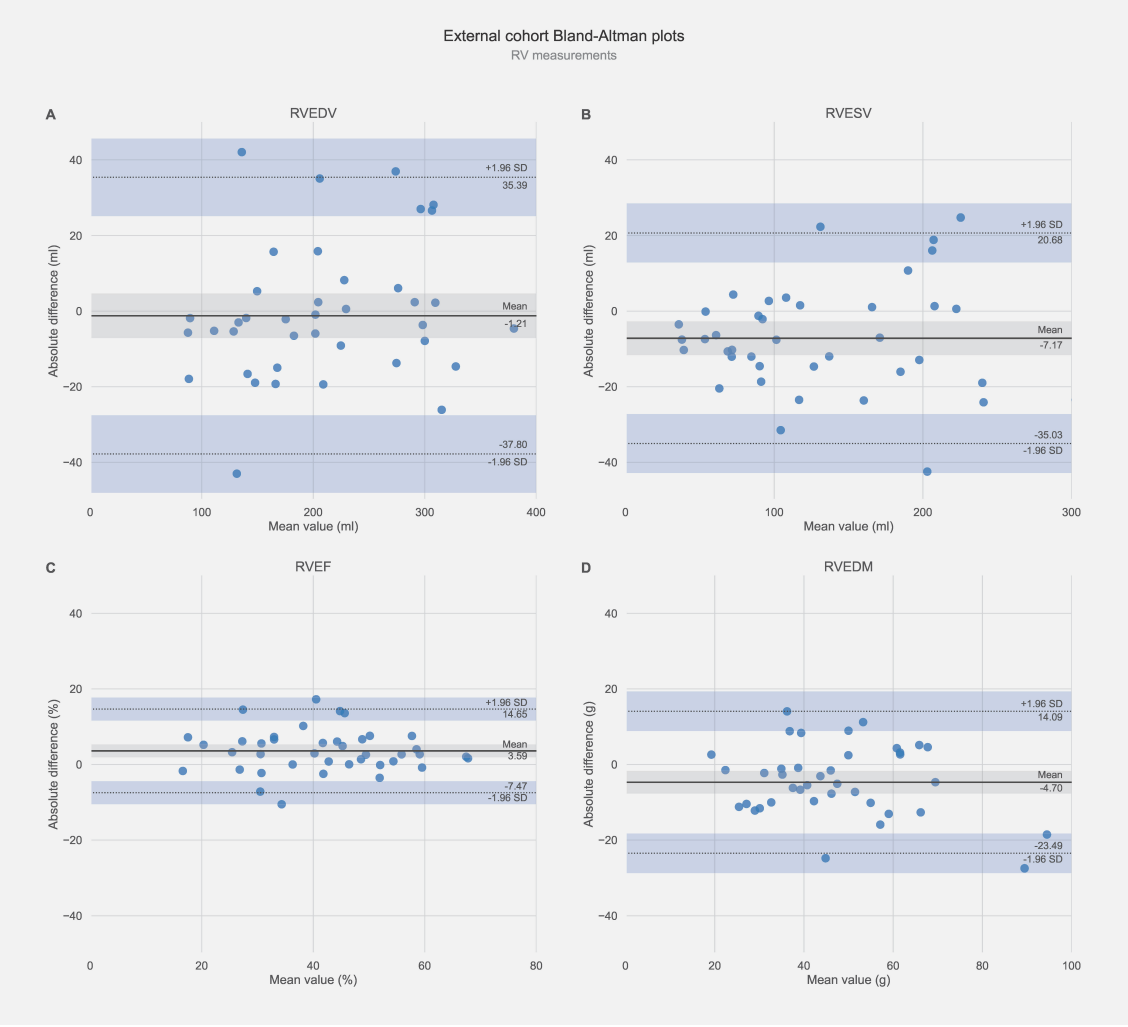


Figure 4.12: Bland–Altman plots showing the absolute mean difference between AI and manual right ventricular (RV) measurements in the external cohort ($n = 40$).

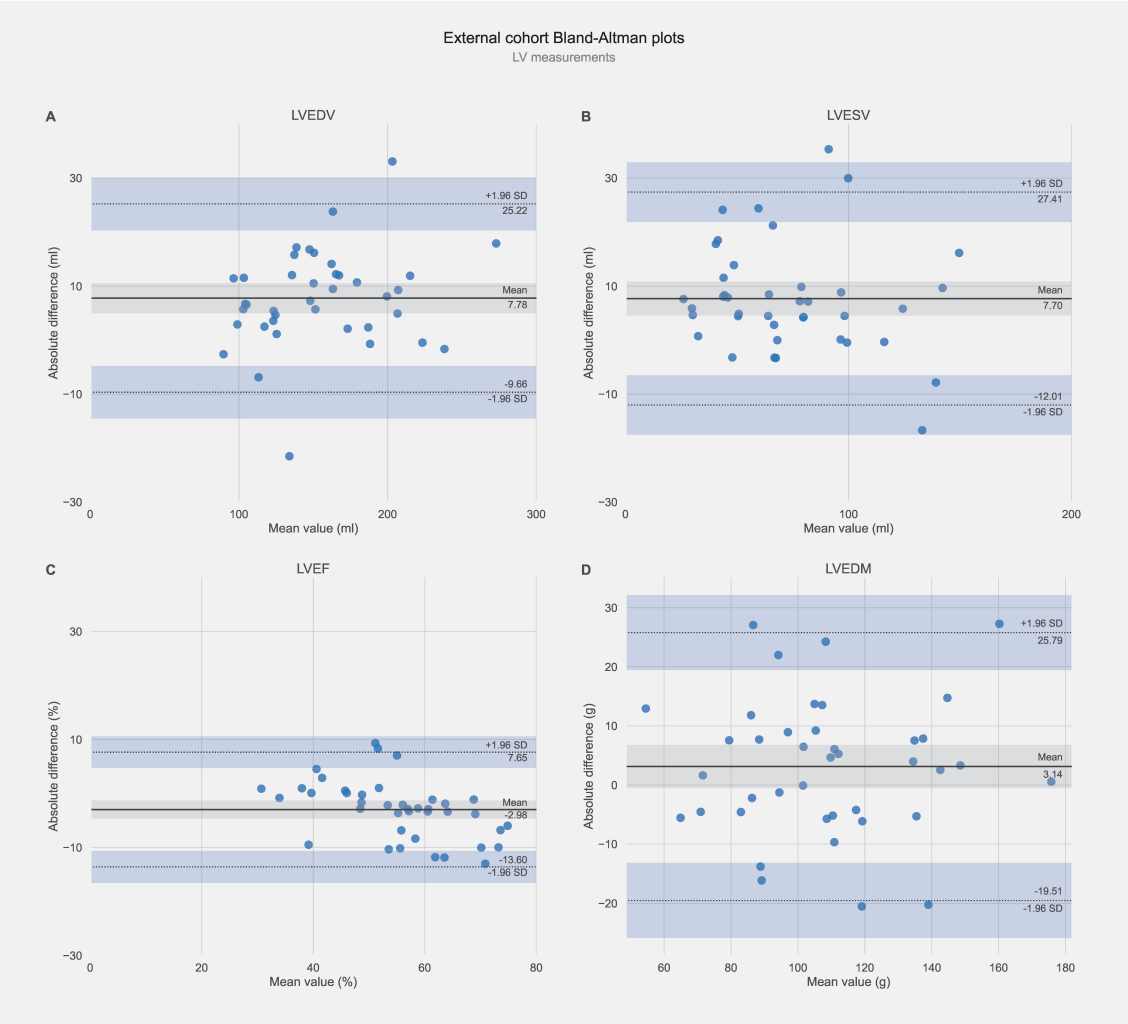


Figure 4.13: Bland–Altman plots showing the absolute mean difference between AI and manual left ventricular measurements in the external cohort ($n = 40$). LV, left ventricle; EDV, end-diastolic volume; ESV, end-systolic volume; EDM, end-diastolic mass.

4.3.7 Segmentation accuracy

Dice analysis showed excellent agreement in the AI and manual LV and RV epi– and endocardial end–systolic and end–diastolic contours in both the internal and external test cohorts (Table 4.9). The Dice values in the internal data sets ranged between 93% and 96% for the LV and 93%−95% in the RV. The Dice values were slightly lower in the external cohort and ranged from 89% to 95% in the LV and 88% to 92% in the RV.

Table 4.10: Comparison of automatic and manual cardiac MRI measurements in the external cohort ($n = 40$)

	AI	***Manual***	***P Value***	***ICC***	
RVEDV (ml)	216 ± 88	217 ± 87	.95	0.99	(0.98, 0.99)
RVESV (ml)	127 ± 69	135 ± 69	.64	0.99	(0.97, 0.99)
RVSV (ml)	89 ± 38	83 ± 40	.50	0.94	(0.88, 0.97)
RVEF (%)	44 ± 13	40 ± 13	.23	0.94	(0.81, 0.97)
RVEDM (g)	46 ± 18	51 ± 21	.28	0.92	(0.82, 0.96)
LVEDV (ml)	163 ± 53	155 ± 51	.50	0.99	(0.92, 1.00)
LVESV (ml)	77 ± 33	69 ± 35	.31	0.96	(0.85, 0.99)
LVSV (ml)	86 ± 29	86 ± 29	.99	0.97	(0.95, 0.98)
LVEF (%)	54 ± 10	57 ± 13	.25	0.93	(0.82, 0.97)
LVEDM (g)	110 ± 27	107 ± 27	.61	0.95	(0.91, 0.97)

4.4 Discussion

Our study developed and comprehensively analysed the performance of a fully automated biventricular cardiac MRI assessment in a large cohort of patients. We demonstrated that fully automated left ventricular (LV) stroke volume and ventricular mass index assessment had a correlation that was stronger than manual assessment with invasive haemodynamics parameters such as LV stroke volume ($r = 0.74$ vs 0.68; $P = .03$), pulmonary vascular resistance ($r = 0.62$ vs 0.41; $P < .001$), and mean pulmonary artery pressure ($r = 0.56$ vs 0.37; $P < .001$). Additionally, we showed excellent scan–rescan repeatability of artificial intelligence (AI) measurements for assessing LV and right ventricular (RV) measurements, including the more challenging RV mass (interclass correlation coefficient, 0.98; 95% CI: 0.96, 0.99). At a population level, we evaluated the prognostic value of AI–based cardiac MRI measurements and showed its ability to predict mortality in a cohort with multiple pathologic diseases, and further evaluated RV parameters in a subgroup of patients with

pulmonary arterial hypertension. Finally, we have shown excellent generalisability of AI contours in an external cohort.

We included varying types and severities of conditions affecting the RV to improve the reliability of our model for automatically measuring RV function and volume. We also trained our model to recognize the RV epicardial contours to capture a variety of RV appearances, such as RV dilatation and hypertrophy, in addition to normal variations. Previous studies that assessed biventricular or focused RV short–axis segmentation used small public data sets and included no or only a limited number of patients with abnormalities of RV function [159]. The largest biventricular segmentation studies were reported by Bai et al. [113] and Budai et al. [203] and each study included approximately 5000 participants. Bai et al. [113] included healthy volunteers from the UK Biobank study, whereas Budai et al. [203] included a cohort with mainly LV pathologic disease and limited RV pathologic disease because of conditions such as arrhythmogenic ventricular disease. Both studies were single–centre and single vendor. Our study differs in two main aspects: the AI segmentation model included large training data sets from multiple vendors (GE, Philips, and Siemens), multiple centres (Budapest and Sheffield), and multiple pathologic causes (LV and RV conditions); and automated cardiac MRI results were assessed by testing their correlation with invasive haemodynamics, prognostic ability, repeatability, and comparison to manual measurements in an external cohort. Our external test data set included patients referred to a specialist for a second opinion for complex pathologic causes.

We validated AI–derived cardiac MRI measurements against invasive haemodynamics performed on the same day. Cardiac MRI has diagnostic accuracy for pulmonary hypertension when compared to reference standard haemodynamics [204–206]. The correlation between RV ejection fraction and pulmonary vascular resistance has been reported to range between -0.32 and -0.55, and the correlation with mean pulmonary artery pressure ranges between -0.28 and -0.66 [207–210]. Ventricular mass index (RV mass–to–LV mass ratio) also correlates with RHC parameters ranging between 0.11 and 0.74 for pulmonary vascular resistance and from 0.53 to 0.87 for mean pulmonary artery pressure [207, 209, 210]. Our

study showed that AI–based cardiac MRI measurements correlate with RHC parameters. Particularly ventricular mass index, which is a known diagnostic and prognostic marker in pulmonary arterial hypertension, showed stronger correlation with RHC when measured automatically, indicating improved accuracy over manually measured ventricular mass index. Although some values showed a high level of disagreement, this is expected in a heterogeneous population including patients with congenital heart disease and considering the significant technical variability between the modalities compared.

The prognostic value of cardiac MRI measurements has been established in several cardiopulmonary diseases, including ischaemic heart disease, cardiomyopathies, heart failure, and pulmonary arterial hypertension [46, 211–213]. In patients with pulmonary arterial hypertension, RV ejection fraction, RV end–systolic volume index, and RV end–diastolic volume index were shown to predict mortality and clinical worsening in a meta–analysis of almost 2000 patients [4]. Our study confirmed the prognostic ability of automatic cardiac MRI measurements in a large cohort of 3417 patients with multiple pathologic diseases, including 920 patients with pulmonary arterial hypertension. RV ejection fraction, RV end–systolic volume index, RV end–diastolic volume index and RV end–diastolic mass index predicted death in pulmonary arterial hypertension when corrected for age and sex. RV end–diastolic mass index is not assessed in commercial software packages but can provide useful prognostic information, particularly in pulmonary hypertension. Additionally, we showed that automatically measured RV ejection fraction is a statistically significant prognostic marker in pulmonary arterial hypertension when added to functional assessment (World Health Organization functional class and walking test) and right heart catheterisation parameters.

Although our analysis showed that differences between the automatic and manual measurements were not statistically significant, these differences can be relatively large (e.g., 6% difference in ejection fraction). Therefore, we believe that establishing normal ranges of AI segmentation is important. Additionally, despite similar repeatability between the automatic and manual segmentation, consistent differences were noted in the manual segmentation of the scan–rescan cohort, such as excluding portions of the right ventricular

outflow tract. Whereas this consistency maintained excellent repeatability, the manual segmentation was less accurate. The scan–rescan segmentation was performed by a cardiac MRI practitioner not involved in the AI model training, highlighting the existence of subjective differences in the interpretation of the base of the heart even within the same institution. Furthermore, the AI segmentation fails in some patients, showing the need for further training. Failed AI segmentation will be continuously identified and incorporated in future training rounds to improve the accuracy of the model.

4.4.1 Limitations

Our study had limitations. First, the validation, including comparison with heart catheterisation and prediction of mortality, was performed in a single centre with two MRI systems and limited cohort description. Second, direct comparison between AI and manual measurements in the large ASPIRE cohort for RHC correlation and mortality prediction could not be performed because of differences in handling trabeculations. Excluding trabeculations from the blood pool and counting them towards myocardial mass might improve prognostic assessment. Right ventricular mass was not a statistically significant predictor of mortality in the multivariable Cox regression when trabeculations were included in the blood pool, while it remained prognostic when trabeculations were excluded from the blood pool. Experimental versions of the AI tool are able to identify trabeculations, however require further development and validation. Third, the segmentation algorithm cannot be made publicly available because the deep learning code would require extensive documentation and compatibility scripts to enable the application by external parties. However, we encourage readers to contact the corresponding author for research access to the Mass software and the AI segmentation tool.

4.5 Conclusion

In conclusion, we described a human–in–the–loop artificial intelligence (AI) approach to develop a biventricular cardiac MRI assessment tool. We provided a comprehensive evaluation of AI–based cardiac MRI measurements in a large cohort of patients with a wide spectrum of right and left ventricular pathologic abnormalities and normal variants. Fully automatic cardiac MRI assessment correlates with invasive haemodynamics and has

prognostic value. Training to target apex errors and more extreme pathologic abnormalities could advance the AI method further. Future research that uses cardiac MRI as a clinical end point can benefit from the high repeatability and generalisability of AI measurements.

4.6 Author Contributions

S.A. received the RSNA Cardiac Research Prize for this study. Guarantors of integrity of entire study, S.A., A.J.S.; study concepts and study design, S.A., R.J.v.d.G., A.J.S.; data acquisition, S.A., A.J.S., N.H.; data analysis/interpretation, S.A., A.J.S., N.H.; cardiac MRI acquisition, D.C.; cardiac MRI labelling, S.A., K.K., A.T., A.J.S.; F.A.; model development and training; R.J.v.d.G.; literature research, S.A.; manuscript drafting, S.A.; manuscript revision for important intellectual content, S.A., A.J.S., H.L., D.G.K., R.J.v.d.G.; statistical analysis, S.A., R.J.v.d.G.; manuscript editing, S.A., K.D., H.L., D.G.K., R.J.v.d.G., A.J.S.; and approval of final version of the submitted manuscript, all authors.

4.7 ASPIRE Cardiac MRI studies

A list of all studies that include part of the ASPIRE cardiac MRI cohort.

1. Alabed S, Alandejani F, Dwivedi K, et al. Validation of Artificial Intelligence Cardiac MRI Measurements: Relationship to Heart Catheterization and Mortality Prediction. Radiology 2022
2. Alabed S, Uthoff J, Zhou S, et al. Machine learning cardiac-MRI features predict mortality in newly diagnosed pulmonary arterial hypertension. European Heart Journal - Digital Health, 2022
3. Alandejani F, Alabed S, Garg P, et al. Training and clinical testing of artificial intelligence derived right atrial cardiovascular magnetic resonance measurements. Journal of Cardiovascular Magnetic Resonoance 2022; 24(1) : 25.
4. Alandejani F, Hameed A, Tubman E, Alabed S, et al. Imaging and Risk Stratification in Pulmonary Arterial Hypertension: Time to Include Right Ventricular Assessment. Frontiers of Cardiovascular Medicine 2022; 9 : 797561.3.
5. Shahin Y, Alabed S, Rehan Quadery S, et al. CMR Measures of Left Atrial Volume Index and Right Ventricular Function Have Prognostic Value in Chronic Thromboembolic Pulmonary Hypertension. Front Med (Lausanne) 2022; 9 : 840196.
6. Swift AJ, Wilson F, Cogliano M, Alabed S, et al. Repeatability and sensitivity to change of non-invasive end points in PAH: the RESPIRE study. Thorax 2021; 76(10) : 1032–1035.
7. Lewis RA, Johns CS, Cogliano M, et al. Identification of Cardiac Magnetic Resonance Imaging Thresholds for Risk Stratification in Pulmonary Arterial Hypertension. American Journal of Respiratory Critical Care Medicine 2020; 201(4) : 458–468.
8. Saunders LC, Johns CS, Stewart NJ, et al. Diagnostic and prognostic significance of cardiovascular magnetic resonance native myocardial T1 mapping in patients with pulmonary hypertension. Journal of Cardiovascular Magnetic Resonoance 2018; 20(1):78.
9. Johns CS, Kiely DG, Rajaram S, et al. Diagnosis of Pulmonary Hypertension with Cardiac MRI: Derivation and Validation of Regression Models. Radiology 2019; 290(1) : 61–68.

10. Johns CS, Wild JM, Rajaram S, et al. Identifying At-Risk Patients with Combined Pre- and Postcapillary Pulmonary Hypertension Using Interventricular Septal Angle at Cardiac MRI. Radiology 2018; 289(1) : 61–68.
11. Johns CS, Rajaram S, Capener DA, et al. Non-invasive methods for estimating mPAP in COPD using cardiovascular magnetic resonance imaging. European Radiology 2018; 28(4) : 1438–1448.
12. Johns CS, Swift AJ, Rajaram S, et al. Lung perfusion: MRI vs. SPECT for screening in suspected chronic thromboembolic pulmonary hypertension. Journal of Magnetic Resonance Imaging 2017; 46(6) : 1693–1697.
13. Swift AJ, Capener D, Johns C, et al. Magnetic Resonance Imaging in the Prognostic Evaluation of Patients with Pulmonary Arterial Hypertension. American Journal of Respiratory Critical Care Medicine 2017; 196(2) : 228–239.
14. Hussain N, Charalampopoulos A, Ramjug S, et al. Pulmonary hypertension in patients with heart failure and preserved ejection fraction: differential diagnosis and management. Pulmonary Circulation 2016; 6(1) : 3–14.
15. Swift AJ, Rajaram S, Capener D, et al. Longitudinal and transverse right ventricular function in pulmonary hypertension: cardiovascular magnetic resonance imaging study from the ASPIRE registry. Pulmonary Circulation 2015; 5(3) : 557–564. Swift AJ, Capener D, Hammerton C, et al. Right ventricular sex differences in patients with idiopathic pulmonary arterial hypertension characterised by magnetic resonance imaging: pairmatched case controlled study. PLoS One 2015;10(5):e0127415.
16. Swift AJ, Wild JM, Nagle SK, et al. Quantitative magnetic resonance imaging of pulmonary hypertension: a practical approach to the current state of the art. Journal of Thoracic Imaging 2014; 29(2) : 68–79.
17. Hurdman J, Condliffe R, Elliot CA, et al. ASPIRE registry: assessing the Spectrum of Pulmonary hypertension Identified at a REferral centre. European Respiratory Journal 2012;39(4):945–955.

CHAPTER 5

Establishing minimally important differences for cardiac MRI endpoints in pulmonary arterial hypertension

Samer Alabed[1,2], Pankaj Garg[3], Faisal Alandejani[1], Krit Dwivedi[1,2], Ahmed Maiter [1,2], Kavita Karunasaagarar[1,2], Smitha Rajarama[1,2], Catherine Hill[2], Steven Thomas[1,2], Rebecca Gossling[1], Michael Sharkey[1,2], Haiping Lu[4] ,Robert A. Lewis[1], Alexander Rothman[1], Lisa Watson[1,5], Abdul Hameed[1,7], Athaniosis Charalampopoulos[1,5], Robin Condliffe[1,5], James M. Wild[1], Rob van der Geest[6], Andrew J. Swift[1,2], David G. Kiely[1,5]

[1] Infection, Immunity & Cardiovascular Disease, University of Sheffield

[2] Department of Clinical Radiology, Sheffield Teaching Hospitals

[3] Norwich Medical School, University of East Anglia

[4] Department of Computer Science, University of Sheffield

[5] Sheffield Pulmonary Vascular Disease Unit, Sheffield Teaching Hospitals

[6] Leiden University Medical Center, Leiden, the Netherlands

The abstract of this chapter was published in *Thorax*. 2022 November;77(suppl) and *The International Journal of Cardiovascular Imaging* 2022 November;264(suppl)

Abstract

Objectives To identify minimally important differences (MIDs) for cardiac MRI (CMR) metrics based on Food and Drug Administration (FDA) recommendations for a surrogate endpoint that should reflect how a patient feels, functions and survives.

Background CMR is the gold standard technique to assess bi-ventricular volumes and function and is increasingly being considered as an endpoint in clinical studies. Currently, with the exception of right ventricle (RV) stroke volume, there are no MIDs reported for CMR metrics.

Methods Consecutive treatment-naive patients with PAH between 2010 and 2021 who had two CMR scans (at baseline prior to treatment and at 12 months following treatment) were identified from the ASPIRE registry. All patients were followed up for one additional year after the second scan. For both scans, cardiac measurements were obtained from a validated fully automated segmentation tool. The MID in CMR metrics was determined using two methods; (i) an anchor-based method combining how a patient "feels" (emPHasis-10 questionnaire) and "functions" (incremental shuttle walking test) and (ii) for "survives" a distribution-based method for one-year mortality. RV ejection fraction (RVEF) and RV and left ventricle (LV) end-diastolic volume, RV end-systolic volume and LV stroke volume were measured at baseline and follow-up. For each metric, the absolute difference (change from baseline to follow-up) and relative difference (ratio of absolute difference to baseline measurement) were compared in a Cox proportional hazard regression and Kaplan-Meier survival analysis.

Results 239 patients were included. The MIDs ($P < 0.05$), for metrics for how a patient "feels and functions" for (i) improvement, were a relative increase in RVEF, LVSV or LVEDV of 5% and a decrease in RVESV or RVEDV of 7% and 5%, respectively and for (ii) clinical worsening, were a relative reduction in RVEF of 10%, reduction of LVSV or LVEDV of 5% or an increase in RVESV of 14%. For "survives" the MID associated with (i) a reduced one-year mortality

was an 8% relative reduction in RVESV and a 12% relative increase in LVEDV and (ii) increased one-year mortality were a 10% relative decrease in RVEF, a 22% increase in RVESV, a 6% increase in RVEDV, a 20% reduction in LVEDV and a 12% reduction in LVEDV.

Conclusion This study establishes clinically relevant CMR MIDs for how a patient feels, functions and survives in response to PAH treatment. These findings provide further support for the use of CMR as a clinically relevant surrogate endpoint and will aid trial-size calculations for studies using CMR.

5.1 Introduction

In patients with pulmonary arterial hypertension (PAH) symptoms and survival are determined primarily by right ventricular function. PAH, a progressive pulmonary vasculopathy results in elevation of mean pulmonary arterial pressure (mPAP) and an increase in right ventricle (RV) afterload [30]. With disease progression and chronically elevated mPAP, the RV undergoes remodelling, resulting in either adaptation and maintenance of output [214] or maladaptation, RV failure and consequently reduced survival [215]. Cardiac MRI (CMR) is the gold standard for assessing the RV and shows potential in the assessment of PAH [12]. Impairments of RV function and associated increases in RV volumes can be detected and quantified by CMR, enabling prediction of clinical worsening and mortality [4] and aids risk stratification [59]. In addition, CMR is sensitive to improvements in RV function following PAH therapy [55, 216–218] and detects a larger treatment effect than the 6-minute-walk [192]. In this context, CMR is an important tool for risk stratification and monitoring of disease and treatment response in PAH [12].

Phase four clinical studies of PAH therapies have recently utilised CMR as a primary endpoint in addition to other composite outcomes [217, 218]. Assessing treatment response with CMR necessitates clinically relevant thresholds in order to determine improvement or worsening. However, only RV stroke volume measured on pulmonary artery phase-contrast flow imaging has established thresholds [219], while those for volumetric CMR measurements on cine imaging remain unvalidated. The introduction of automatic volumetric CMR measurements assessing RV changes over time has several advantages. It offers excellent repeatability in scan-rescan assessment and has higher accuracy than manual assessment [7]. In addition, results are generalisable across different centres and MRI systems, allowing standardised comparisons independent of the location of scan [7, 163, 220].

The Federal Drug Administration (FDA) has highlighted the need to identify surrogate endpoints for PAH therapy trials that reflect how a patient “feels, functions and survives” [150]. To reflect this, we have aimed to identify clinically relevant thresholds for change in automatically derived CMR RV and LV measurements, benchmarking against

patient-centred clinical parameters ("feels"), exercise testing ("functions") and mortality ("survives").

5.2 Materials and Methods

5.2.1 Study Sample

Adult PAH patients were identified from the "Assessing the Spectrum of Pulmonary hypertension Identified at a REferral centre" (ASPIRE) registry [191] between January 2010 and February 2021. Diagnosis of PAH was based on mPAP $\geq$ 25 mmHg and PAWP $\leq$ 15 mmHg and PVR $\geq$ 3 Wood Units, measured by right heart catheterisation (RHC). Patients were eligible for inclusion if they had: (1) baseline CMR prior to starting treatment and within 48 hours of PAH diagnosis, (2) follow-up CMR at 12-24 months and (3) at least one-year follow-up after the follow-up scan. Patients were excluded if they did not have complete short axis stack imaging for both baseline and repeat scans. The local ethics committee and institutional review board approved this study (ASPIRE, ref: c06/Q2308/8).

5.2.2 Imaging Procedures

MRI Protocol

CMR was performed with 1.5 Tesla MRI systems from GE (Signa HDx, General Electrics Healthcare). Short-axis cine images were acquired using a cardiac-gated multislice balanced steady-state free precession sequence (20 frames per cardiac cycle, section thickness 10mm, 0mm inter-section gap, field of view 480mm, acquisition matrix 256 $\times$ 200, flip angle 60°, $BW125KHz/pixel$, TR/TE3.7/1.6ms). A stack of images in the short-axis plane was acquired, fully covering both ventricles from base to apex. End-systole was considered to be the smallest cavity area. End-diastole was defined as the first cine phase of the R-wave triggered acquisition or largest volume. Patients were supine with a surface coil and with retrospective ECG gating.

Image Analysis

RV function was assessed visually on four-chamber and short-axis cine and short-axis stacks. The magnitude of change in RV function between baseline and follow-up scans

was graded as improved, stable or worsened. Visual assessment was performed by expert radiologists (C.H., S.T., A.J.S., S.R.) with 17, 16, 12, 11 years of experience who were blinded to automatic but not manual measurements. Manual CMR measurements were obtained from contouring the biventricular endo- and epi-cardial borders at end-diastole and end-systole on short-axis stack images and by excluding trabeculations from the blood pool. Manual contouring was performed by an MRI radiographer with 19 years of experience using Medis (Medis Medical Imaging Systems, Leiden, the Netherlands). An in-house deep learning cardiac MRI segmentation tool was used to obtain fully automatic measurements [7]. The segmentation tool was trained in a multi-centre, multi-vendor and multi-pathology dataset and was previously validated by assessing: (i) Accuracy against same-day invasive pulmonary haemodynamics and phase contrast flow imaging. (ii) Repeatability in a same-day scan-rescan cohort. (iii) Generalisability in an external testing cohort. (iv) Mortality prediction in a large cohort with multiple cardiac and lung pathologies. The automatic contours included trabeculations in the blood pool and were obtained using MASS software (MASS, research version 2020; Leiden University Medical Center).

Clinical Parameters

The emPHasis-10 questionnaire (E-10) to assess health-related quality of life reported by patients with PAH was completed at baseline and at the time of the follow-up scan from 2014 onwards. Each patient completed ten questions ranked on a scale of 0 to 5, with a lower score indicating a better quality of life [221]. The incremental shuttle walking test (ISWT) was performed as part of routine patient evaluation according to the standard method [222]. The REVEAL score was calculated from composite clinical parameters [223] and modified to include the incremental shuttle walk test instead of the 6 minute walking test [59, 224]. Mortality data were collected from the electronic records of the National Health Service (NHS) Personal Demographics Service. The NHS automatically updates the mortality records once a death is registered in the United Kingdom. All patients were followed up as part of the national service specification for patients with pulmonary hypertension for a minimum of 12 months.

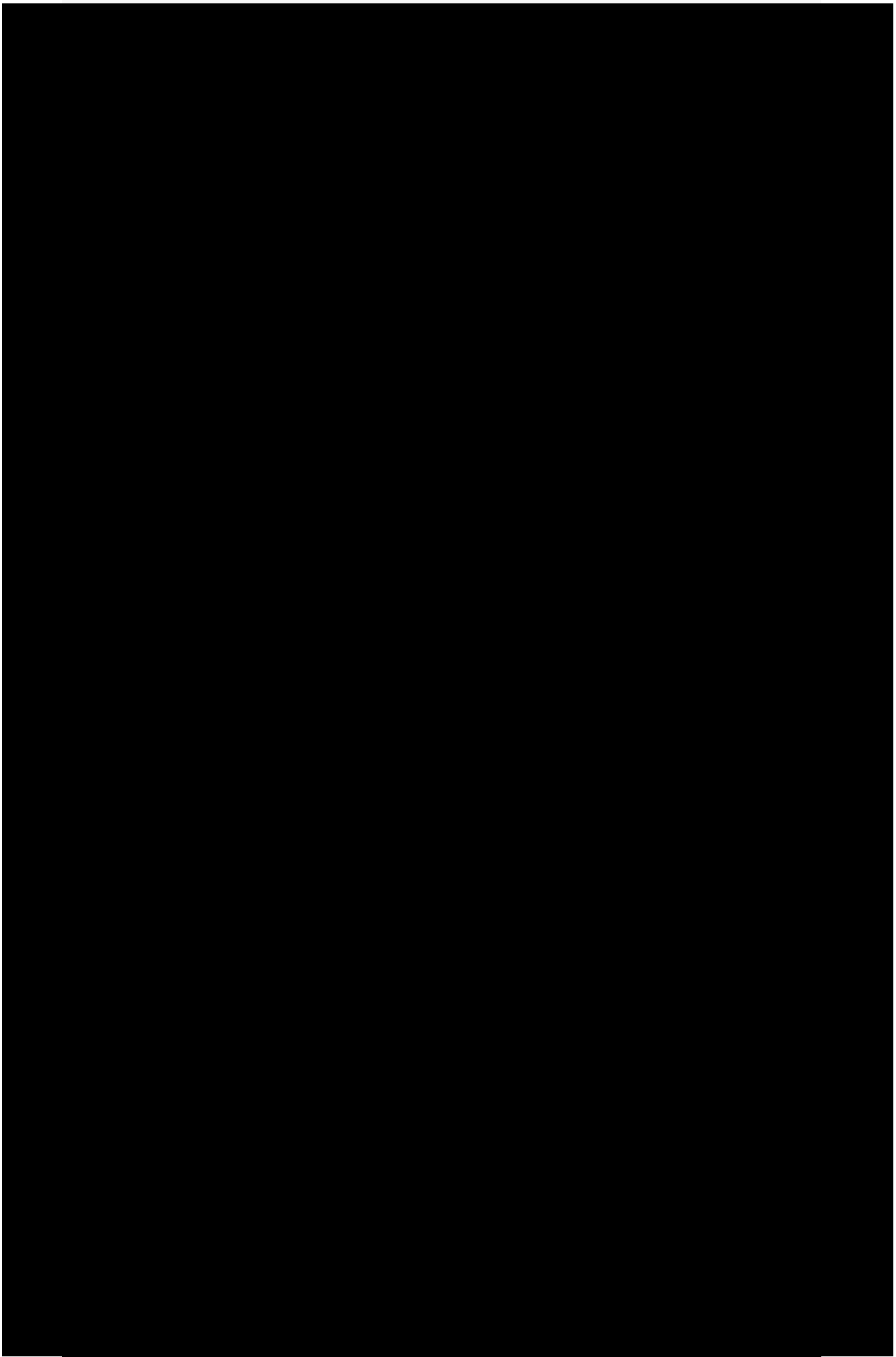
Statistical Analysis

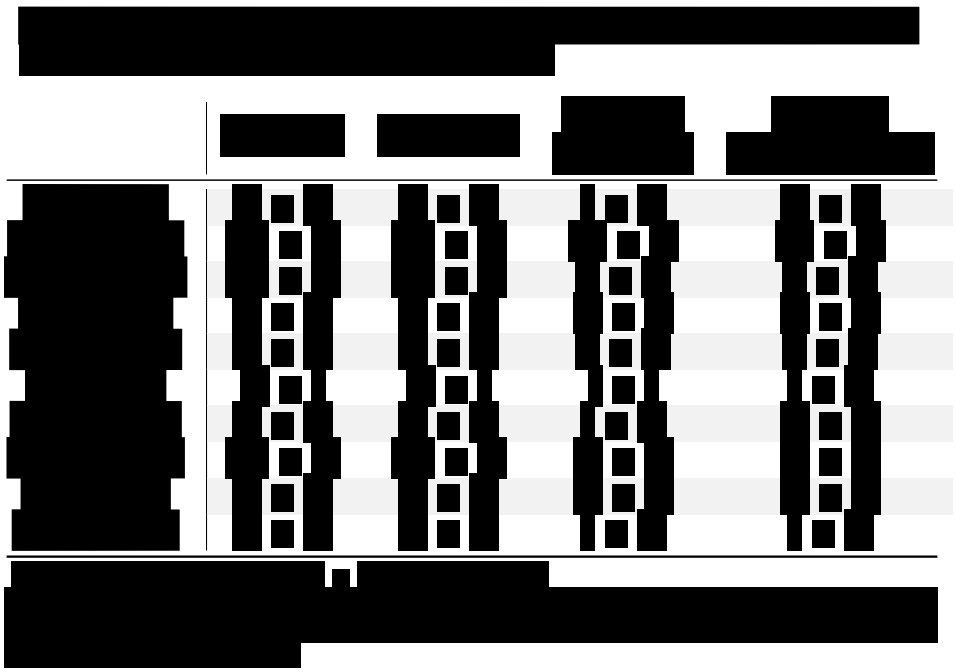
Baseline characteristics are presented as proportions, means standard deviations (SD), or medians and interquartile ranges (IQR). Two thresholds for minimal important difference (MID) were assessed as follows, with a significance threshold of 0.05. Firstly, MIDs for how a patient "feels and functions" were determined using a combination of a patient-reported outcome measure and assessment of exercise capacity. An anchor-based method was utilised using the E-10 in combination with the ISWT. Patients that had either E-10 or ISWT completed at the time of baseline and follow-up scans were considered to have improved, remained stable or worsened, based on a change of 6 points in E-10 [225, 226] or $47.5m$ in ISWT [227, 228]. If a patient had both tests and were discordant for improvement or worsening the patient was considered to have worsened. To calculate the MID, the mean changes in CMR measurements between baseline and follow-up in the improved and worsened groups were compared to the mean CMR changes in the overall sample used in the E-10 and ISWT assessment [229]. The Wilcoxon test was performed to test for differences between baseline and follow-up parameters. Statistical analyses were carried out using the lifelines and pingouin Python libraries [200, 201] and graphs were produced using the Matplotlib library [202] and Prism 9 (GraphPad Software, La Jolla CA, USA). Secondly, MIDs for how a "patient survives" were based on one-year mortality for CMR metrics encompassing those shown to be prognostic in PAH [4] and including right ventricular ejection fraction (RVEF), right ventricular end-systolic volume (RVESV), right ventricular end-diastolic volume (RVEDV), left ventricular stroke volume (LVSV) and left ventricular end-diastolic volume (LVEDV) were estimated through a distribution method using SD. Threshold values were defined at $0.25, 0.5$ and $0.75SD$ of the baseline measurements for absolute change and of the % change from baseline to follow-up for relative change [230, 231]. For each metric, the absolute (follow-up minus baseline measurement) and relative differences (ratio of the absolute difference to baseline measurement) were used to categorise patients as improved, stable or worsened. To assess the prediction of worsening, patients who improved or remained stable were compared against patients that worsened using Cox proportional hazard regression for one-year mortality. Similarly, to assess prediction of improvement, patients who worsened

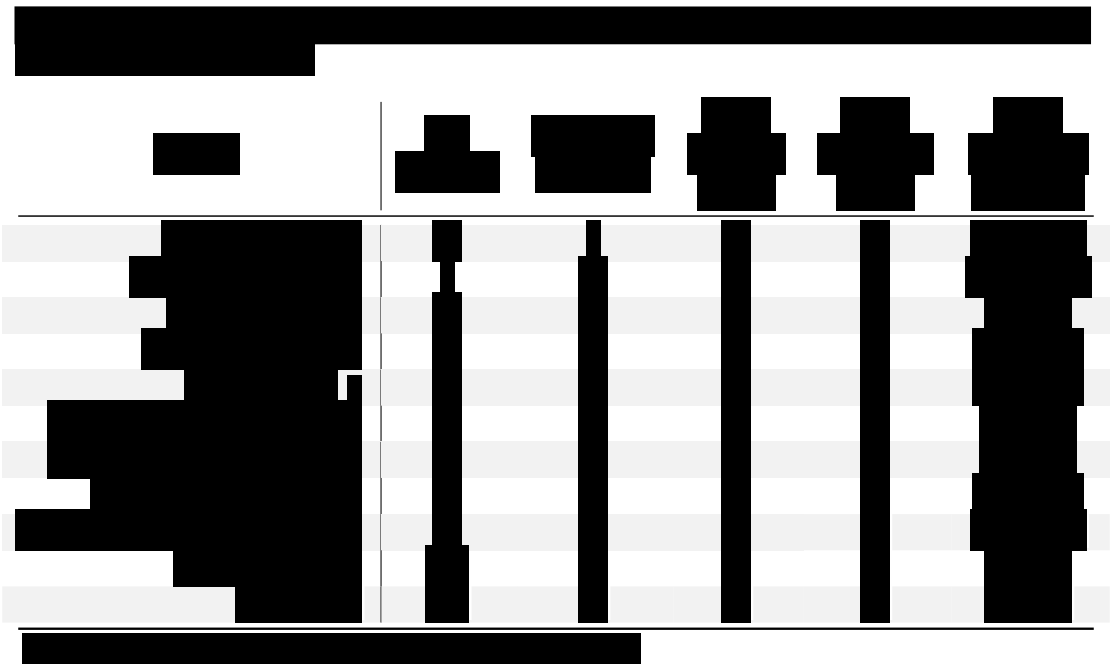
or remained stable were compared to patients that improved. Thirdly, the prognostic accuracy of predicting mortality compared to the visual assessment of the expert CMR reader was compared using (1) the receiver-operating characteristic curve (ROC) and the area under the curve (AUC) for one-year mortality and (2) Kaplan-Meier survival analysis for five years. DeLong's test was used to compare the AUC of RV change classification and visual assessment [232]. Between-group comparisons in the Kaplan-Meier analyses were made using the log-rank test. Agreement between the classification using RV thresholds and the visual assessment was tested using Cohen's Kappa statistic (k) (< 0.2 slight, $0.2 - 0.39$ fair, $0.4 - 0.59$ moderate, $0.6 - 0.8$ substantial and > 0.80 excellent agreement).

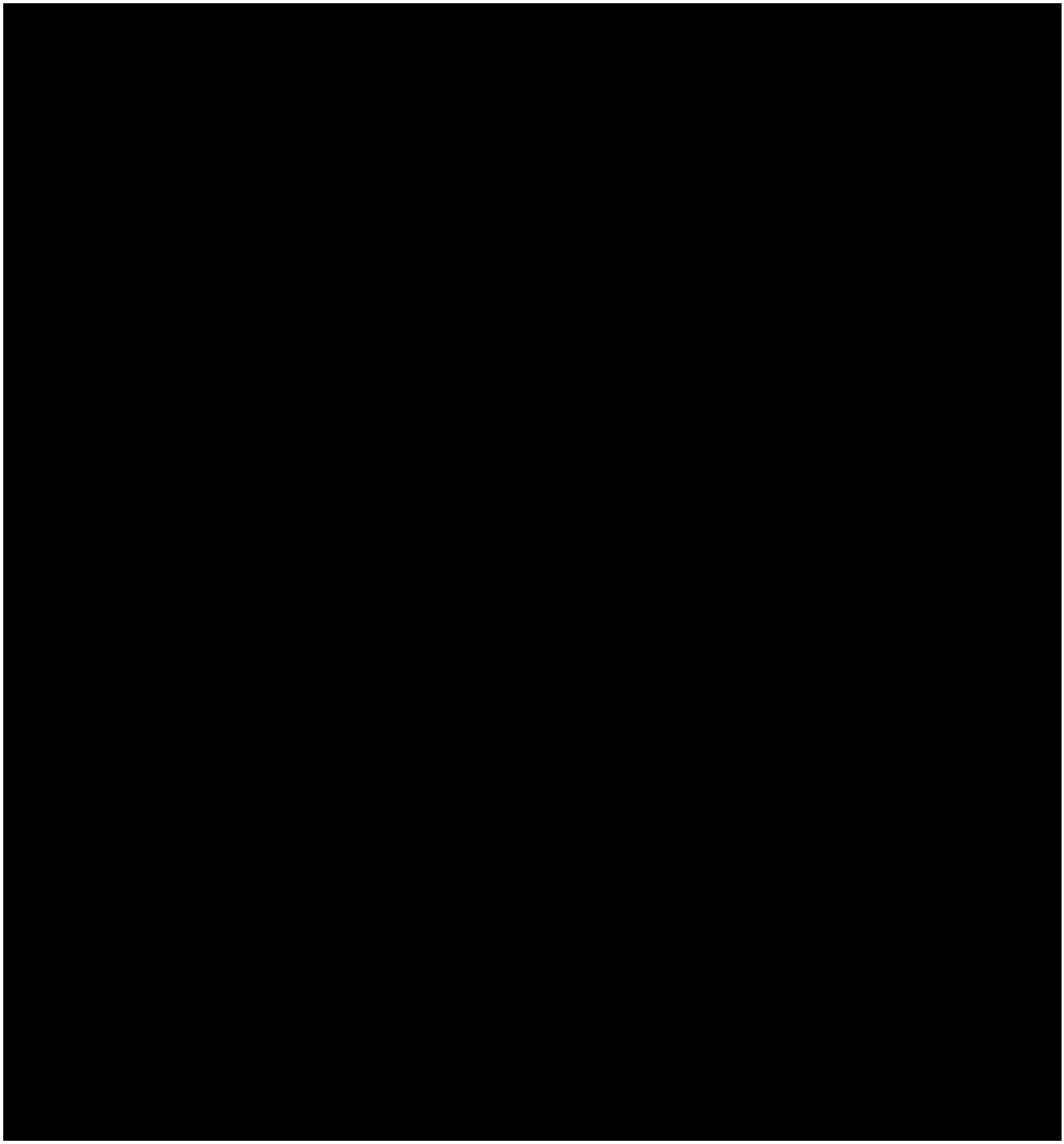
5.3 Results

5.3.1

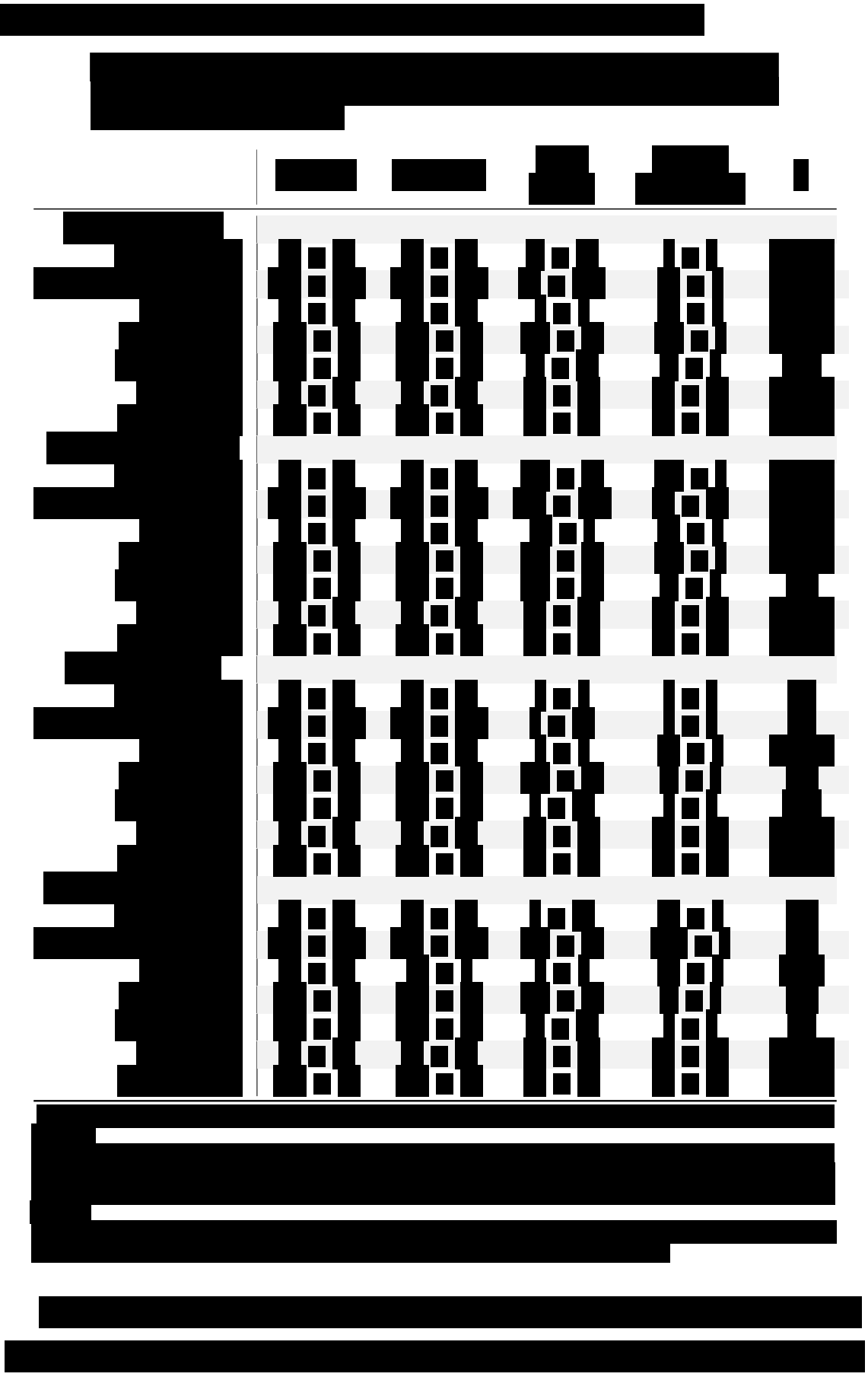




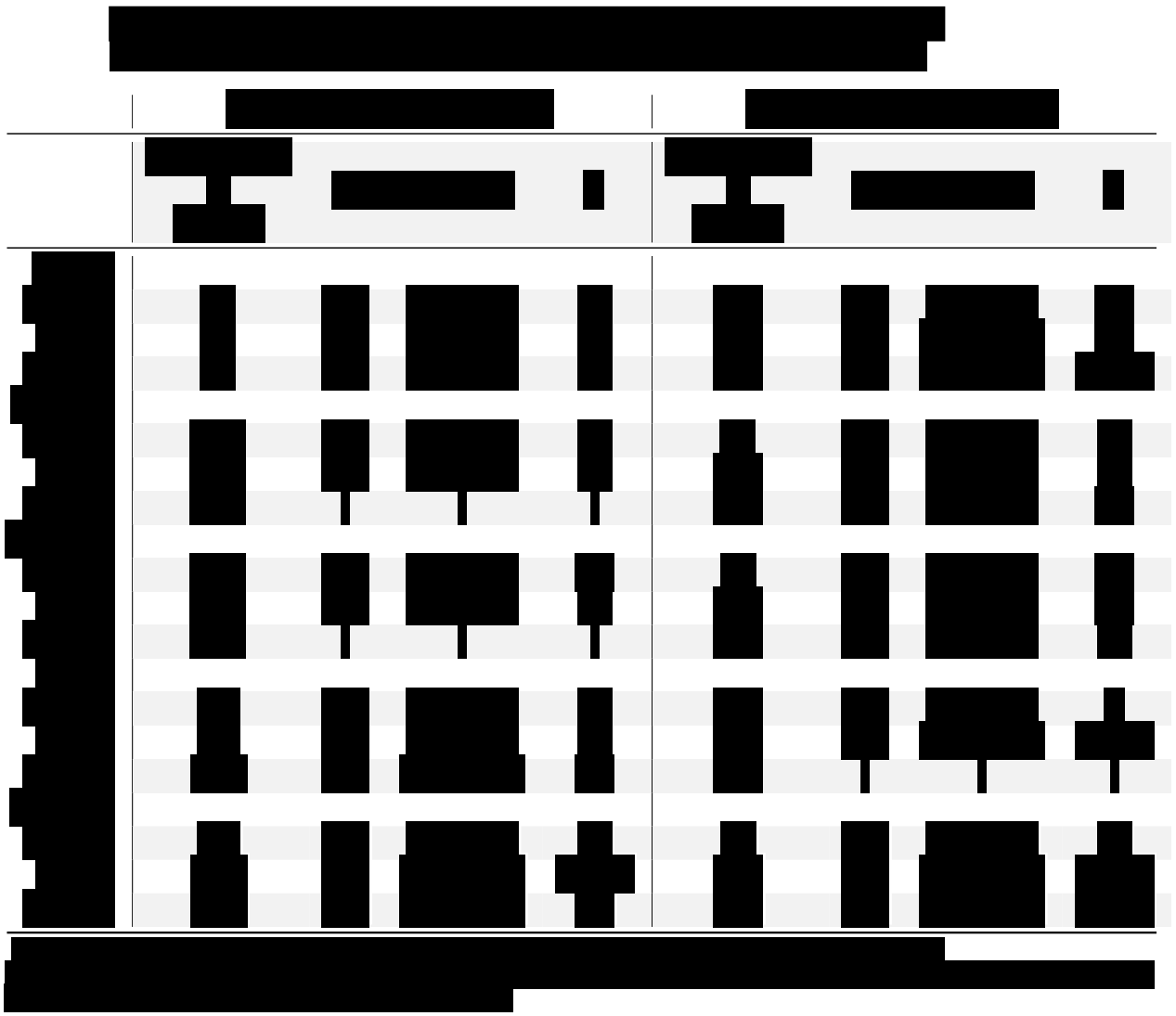


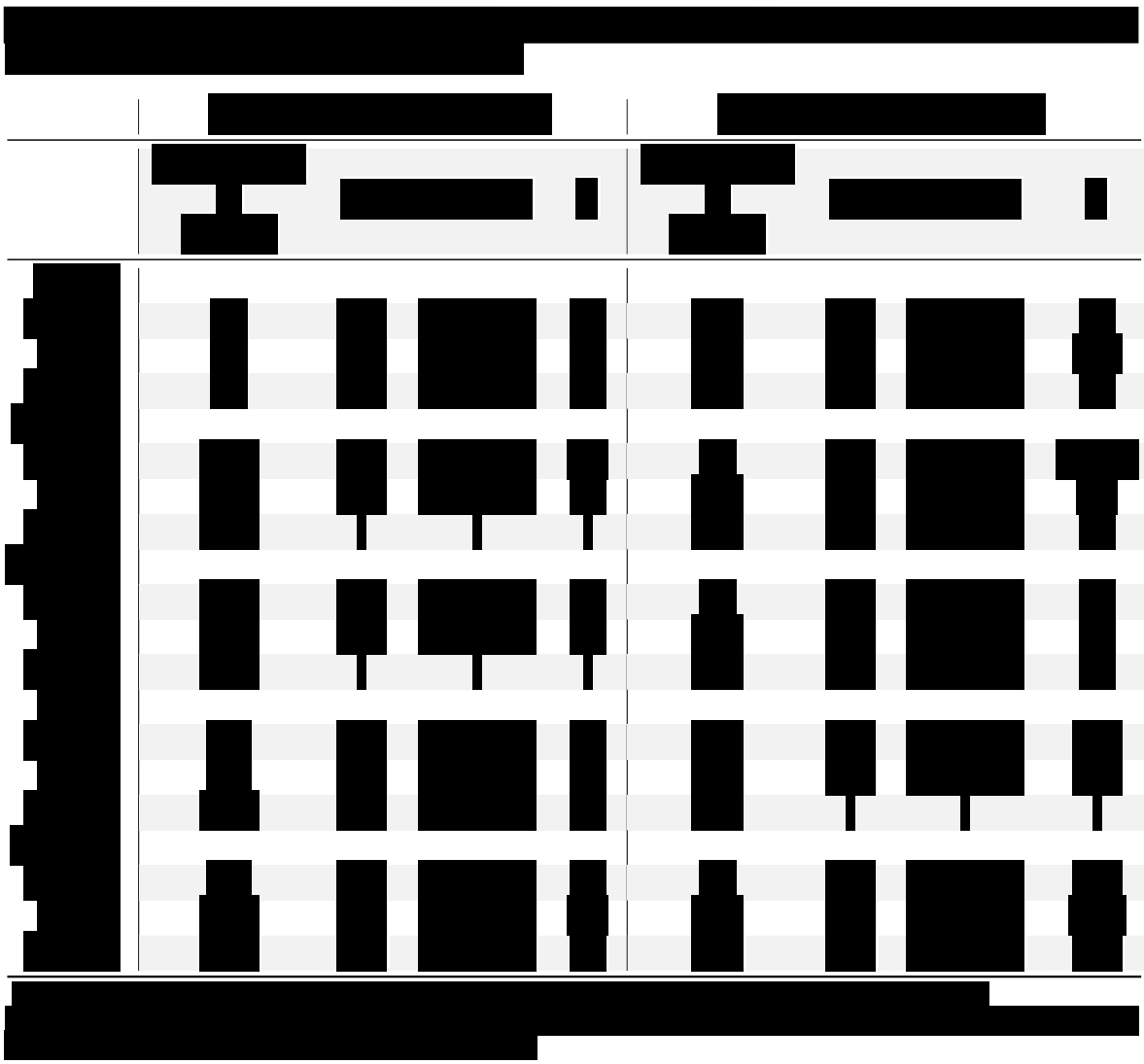


5.3.2 [redacted]



[redacted]

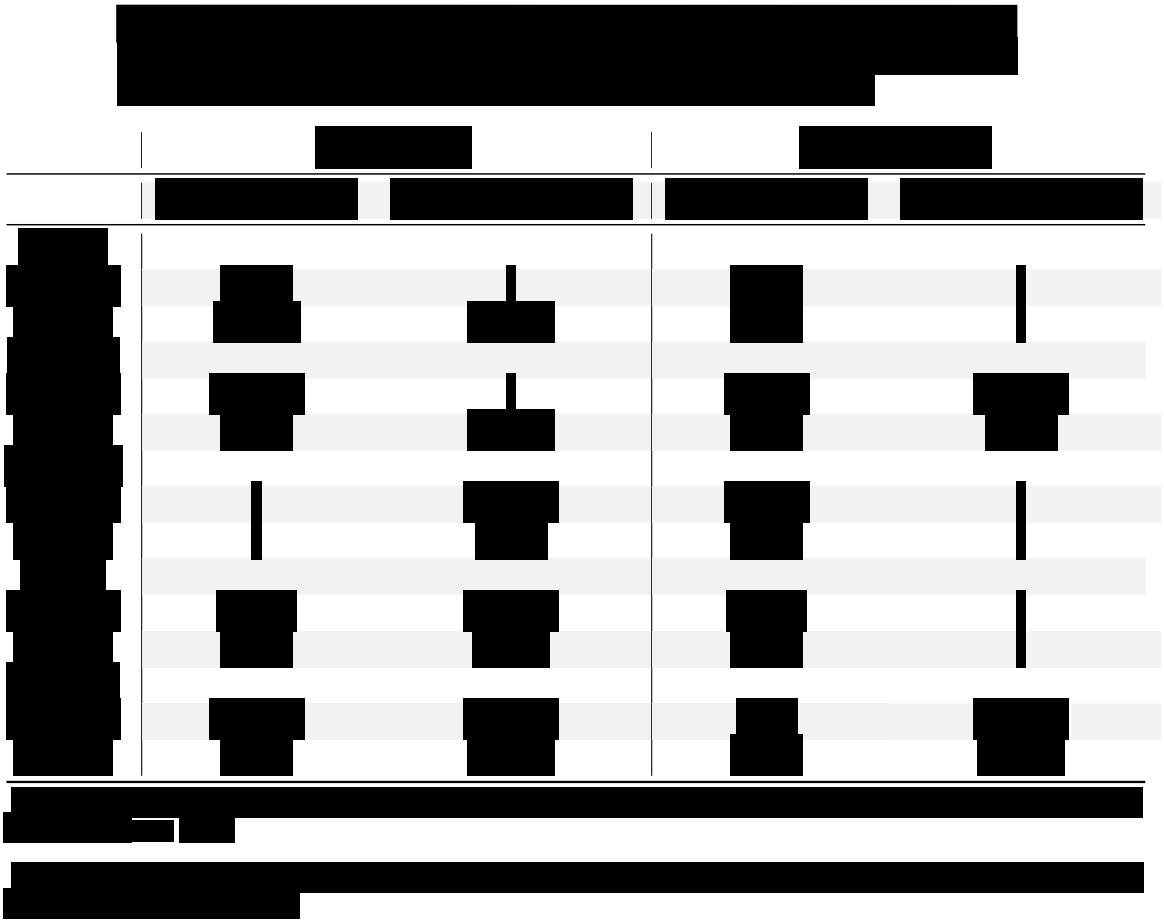




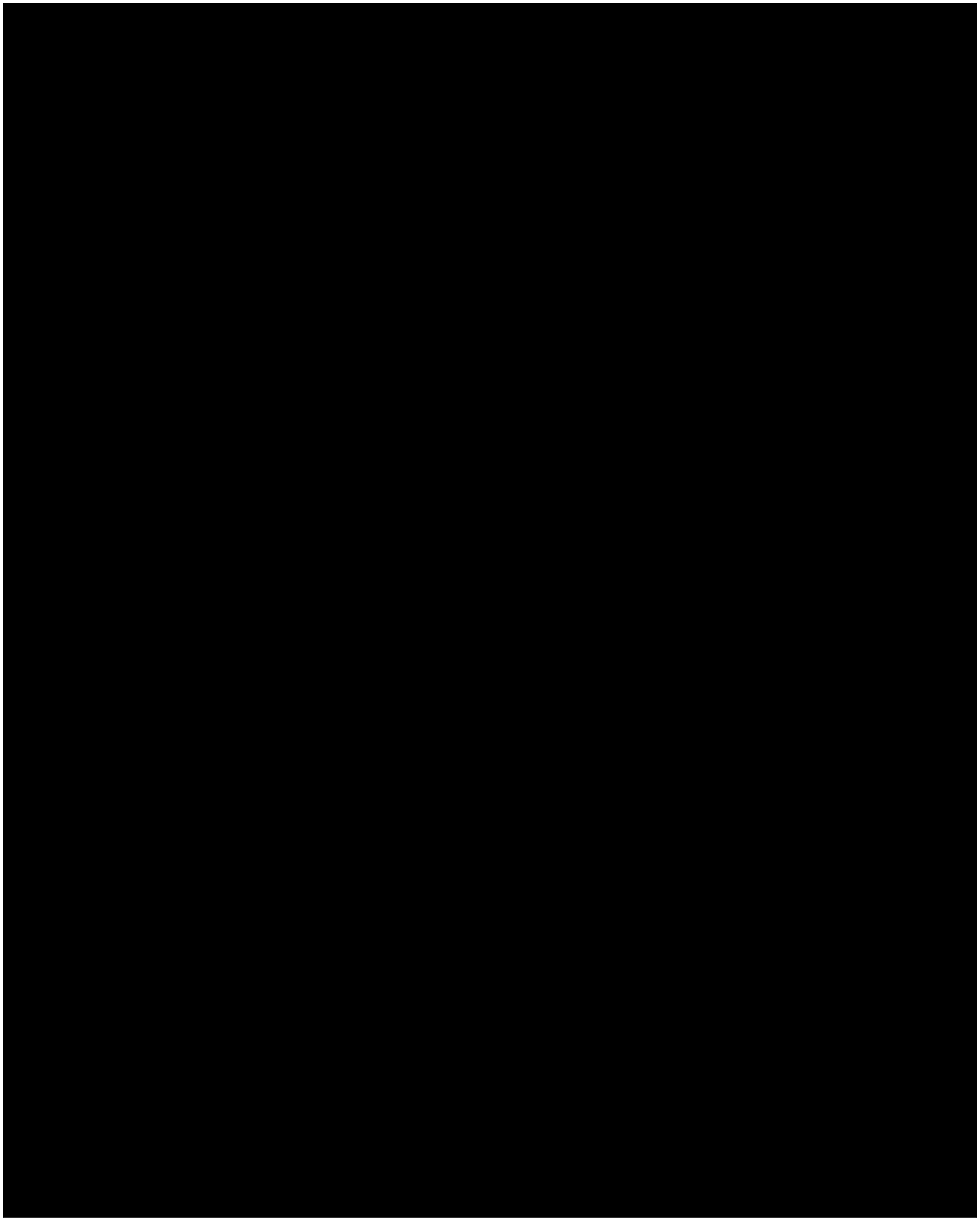
[redacted]

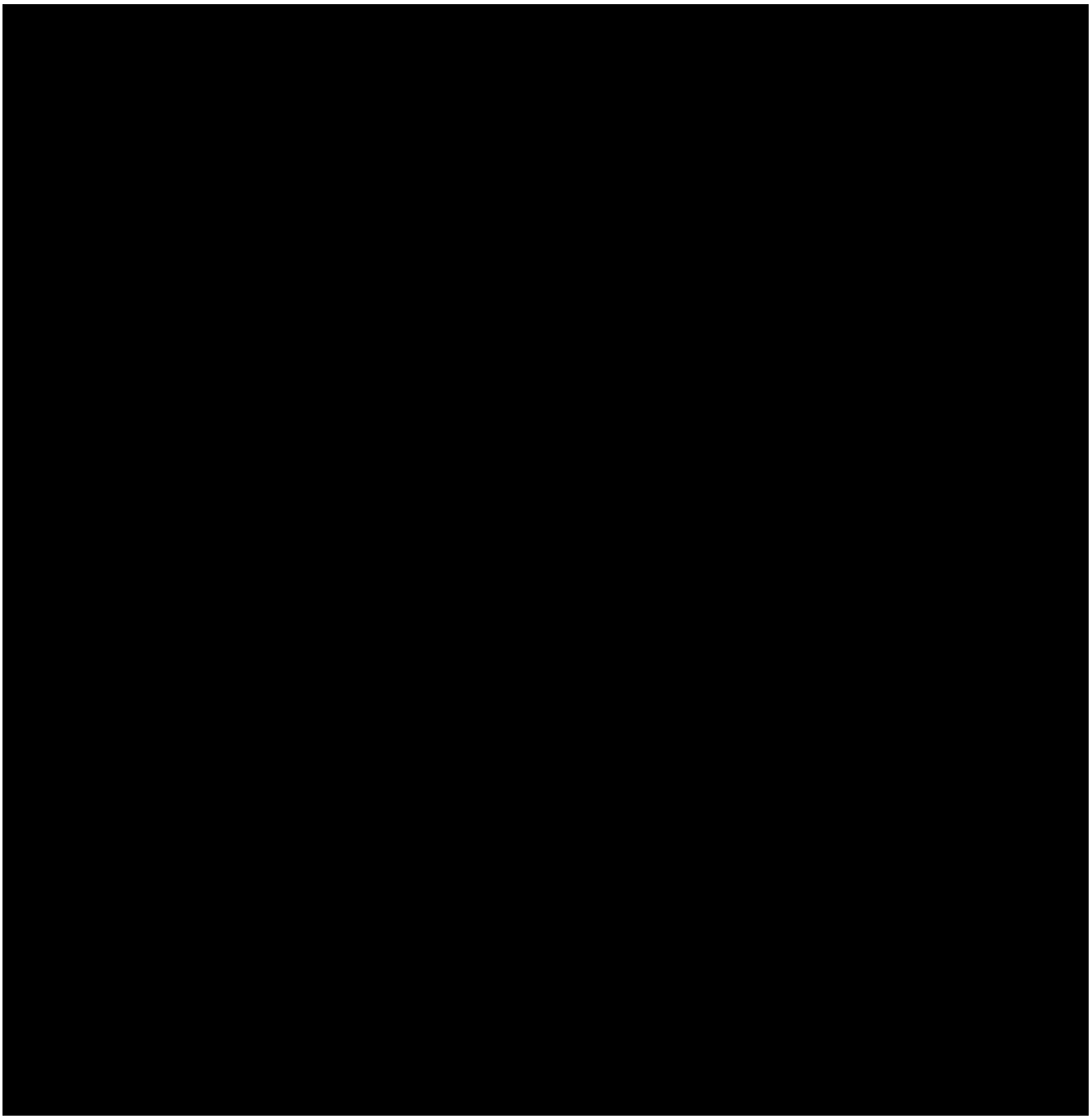
[redacted]

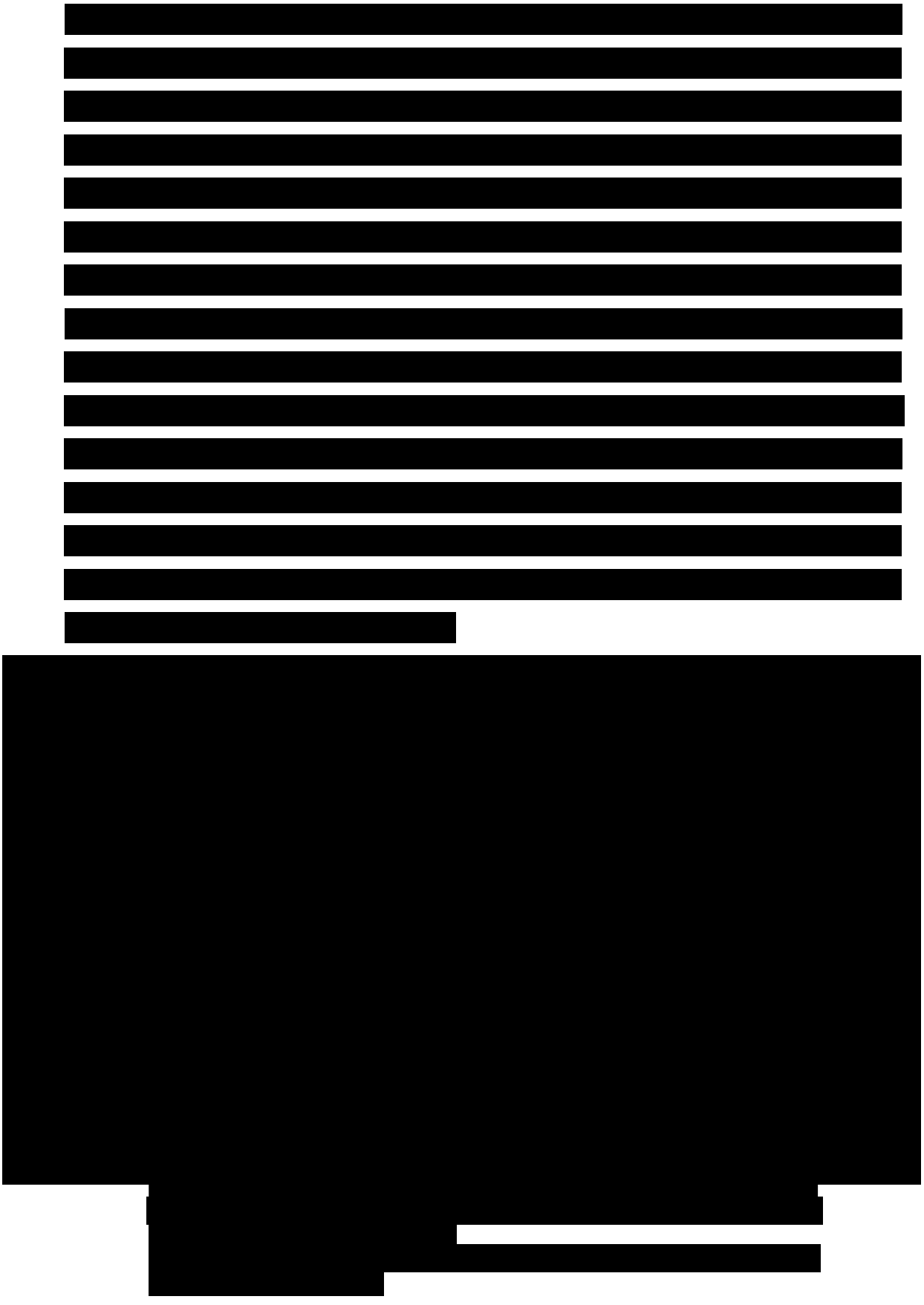
[redacted]



[redacted]







[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

CHAPTER 6

Machine learning cardiac-MRI features predict mortality in newly diagnosed pulmonary arterial hypertension

Samer Alabed[1,2], Johanna Uthoff[3], Shuo Zhou[3], Rebecca Gosling[1], Lawrence Schobs[3], Martin Brook[1], Pankaj Garg[4], Krit Dwivedi[1,2], Faisal Alandejani[1], Yousef Shahin[1,2], David Capener[1], Christopher Johns[1,2], James M. Wild[1], Alexander Rothman[1], Rob van der Geest[5], Robin Condliffe[1,6], David G. Kiely[1,6], Haiping Lu[3], Andrew J. Swift[1,2]

[1] Infection, Immunity & Cardiovascular Disease, University of Sheffield

[2] Department of Clinical Radiology, Sheffield Teaching Hospitals

[3] Department of Computer Science, University of Sheffield

[4] Norwich Medical School, University of East Anglia

[5] Leiden University Medical Center, Leiden, the Netherlands

[6] Sheffield Pulmonary Vascular Disease Unit, Sheffield Teaching Hospitals

This chapter is based on a publication in *European Heart Jouranl - Digital Health.* 2022 June;3(2):265-275.

Abstract

Objectives Pulmonary arterial hypertension (PAH) is a rare but serious disease associated with high mortality if left untreated. This study aims to assess the prognostic cardiac magnetic resonance (CMR) features in PAH using machine learning.

Methods Seven hundred and twenty-three consecutive treatment-naive PAH patients were identified from the ASPIRE registry; 516 were included in the training, and 207 in the validation cohort. A multilinear principal component analysis (MPCA)-based machine learning approach was used to extract mortality and survival features throughout the cardiac cycle. The features were overlaid on the original imaging using thresholding and clustering of high- and low-risk of mortality prediction values.

Results The 1-year mortality rate in the validation cohort was 10%. Univariable Cox regression analysis of the combined short-axis and four-chamber MPCA-based predictions was statistically significant (hazard ratios: 2.1, 95% CI: 1.3, 3.4, $c-index = 0.70$, $P = 0.002$). The MPCA features improved the 1-year mortality prediction of REVEAL from $c-index = 0.71$ to 0.76 ($P < 0.001$). Abnormalities in the end-systolic interventricular septum and end-diastolic left ventricle indicated the highest risk of mortality.

Conclusion The MPCA-based machine learning is an explainable time-resolved approach that allows visualisation of prognostic cardiac features throughout the cardiac cycle at population level, making this approach transparent and clinically interpretable. In addition, the added prognostic value over the REVEAL risk score and CMR volumetric measurements allows for a more accurate prediction of 1-year mortality risk in PAH.

6.1 Introduction

Cardiac MRI (CMR) is a powerful prognostic tool owing to its ability to assess cardiophysiological parameters such as the volume and function of the cardiac chambers, tissue characterisation and anatomical structure. Machine learning methods harnessing CMR's prognostic abilities remain rare and mainly focus on segmenting cardiac chambers to automate CMR measurements [240]. The process of automating CMR measurements has matured over recent years, proving to be accurate and comparable with results obtained from manual segmentation [159, 190, 241]. However, there is a wealth of data available in CMR studies other than those based on volumetric measurements. A recent machine learning model based on the motion of segmented right ventricle predicted mortality in a mixed cohort of pulmonary hypertension patients [116]. This study linked impaired basal longitudinal shortening and transverse contraction at the interventricular septum and free wall with an increased risk of mortality [116]. Another recent machine learning model based on CMR disease features extracted by multilinear principal component analysis (MPCA) has been used to predict the presence or absence of pulmonary arterial hypertension (PAH) [9] without the need for segmentation.

The MPCA-based model is interpretable because MPCA is a linear and transparent feature extraction method, thus a particularly promising machine learning approach for CMR imaging. Each CMR image sequence is a three-dimensional array (e.g. $512 \times 512 pixels$ $\times 6mm$ slice thickness $\times 20$ images throughout the cardiac cycle), with each element being a voxel capturing different tissue characteristics, anatomical location, and temporal variation in the cardiac cycle. Such a multidimensional array can be naturally represented as a mathematical object called a tensor. MPCA extracts features directly from such multidimensional tensor representation which preserves the multidimensional structure of the original CMR data more accurately than reshaping it into one-dimensional, vector representation [242]. The extracted MPCA features can then be weighted in classification or regression models to optimise the prediction of the desired outcome. MPCA is robust and has been successfully applied to brain imaging and more recently to CMR [9, 242].

Pulmonary arterial hypertension is a rare but serious disease that is associated with high mortality if left untreated [30]. This study aims to assess the prognostic accuracy of the above MPCA-based model to predict 1-year mortality in PAH. Therefore, evaluating prognosis is key to identifying high-risk patients and optimising their management strategies as recommended by the European Society of Cardiology guidelines [12, 13]. Multiple clinical parameters are routinely obtained to evaluate PAH disease progression, including pulmonary haemodynamics from right heart catheterization (RHC), functional data from exercise tolerance and pulmonary function tests, biochemistry including N-terminal pro-B-type natriuretic peptide (NT-proBNP) and imaging including echocardiogram and CMR. The REVEAL score is a composite clinical risk score for mortality that combines these clinical parameters to predict 1-year mortality [223]. In addition, CMR measurements such as right ventricular volumes and function have been shown to predict mortality in PAH [4]. Thus, the availability of detailed patient phenotyping and prediction scores allows setting a clinical benchmark for the performance of machine learning prognostic models in PAH. This study assesses the additive value of the MPCA-based model to predict mortality compared with established prognostic parameters such as the REVEAL risk score and CMR measurements.

6.2 Materials and Methods

The Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) checklist for reporting prediction model development and validation was followed [243] and is available in the supplemental material.

6.2.1 Study Sample

Study population All consecutive treatment-naïve patients with PAH referred for a baseline CMR between 2008 and 2019 were identified from the ASPIRE registry[191]. Eligibility criteria included: (i) a baseline CMR study performed within 14 days of a PAH diagnosis, confirmed by RHC, and before the commencement of PAH treatment. (ii) Minimum 12 months follow-up or death within 12 months post-CMR study. The study population was divided into two cohorts: (i) a training cohort whose CMR images were used to develop and optimise the prognostic algorithm and (ii) a validation cohort that was left out of

training and used to validate the performance of the prognostic model. The cohort was split 70 : 30 into the model development and model validation cohort.

Ethical approval was obtained from the local ethics committee and written consent was waived for this retrospective study (ref $c06/Q2308/8$).

6.2.2 MR imaging protocol

Cardiac magnetic resonance was performed with a 1.5 Tesla GE HDx (GE Healthcare, Milwaukee, USA) system using an eight-channel cardiac coil. Four-chamber (4Ch) and short-axis (SA) cine images were acquired using a cardiac-gated multislice balanced steady-state free precession sequence (20 frames per cardiac cycle, slice thickness $10mm$, $0mm$ inter-slice gap, field of view $480mm$, acquisition matrix 256×200, flip angle 60°, $BW125KHz/pixel$, TR/TE $3.7/1.6ms$). A stack of images in the SA plane were acquired fully covering both ventricles from base to apex. End-systole was considered to be the smallest cavity area. End-diastole was defined as the first cine phase of the R-wave triggered acquisition or largest volume. Patients were in the supine position with a surface coil and with retrospective ECG gating.

Volumetric and ventricular function analysis was performed with a fully automated and validated segmentation tool ([7]) that contoured the ventricular endocardial borders at end-diastole and end-systole on the SA images using MASS software (MASS, 2020; Leiden University Medical Center, Leiden, the Netherlands). Papillary muscles and trabecula were included in the blood volume.

6.2.3 Image preprocessing

Mid-chamber SA and 4Ch cine images were used in this study. Images were processed following methods in a previous study [244]. In brief, images were preprocessed by (I) standardising CMR voxel units between subjects, (II) registering to each other using three anatomical landmarks, (III) masking surrounding tissues, and (IV) downscaling image size (Figure 6.1 and Figure 6.2).

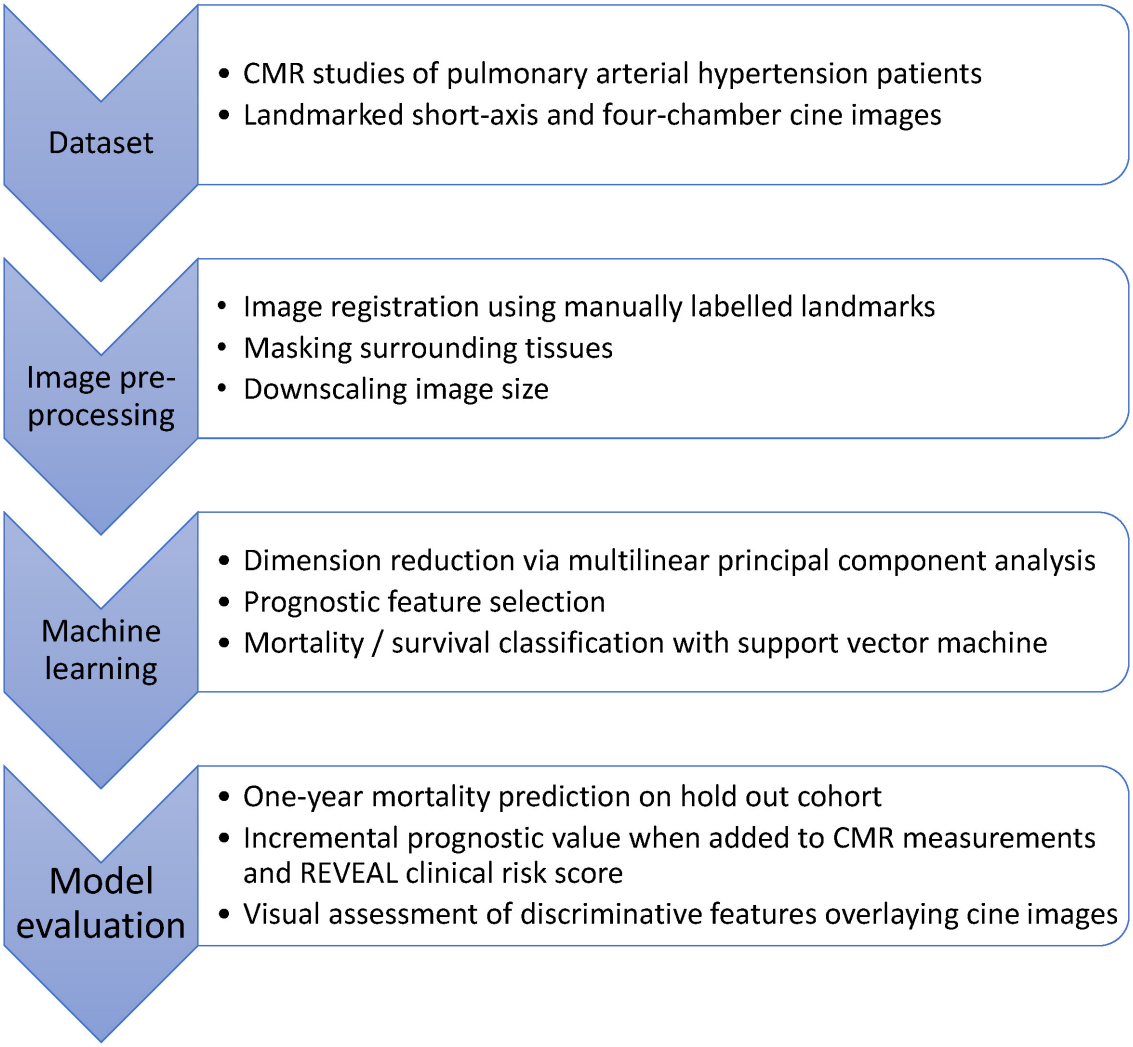


Figure 6.1: Model development flow chart.

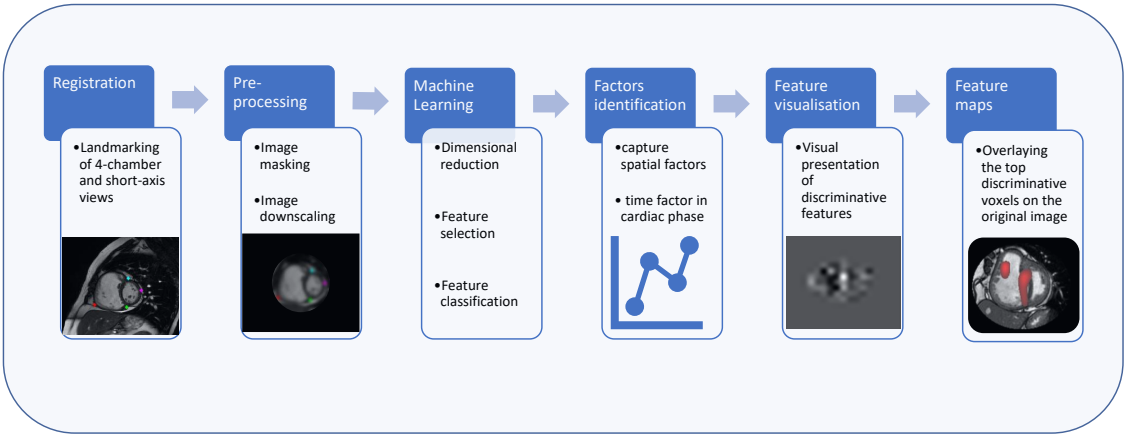


Figure 6.2: Model pipeline.

Standardisation and inhomogeneity bias correction

Cardiac MRI voxel units were standardised between subjects by z-scores to allow comparison between the different scans.

Landmarking and Registration

Rigid image registration was used based on three predefined fixed anatomic landmarks aligning the hearts on different scans to the same image space. The landmarks were manually placed on SA (superior insertion point; right ventricular free wall inflexion; mid-left ventricular lateral wall) and 4Ch (left ventricle apex; lateral mitral annulus; lateral tricuspid annulus) (Figure 6.3). Cardiac MRI images were landmarked by a single reader (S.A.) using MIM software (MIM Software Inc., Cleveland, Ohio, United States) with a customised workflow developed by a University of Sheffield clinical scientist (M.S) that facilitated storing the coordinates of the landmarks in an online spreadsheet. Independent visual quality assurance checks were performed (S.A.; J.U.).

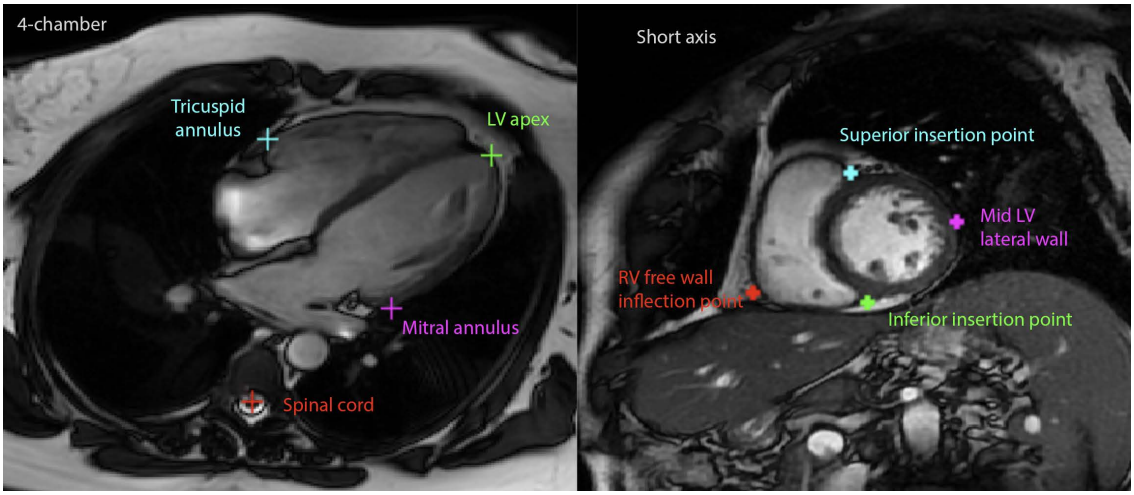


Figure 6.3: Landmarks on 4-chamber and short-axis view

Image masking

To focus on spatially relevant features, an ellipsoidal mask was fitted around the heart which blurred and smoothed out the image outside the heart. The smoothing increased incrementally with increased distance away from the heart by dilating the edges of the mask towards the image boundaries.

Scaling

The spatial resolution of the CMR images had to be reduced to increase the signal to noise ratio. A low level of signal to noise ratio can impair the ML model by learning features that are part of the noise (e.g. overfitting the model) rather than actual prognostic features. The resolution of the CMR in preprocessing has therefore been downsampled from 512×512 to four smaller scales: 32×32, 64×64, 128×128 and 256×256 to identify the best resolution for prediction.

6.2.4 Multilinear principal component analysis pipeline

The machine learning (ML) model includes three main steps: (i) feature extraction (ii) feature selection and (iii) decision function [244].

Feature extraction

In the feature extraction step, the ML model learns a low-dimensional representation from a high-dimensional input. The dimensional reduction in ML is necessary for high dimensional data analysis to achieve meaningful and efficient machine learning. Principal component analysis (PCA) is one of the most common methods of linear dimension reduction. The MPCA method extends the utility of PCA from linear data to multilinear data including tensors. MCPA converts the tensor to a low dimensional space while retaining most of the data variation of the original data [245]. For example in a CMR sequence with a three dimensional tensor including spatial dimensions of an image ($A \times B$) and a time dimension defined by the number of frames in the sequence (C). MCPA converts the CMR sequence tensor ($A \times B \times C$) into three matrices ($A \times P, B \times Q$ and $C \times R$) for each sample of the CMR sequence. These three matrices are optimised during training and capture the variation from each sample of the CMR sequence. The MCPA uses the optimised three matrices to map a new tensor ($P \times Q \times R$) which, in effect, is a low dimensional representation of the original tensor ($A \times B \times C$) (Figure 6.4).

Feature selection

In the feature selection step, the ML model focuses on a small sample of extracted features and rank them based on the highest discrimination performance by performing Fisher's discriminant analysis.

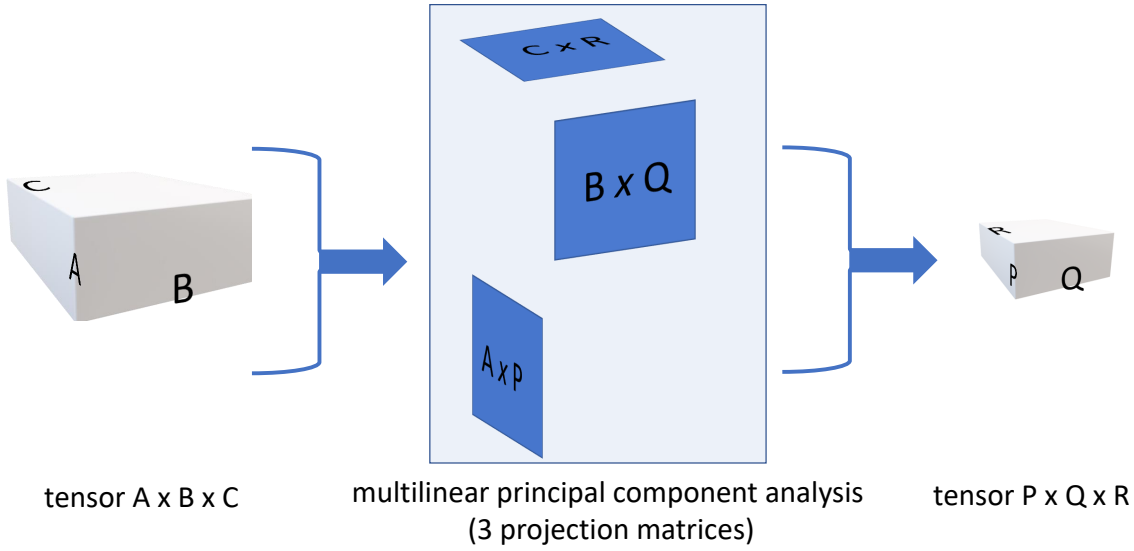


Figure 6.4: Multilinear principal component analysis.

Feature classification

The final step involves training the model to classify newly observed features into categories of survival and non-survival based on the top-ranked features extracted and selected in the previous steps. A simple linear method is chosen in the decision function as it is more transparent and interpretable and less prone to overfitting than complex, non-linear methods. The linear classifier model determines voxel-wise weighting for the learnt features, which then can be mapped back onto the original CMR images. The linear prognostic prediction was achieved by training support vector machines (SVMs) on MPCA features extracted from CMR studies [9, 242]. The methodology followed the MPCA-based pipeline in previous studies [9, 244]. This pipeline was trained through 10 rounds of 10-fold cross-validation on the development cohort ($n = 516$). For each fold during training, MPCA features were extracted and ranked for prognostic capability and selected using a step-wise feature inclusion method. This was performed using a random tuning-set ($n = 50$) of cases. The feature set with the highest tuning-set performance was used to train an SVM and tested on the left-out fold. The feature set with the highest fold-performance was used to train the final development SVM. This MPCA-based machine learning model was then applied to the completely left-out validation cohort ($n = 207$). On a standard computer, the time it takes to process each image and perform inference is much < 0.1 second. A Jupyter notebook tutorial of the open-source pipeline code

is available at: `https://colab.research.google.com/github/pykale/pykale/blob/main/examples/cmri_mpca/tutorial.ipynb`

6.2.5 Visualisation of tensor features

Trained features were visualised by using MPCA reconstruction to obtain spatially relevant feature maps. To visually inspect the impact of specific regions on high- and low-risk of mortality prediction, a two-step procedure of thresholding and clustering was implemented. Voxels containing high absolute values (high positive = high-risk, high negative = low-risk) of MPCA features were thresholded. Morphological dilation-erosion using a spherical structural element ($r = 2$) was performed and clusters of visually significant size were overlaid on individual patients' original CMR scans.

6.2.6 Clinical and mortality data

Clinical data including intermittent shuttle walking test, pulmonary function test (PFT), estimated glomerular filtration rate (eGFR) and serum level of NT-proBNP were collected before treatment was commenced. Data was collected from the electronic clinical database (Infoflex), laboratory results and radiological information system at Sheffield Teaching Hospitals (STH). Demographic data, WHO functional status, PAH subgroup diagnosis, and outcome were collected from the electronic medical system. Mortality data were collected from the electronic records of the National Health Service (NHS) Personal Demographics Service. The NHS automatically updates the mortality records once a death is registered in the United Kingdom. All patients were followed up as part of the national service specification for patients with pulmonary hypertension for a minimum of 12 months. No patients were lost to follow-up.

REVEAL score

The Registry to Evaluate Early and Long-term Pulmonary Arterial Hypertension Disease Management (REVEAL Registry) is the largest multicentre registry of a PAH population in the USA [27]. The REVEAL data is acquired from 55 centres and includes both consecutive incident and prevalent PAH cases. Based on the observational data from the REVEAL registry, a mortality risk equation was developed on the incident cases and validated on the prevalent cases [223]. The REVEAL combines several prognostic variables

including demographics, PAH subcategory, RHC, PFT, WHO class, walking distance, pericardial effusion and eFGR (Figure 6.5). The REVEAL score has become a commonly used assessment tool in clinical practice and gives a good indication of the mortality risk at baseline.

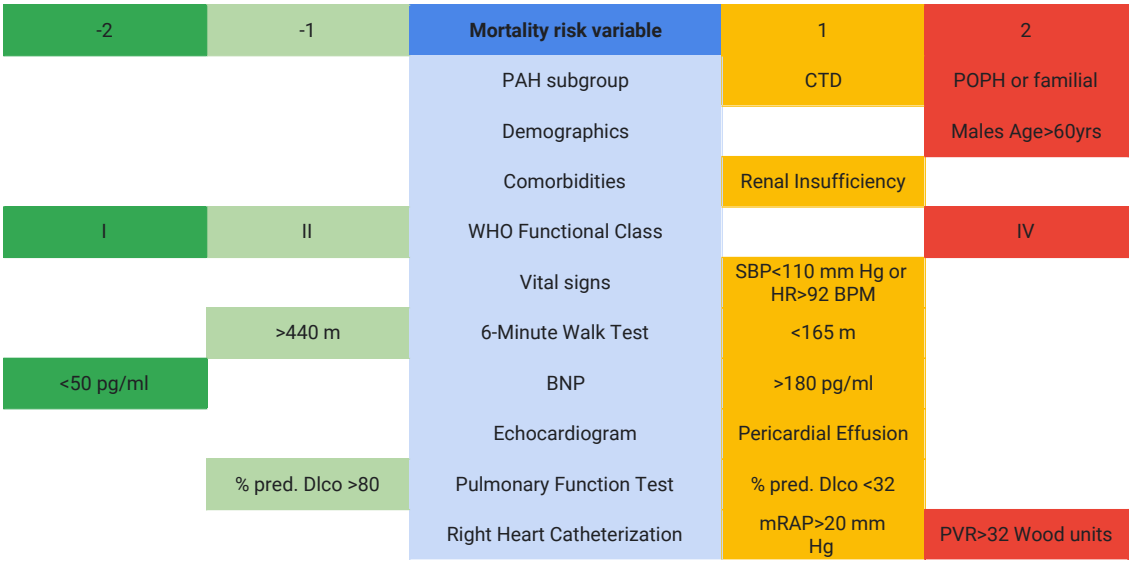


Figure 6.5: REVEAL PAH mortality risk score
CTD: connective tissue disease; PoPH: portal hypertension; WHO: World Health Organisation; BNP: brain natriuretic peptide; SBP: systolic blood pressure; HR: heart rate; Dlco: diffusing capacity of the lung for carbon monoxide; % pred: % predicted; mRAP: mean right atrial pressure; PVR: pulmonary vascular resistance.

Patients with PAH and particularly idiopathic PAH (IPAH) have an impaired TLCO due to pulmonary artery vasculopathy resulting in increased pulmonary vascular resistance [246, 247]. A reduction in TLCO is a significant independent prognostic marker in PAH that increases the risk of death in 5-years by four times [248]. Pulmonary function tests were adjusted for age, sex and ethnicity to calculate the per cent predicted PFTs using the online calculators provided by the Global Lung Function Initiative (GLI) [249, 250]. The values calculated were the %*predicted* forced expiratory volume (FEV1); the amount of air exhaled during the first second of forced breath, the %*predicted* forced vital capacity (FVC); the total amount of air exhaled during the FEV1 test and the %*predicted* transfer factor of the lung for carbon monoxide (TLCO) which measures the alveolar gas exchange with the pulmonary capillaries.

Pericardial effusion is described as a significant prognostic marker in current guidelines [13]. It has the highest prognostic value compared to other echocardiogram markers and

can indicate a doubled risk of mortality over five years [251, 252]. Presence of pericardiacl effusion was elicited from time matched (within 30-days) echocardiogram reports, chest computed tomography (CT) and CMR reports.

Exercise limitation is a common presenting symptom in PAH and is strongly associated with poor prognosis [253, 254]. Several exercise tests exist, but the most common test is the 6-minute walking test (6MWT) which is also incorporated in the REVEAL score. The 6MWT measures the distance covered by walking for six minutes on a level and flat surface. A distance of fewer than 165m is an indicator or poor function and prognosis while a distance exceeding 440m is a marker of good prognosis. A main limitation of the 6MWT functional test is that the maximal exercise capacity is limited by time and distance, rather than the functional ability of patients. This artificial ceiling of effect has led to concerns about the sensitivity to change in functional capacity in younger patients and those with less severe disease [255]. A more complex functional exercise that avoids this ceiling effect is the incremental shuttle walking test (ISWT) that is performed at the Sheffield Pulmonary Vascular Disease Unit (SPVDU). During the ISWT patients walk back and forth on a 10m walking course while increasing their pace every minute until the other end of the track can not be reached within one minute. The ISWT allows for assessing higher levels of exercise capacity measured by distance and speed achieved by the patient.

6.2.7 Statistical analysis

Continuous variables are presented as proportions, means ± standard deviations, or median and interquartile range for data not following a normal distribution. The sample size for developing the prediction model was calculated using a 1-year mortality prevalence of 10% and seven predictor parameters and required 420 patients to develop the mortality prediction model [256]. The REVEAL score was calculated from composite clinical parameters [223] and modified to include the incremental shuttle walk test instead of the 6 minute walking test[59, 224] using the following thresholds: a distance of $0 - 40m = +2$; $40 - 180m = +1$; $181 - 339m = 0$, $340 - 420m = -1$ and a distance $> 420m = -2$ REVEAL points [59, 224]. The CMR volumetric measurements were indexed for body surface area and corrected for age and sex by calculating the percentage predicted values

as per published reference data [197, 198]. The outcome of the MPCA-based pipeline was calculated as the SA and 4Ch probabilities based on the SVM prediction. A combined probability was calculated by further training a dual-scan SVM from the selected features of both individual models - SA and 4Ch. All variables were standardised by subtracting the mean for each variable and dividing it by its standard deviation (SD) to allow for more meaningful comparisons. A univariable Cox proportional hazards regression was performed to estimate the 1-year mortality prediction of the REVEAL score, CMR measurements, and the MPCA probabilities. For the multivariable analysis, we planned to include the CMR measurements that were identified in previous prognostic studies, namely right ventricular ejection fraction (RVEF), right ventricular end-systolic volume index (RVESVi), right ventricular end-diastolic volume index (RVEDVi), left ventricular end-diastolic volume index (LVEDVi), left ventricular stroke volume index (LVSVi) and pulmonary artery relative area change (PA RAC) [4, 46]. Owing to the high correlation between RVESVi and RVEDVi ($r = 0.89$), we only included RVESVi as the stronger predictor in the multivariable analysis. The proportional hazards assumption was confirmed using scaled Schoenfeld residuals. The c-index was used to measure the relative goodness of fit between the different regression models. The c-index indicates the rate of correct predictions of survival the model makes. We also computed the Akaike information criterion (AIC) for each model. The AIC estimates the rate of incorrect prediction and compares the quality of different models relative to each other while penalising the models with more variables. While a higher c-index indicates a better model fit, a lower AIC value indicates fewer prediction errors [257].

In addition, the likelihood ratio test was performed to assess if there is a statistically significant difference between the different models and to determine the additive predictive value of the MPCA probabilities. The models compared were the univariable REVEAL score, the REVEAL score combined with prognostic CMR measurements, and finally a multiple variable model including the REVEAL score, CMR measurement and the MPCA probabilities. Kaplan-Meier curves were analysed to demonstrate the prognostic value of MPCA predictions dividing patients based on the median MPCA value as the threshold. The high and low mortality risk groups were compared using the log-rank (Mantel-Cox) test. The receiver-operating characteristic curve (ROC) and the area under the curve

(AUC) were used to estimate the prognostic accuracy of the different MPCA features. All tests were performed at .05 level.

6.3 Results

6.3.1 Study Sample Characteristics

A total of 737 consecutive incident patients with PAH were identified. Incomplete scans because of claustrophobia or patient intolerance were excluded, leaving 723 scans for the analysis. The training cohort included 516 and the validation cohort 207 patients (Figure 6.6).

The baseline characteristics of both cohorts are presented in Table 6.1. In summary, the study population were 74% females aged 59 $\pm$ 16 years and included PAH secondary to connective tissue disease (CTD) (46%), idiopathic PAH (IPAH) (27%), congenital heart disease (CHD) (16%), secondary to portal hypertension (7%) and other PAH subtypes (4%).

6.3.2 Mortality prediction

Survival analysis

The 1-year mortality rate in the validation cohort was 10% with an overall mortality rate over the total follow-up period of 29%. Kaplan-Meier survival analysis demonstrated a significant difference in survival in patients with high and low mortality risk in the validation cohort ($log-rank$ test $P < 0.001$) (Figure 6.7). The ROC curve for each model is shown in Figure 6.8. The AUC was 0.73 for the SA model, 0.64 for 4Ch, and 0.70 for the combined MPCA-based features to predict 1-year mortality in the validation cohort.

Univariable Cox regression analysis confirmed a strong prognostic utility of the SA and combined SA and 4Ch MPCA-based predictions (Table 6.2). However, the 4Ch features alone were not significant predictors of mortality. The univariable Cox regression hazard ratios for the demographics, RHC and CMR measurements, functional tests and clinical parameters are shown in Table 6.2. The REVEAL score and, PA RAC and age and sex-adjusted RVESVi were significant predictors of 1-year mortality.

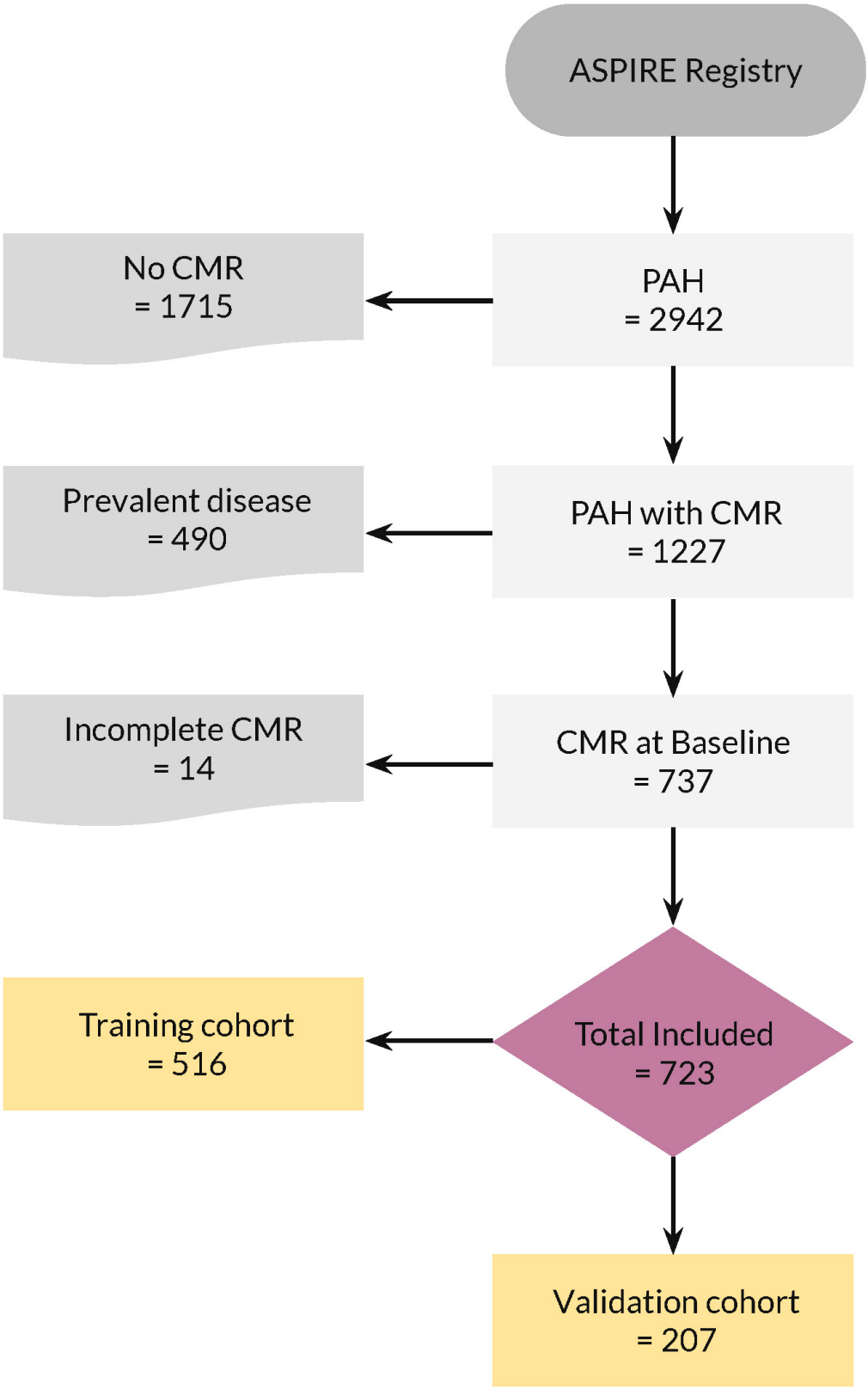


Figure 6.6: Study participants flow chart.

Table 6.1: Baseline characteristics.

	Training* *N = 516	***Validation* *N = 207***	***P Value***
Age (years)	62 (49 to 71)	62 (48 to 72)	.67
Sex (female)	376 (72%)	166 (80%)	**.04**
BSA (m^2)	1.82 ± 0.2	1.83 ± 0.2	.93
Diagnosis			
CHD	71 (13%)	34 (16%)	
CTD	242 (46%)	96 (46%)	
IPAH	137 (26%)	55 (26%)	
Portal hypertension	35 (6%)	17 (8%)	
other PAH	31 (6%)	5 (2%)	
WHO functional class			.33
I	2 (0%)	0 (0%)	
II	37 (7%)	11 (5%)	
III	409 (79%)	170 (82%)	
IV	58 (11%)	26 (12%)	
RHC			
mPAP (mmHg)	46 (34 to 56)	48 (38 to 56)	.16
PVR ($dyns.s.cm^{-5}$)	608 (330 to 883)	825 (455 to 1253)	**<.001**
PAWP (mmHg)	11 (8 to 13)	10 (8 to 12)	**<.001**
RA mean (mmHg)	9 (6 to 14)	9 (6 to 14)	.55
CO (L/min)	5 (4 to 6)	4 (3 to 5)	**<.001**
SvO_2 (%)	66 (58 to 71)	66 (58 to 73)	0.63
CMR			
RVEF (%)	37 ± 13	36 ± 11	.54
RVESVi (ml/m^2)	74 ± 35	76 ± 31	.12
RVEDVi (ml/m^2)	113 ± 41	115 ± 39	.16
RVEDMi (g/m^2)	27 ± 8	28 ± 8	**.02**
LVEF (%)	53 ± 10	53 ± 9	.97
LVESVi (ml/m^2)	31 ± 11	31 ± 16	.21
LVEDVi (ml/m^2)	67 ± 19	64 ± 21	.08
LVSVi (ml/m^2)	36 ± 12	34 ± 9	.13
VMI (ratio)	0.58 ± 0.2	0.62 ± 0.2	**.01**

Note.—Data presented as mean ± standard deviation or median (range). BSA, body surface area; CHD, congenital heart disease; CO, cardiac output; CTD, connective tissue disease; CTEPH, chronic thromboembolic pulmonary hypertension; EDVi, end-diastolic volume index; ESVi, end-systolic volume index; IPAH, idiopathic pulmonary arterial hypertension; LV, left ventricle; mPAP, mean pulmonary artery pressure; PAH, pulmonary arterial hypertension; PAWP, pulmonary arterial wedge pressure; PH, pulmonary hypertension; PVR, pulmonary vascular resistance; RA, right atrium; RHC, right heart catheterization; RV, right ventricle; RVEF, right ventricle ejection fraction; RVEDMi, right ventricular end-diastolic mass index; SV, stroke volume; SvO2 = mixed venous oxygen saturation; VMI, ventricular mass index; WHO, World Health Organisation

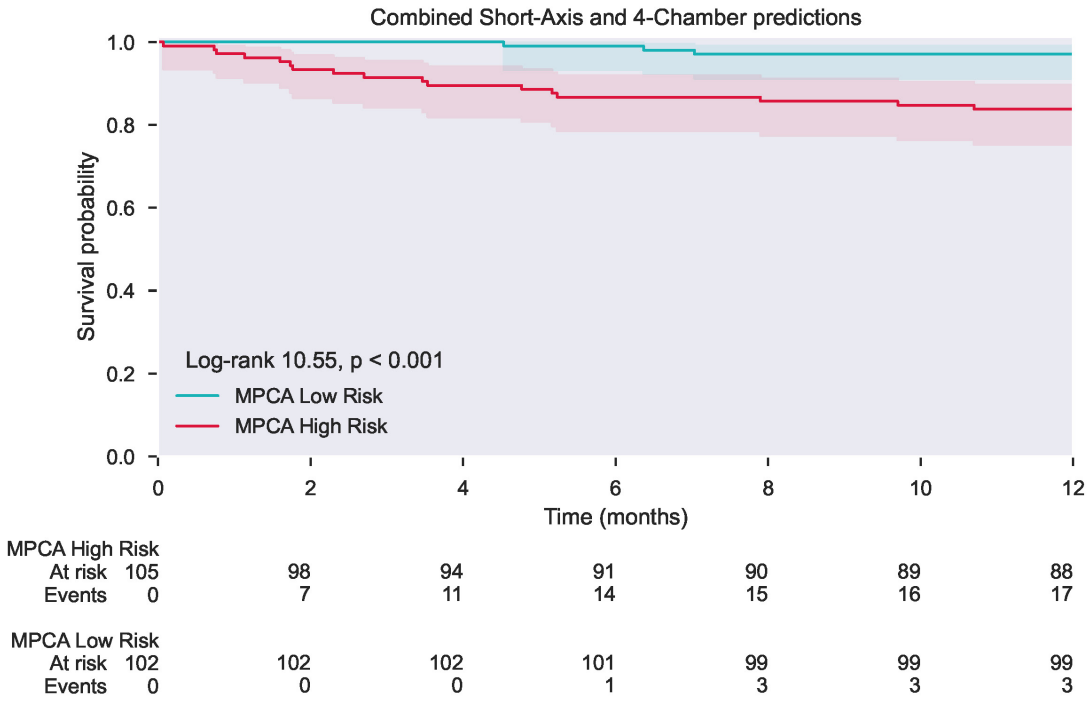


Figure 6.7: Kaplan–Meier curve. The Kaplan–Meier curve shows the survival of high- and low-risk patients based on the combined short-axis and four-chamber model predictions. The risk threshold was determined based on the median value of the MPCA predictions. The Kaplan–Meier analysis shows a significant difference in survival between the high and low risk of mortality patient groups ($log-rank\ P < 0.001$).

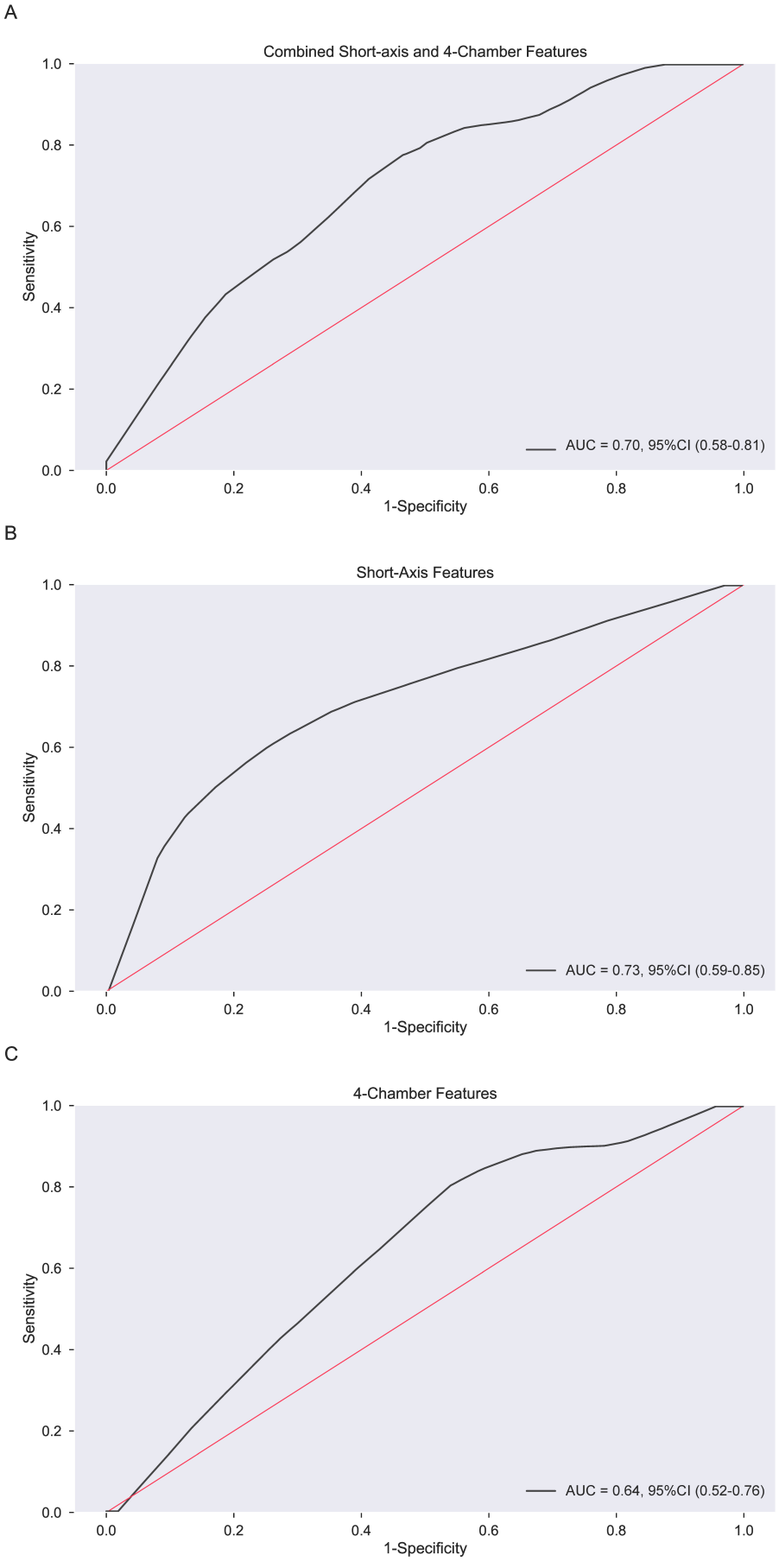


Figure 6.8: Receiver-operating characteristic curves for 1-year mortality prediction. The prognostic accuracy of the different machine learning models were compared (A) combined model, (B) short-axis model and (C) four-chamber model. The highest area under the curve was achieved with the short-axis model ($AUC = 0.73$).

Table 6.2: Univariable Cox proportional hazard regression ratios for 1-year mortality.

	Hazard Ratio		P Value
Age (years)	1.04	[1.01, 1.08]	.03
Sex	1.04	[0.35, 3.10]	.95
WHO class	2.03	[0.78, 5.29]	.15
REVEAL	1.34	[1.11, 1.62]	.002
RHC			
mPAP (mmHg)	0.98	[0.95, 1.02]	.31
PVR (dyns.s.cm^{-5})	1.00	[0.99, 1.00]	.62
PAWP (mmHg)	1.01	[0.87, 1.18]	.88
RA mean (mmHg)	1.08	[1.02, 1.15]	.01
CO (L/min)	0.84	[0.60, 1.19]	.32
SvO_2 (%)	0.97	[0.93, 1.01]	.13
CMR			
RVEF (% pred)	0.76	[0.45, 1.31]	.33
RVESVi (% pred)	1.70	[1.10, 2.63]	.02
RVEDVi (% pred)	1.44	[0.97, 2.14]	.07
RVEDMi (% pred)	1.11	[0.78, 1.59]	.55
LVEF (% pred)	1.28	[0.79, 2.06]	.32
LVESVi (% pred)	1.04	[0.65, 1.65]	.88
LVEDVi (% pred)	0.92	[0.56, 1.51]	.74
LVSVi (% pred)	0.93	[0.59, 1.47]	.77
VMI (ratio)	0.96	[0.53, 1.75]	.90
PA RAC (%)	0.91	[0.84, 0.99]	.03
Septal angle systole	0.99	[0.97, 1.03]	.94
Septal angle diastole	0.98	[0.94, 1.04]	.64
MPCA-based features			
SA features	2.40	[1.46, 3.95]	.001
4-chamber features	1.47	[0.98, 2.22]	.06
Combined features	1.97	[1.28, 3.03]	.002

Note.—CMR parameters are corrected for age and sex (%pred). For abbreviations see Table 6.1

Additive prognostic value

Several multivariable prognostic models were compared in Table 6.3 to compare the predictive value of the REVEAL score alone, REVEAL score combined with CMR measurements or MPCA features and finally REVEAL score combined with CMR measurements and MPCA features. The prognostic models were compared using the c-index and AIC test for goodness of fit and the log-rank test to assess the statistical significance of the difference between the models. The univariable REVEAL model allows the assessment of the 1-year risk of mortality based on available composite clinical data alone. Adding the MPCA-based predictions allows evaluating the added incremental value in predicting death compared with REVEAL and segmentation-based CMR parameters. The REVEAL score alone had a $c-index$ of 0.71 and AIC of 203. Adding CMR measurements improved the model statistically significantly, to 0.78 and AIC of 205 ($log-rank$ test $P = 0.003$). The model including MPCA prediction, REVEAL score and CMR measurements, showed the strongest prognostic utility ($c-index$: 0.83 and AIC 193, $log-rank$ test $P < 0.001$). The MPCA model alone had similar accuracy to the REVEAL score with a $c-index$ of 0.71 and AIC of 204.

Table 6.3: C-index and Akaike information criterion (AIC) for the univariable and multiple variable Cox regression analysis for the REVEAL score, CMR measurements and the MPCA machine learning model.

	C index		AIC	P Value
CMR measurements*	0.70	[0.60-0.80]	211	
MPCA**	0.70	[0.59-0.81]	204	
REVEAL score	0.71	[0.61-0.81]	203	
REVEAL + MPCA	0.76	[0.67-0.85]	197	**.003**
REVEAL+ CMR measurements	0.78	[0.70-0.86]	205	**.003**
REVEAL + CMR measurements + MPCA	0.83	[0.76-0.90]	193	**<.001**

Note.— A higher c-index indicates a better model fit and a lower AIC indicates a relative lower prediction error. The log-rank test indicates that the combination of MPCA, CMR measurements and REVEAL is statistically significantly more predictive than REVEAL score alone ($c-index$ 0.83 vs. 0.72, $P < 0.001$).
∗ CMR measurements included age and sex corrected right ventricular ejection fraction, right ventricular end-systolic volume index, left ventricular end-diastolic volume index, left ventricular stroke volume index and pulmonary artery relative area change.
∗∗ MPCA combined short-axis and four-chamber features.

Temporal prognostic dynamics

The MPCA-based features were assessed throughout the cardiac cycle and grouped according to the anatomical region into the right ventricle (RV), left ventricle (LV) and septum. For visualisation purposes, we manually segmented the averaged SA and 4Ch slice

to group the MPCA features into anatomical regions. The features were divided into low and high-risk features based on the median MPCA feature values used in the Kaplan-Meier analysis (Figure 6.9, Figure 6.10). On the SA views, abnormal interventricular septum during systole and particularly at end-systole and the LV chamber during diastole and particularly at end-diastole indicated a higher risk of mortality. On 4Ch views, the features with the highest impact on predicting mortality were at the RV at early systole. A normal LV and interventricular septum in diastole on SA and 4Ch imaging were the strongest predictors of survival, whereas the RV was a poor indicator of survival (Figure 6.9).

6.4 Discussion

This study assessed the prognostic utility of an MPCA-based machine learning model in CMR in patients with treatment-naïve PAH. This is the first study to localise prognostic PAH features with an explainable AI approach dynamically over the cardiac cycle. In addition, we have shown the incremental prognostic value of the MPCA model compared to known prognostic markers such as the REVEAL score and CMR volumetric measurements (Figure 6.11).

The advantage of using MPCA is its interpretability. The ability to directly relate prognostic features identified in the machine learning process helps understand and explain the machine learning model's findings. Diagnostic and prognostic models based on deep learning methods have been criticised for creating a 'black-box' situation where the predictions are often difficult to comprehend and retrace [258]. Visualising the MPCA features throughout the cardiac cycle allowed discerning the most significant discriminatory predictors of death on CMR in PAH. The known prognostic features identified in pulmonary hypertension of diastolic interventricular septal flattening [251], reduced LV size and increased RV size [4] can all be visually assessed on SA images. The most significant features identified in non-survivors on SA imaging were located at the septum at end-systole and LV at end-diastole. Changes in the interventricular septum at end-systole are the result of RV pressure overload. The altered pressure gradient between the LV and RV results in flattening of the septum giving a characteristic D-shaped LV and eventually results in impaired LV diastolic function and reduced LV filling [153, 259].

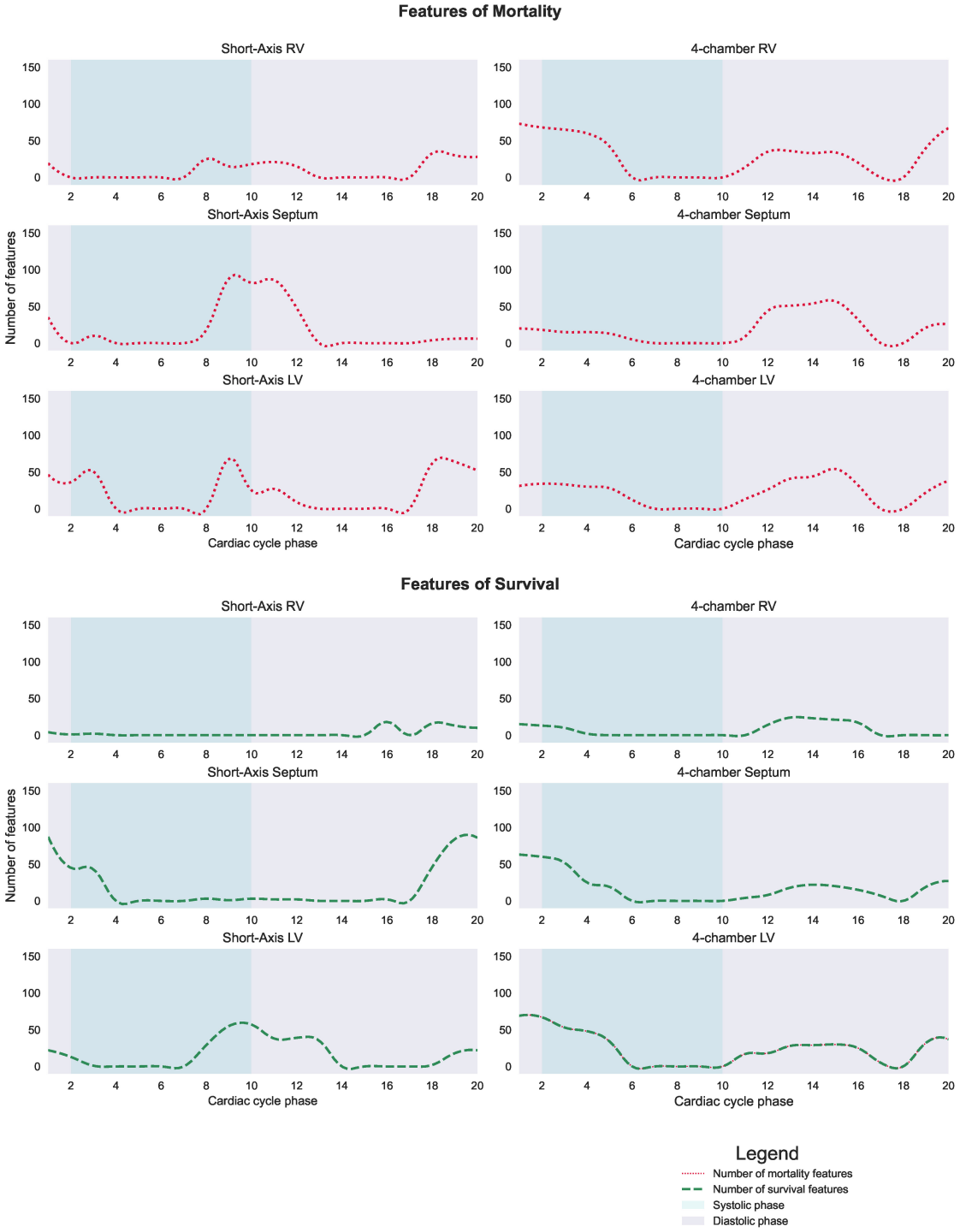


Figure 6.9: Time-resolved prognostic cardiac features. Features of poor prognosis and also protective features were examined throughout the cardiac cycle on the short-axis and four-chamber views. The most significant cardiac features were the end-systolic and early diastolic septum on both the short-axis and four-chamber views. The RV during systole and LV during diastole also were predictive of 1-year mortality. In contrast, the most important features of survival were the end-diastolic septum on short-axis and four-chamber views and the LV at systole.

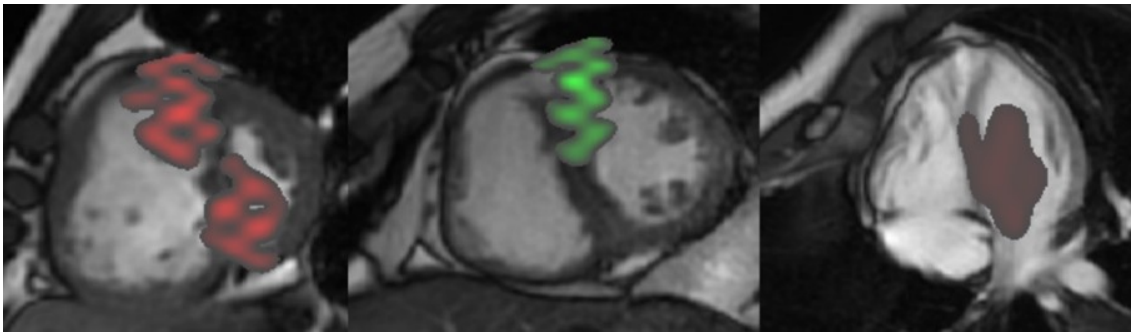


Figure 6.10: Visualisation of prognostic mortality and survival features learnt from the training data, overlaid on three example short-axis / four-chamber images from three different patients with PAH to interpret the corresponding anatomical regions. Left: Septum and LV features of high risk of mortality at end-systole. Middle image: Features of survival visualised at the septum at end-diastole. Right: four-chamber view showing high-risk features in the septum and LV in diastole.

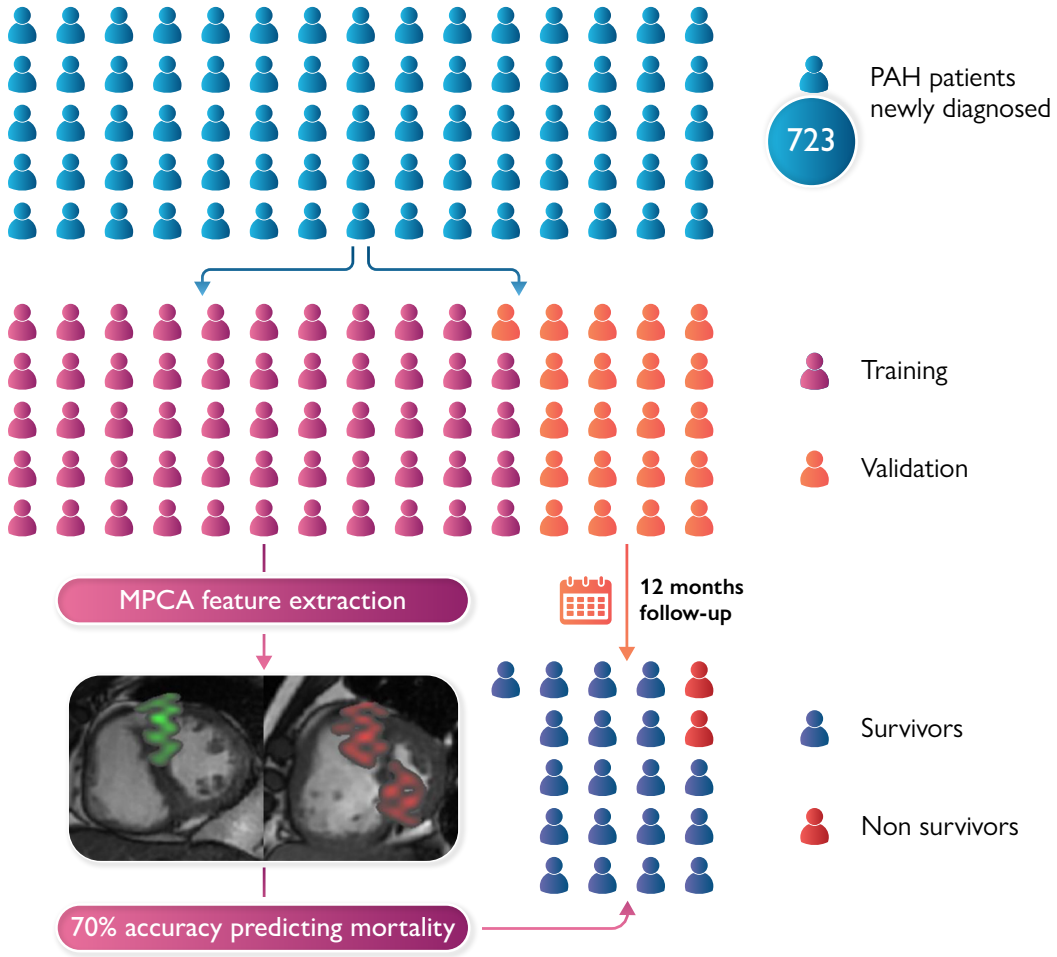


Figure 6.11: Prognostic features on CMR can be extracted using multilinear principal component analysis (MPCA) machine learning to predict outcome in pulmonary arterial hypertension (PAH).

Survivors showed the opposite with features in the septum at end-diastole and LV at end-systole. We found fewer overall features on SA images at the RV. However, on 4Ch imaging the most significant features were identified in RV systole. Whereas the septal and LV features were less important on 4Ch imaging. The 4Ch view allows assessing the longitudinal RV contractility which for example can be inferred on echocardiogram by assessing the tricuspid annular plane systolic excursion (TAPSE). Right ventricle longitudinal contraction is known to be the larger component of RV contraction and a key prognostic indicator [68, 70, 260] which explains its prognostic importance in PAH.

The MPCA-based model was developed and validated on CMR imaging performed at diagnosis and in treatment-naive PAH patients. Disease severity assessed at baseline assessment is important for planning an optimal treatment strategy. Almost all published prognostic CMR studies in PAH are based on disease prevalent PAH patients in later stages of the disease process [4]. A meta-analysis of 22 studies and almost 2000 patients with PAH showed that RVEF, RVESVi, RVEDVi, LVEDVi and LVSVi were significant predictors of mortality [4]. Right ventricle ejection fraction, RVEDVi, LVEDVi and LVSVi did not predict mortality in our baseline PAH cohort. The MPCA pipeline can therefore elicit cardiac changes before they affect RV function and size and adds prognostic value at baseline evaluation. In addition, comparing the MPCA to the REVEAL score allowed us to evaluate the incremental value benchmarked against a clinically validated baseline prognostic tool. The MPCA-based predictions significantly improved the 1-year mortality prediction of the REVEAL score. The prognostic model accuracy ($c - index$) using REVEAL improved from 71% to 83%($log - ranktestP < 0.001$) when it was combined with CMR data including MPCA predictions and CMR measurements. However, even without REVEAL data, mortality can still be accurately predicted based on MPCA features alone with an accuracy ($c - index$) of 70%.

The application of step-wise cardiac features extraction using CMR has further potential that can be evaluated in future developments. Comparing prognostic features at follow-up with baseline features might provide a better understanding of disease progression on CMR and might offer a standardised disease monitoring tool. In addition, technical improvements

would allow to fully automate the prognostic model, which currently requires manual image registration. Deep learning automated landmarking for image registration would reduce the manual processing of CMR images and reduce the time and cost associated with it [261, 262].

6.4.1 Limitations

This was an exploratory retrospective single centre study on patients with PAH. Findings will need to be confirmed in a prospective trial with an external validation cohort. In addition, applying the model to other diseases and MRI systems would further validate its generalisability. The MPCA method was applied on cine images of the mid-chamber slice throughout the cardiac cycle. Stack imaging of the whole heart can currently not be included in the MPCA model training. However, because of the strong prognostic signal from the SA and 4Ch cine images we envisage that future developments including 3D data of the heart will further improve prognosis prediction.

6.5 Conclusion

Patient outcome prediction in PAH can be enhanced by adding MPCA-based machine learning to CMR volumetric data and clinical risk scores. The MPCA analysis gives a population insight into the prognostic cardiac features in PAH in an explainable and visualisable approach.

6.6 Author Contributions

Guarantors of integrity of entire study, S.A., H.L., A.J.S.; study concepts and study design, S.A., A.J.S., J.U., H.L.; data acquisition, S.A., J.U., A.J.S.; data analysis/interpretation, S.A., J.U., H.L., A.J.S.; cardiac MRI acquisition, D.C.; CMR landmarking, S.A.; model development and training; J.U., H.L.; literature research, S.A.; manuscript drafting, S.A.; manuscript revision for important intellectual content, S.A., A.J.S., J.U., H.L., D.G.K., R.J.v.d.G.; statistical analysis, S.A., J.U.; manuscript editing, S.A., K.D., H.L., D.G.K., R.J.v.d.G., A.J.S.; and approval of final version of the submitted manuscript, all authors.

CHAPTER 7

A Meta-analysis of Myocardial T_1-mapping and Extracellular volume in Pulmonary Hypertension

Samer Alabed[1], Laura Saunders[1], Pankaj Garg[1], Yousef Shahin[1], Faisal Alandejani[1], Andreas Rolf[2], Valentina O. Puntmann[3], Eike Nagel[3], Jim M. Wild[1], David G. Kiely[1] and Andy J. Swift[1]]

[1] Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield

[2] Department of Cardiology, Kerckhoff-Heart Center, Bad Nauheim, Germany

[3] Institute for Experimental and Translational Cardiovascular Imaging, Frankfurt am Main, Germany

This chapter is based on a publication in *Magnetic Resonance Imaging.* 2021 Jun;79:66-75.

Abstract

Introduction Elevated myocardial T_1-mapping and extracellular volume (ECV) measured on cardiac MR (CMR) imaging is associated with myocardial abnormalities such as oedema or fibrosis. This meta-analysis aims to provide a summary of T_1-mapping and ECV values in pulmonary arterial hypertension (PAH) and compare their values with controls.

Methods We searched CENTRAL, MEDLINE, Embase, and Web of Science in August 2020. We included CMR studies reporting T_1-mapping or ECV values in adults with any type of PAH. We calculated the mean difference of T_1-values and ECV between PAH and controls.

Results We included 12 studies with 674 participants. T_1-values were significantly higher in PAH with the highest mean difference (MD) recorded at the RV insertion points (RVIP) (108 milliseconds (ms), 95% confidence intervals (CI) 89 to 128) , followed by the RV free wall (MD 91 ms, 95% CI 56 to 126). The pooled mean T_1-value in PAH at the RVIP was 1,084, 95% CI (1,071 to 1,097) measured using 1.5 Tesla Siemens systems. ECV was also higher in PAH with an MD of 7.5%, 95% CI (5.9 to 9.1) at the RV free wall.

Conclusion T_1-mapping values in PAH patients are on average 9% higher than healthy controls when assessed under the same conditions including the same MRI system, magnetic field strength or sequence used for acquisition. The highest T_1 and ECV values are at the RVIP. T_1 mapping and ECV values in PH are higher than the values reported in cardiomyopathies and were associated with poor RV function and RV dilatation.

7.1 Introduction

Native myocardial T_1 and extracellular volume (ECV) mapping have shown promise as novel biomarkers to support diagnostic, therapeutic and prognostic decision making in several cardiovascular disorders [78]. T_1 mapping produces a pixel-by-pixel representation of the longitudinal relaxation times (T_1) within a tissue [79]. This relaxation time can be measured on a MRI system and sequence-specific standardised scale to provide surrogate tissue characterisation data [80]. In the myocardium, T_1 times are affected by two main factors; oedema and collagen in the interstitial space [81]. Oedema can be secondary to inflammation or infarction, whereas increased collagen is associated with fibrosis or infiltrative processes [82, 263–265]. An elevated T_1 value is, therefore, a non-specific tissue composition marker for conditions such as myocardial infarction, myocarditis, cardiomyopathies and diastolic heart failure [83, 263, 266–272]. A low T_1 is being used as a diagnostic tool and follow-up biomarker in Anderson-Fabry disease [273–277] and can be used as a complementary sequence to T2* in thalassemia [278, 279].

Performing T_1 mapping after contrast administration enables the assessment of the extracellular space [78, 81]. As gadolinium collects in the extracellular fluid, its paramagnetic effect causes shortening of the T_1 values of the myocardium. The T_1 shortening is proportional to the concentration of the gadolinium in the extracellular fluid. Therefore, combining pre- and post-contrast T_1 values of the myocardium and blood pool with the haematocrit allows estimation of the extracellular volume (ECV) [78]. Elevated ECV is seen with myocardial fibrosis or oedema and is associated with an increased risk for mortality[280–282]. Native T_1 and ECV can provide prognostic information in coronary artery disease and nonischemic cardiomyopathies [81, 283, 284] and might play a role in disease risk stratification, early diagnosis and monitoring progression [79, 269, 270].

In pulmonary arterial hypertension (PAH), elevated pulmonary artery pressure causes significant afterload on the right ventricle (RV). Eventually, the RV hypertrophies and dilates triggering a fibrogenic process [285]. T_1 mapping and ECV might, therefore, play a role in evaluating the changes in the RV [12] and the degree of fibrosis in PH [285].

Previous reviews have assessed normal T_1 mapping and ECV values in healthy people [286] and pathological values in cardiomyopathies [287–289]. Several studies report T_1

values in PAH, but there is currently no meta-analysis to summarise their results. This meta-analysis aims to compare the range of T_1 values and ECV in PAH patients to control participants and identify the regions of the myocardium with the highest T_1 values.

7.2 Methods

The review was prospectively registered with The International Prospective Register of Systematic Reviews (PROSPERO) on 10/02/2020 (ID: CRD42020166392).

Criteria for considering studies for this review

We considered any study comparing T_1 values or ECV rates in adult patients with PAH and controls; such as controlled trials, cohort studies or case-control studies. Studies with less than 10 participants and case reports were excluded. Inclusion was considered irrespective of prospective or retrospective recruitment, publication date, publication status or language.

Outcomes

1. The pooled mean difference of T_1 value between PAH and controls
2. Identifying the myocardial region with the largest T_1 values and ECV in PAH
3. Comparing the T_1 values in the subgroups of PAH

Search methods for identification of studies

Electronic searches

The following databases were systematically searched for relevant studies on 08/08/2020: (i) Cochrane Central Register of Controlled Trials (CENTRAL) (ii) MEDLINE (ProQuest, 1946 to Aug 2020) (iii) Embase (Ovid, 1974 to 2020 Week 32) and Web of Science. The reference lists of all relevant articles identified during the full-text screening were scrutinised for relevant studies.

The following search strategy was used:

1. exp "HYPERTENSION, PULMONARY"/
2. exp "PULMONARY HEART DISEASE"/
3. exp "PULMONARY VASCULAR DISEASE"/

4. (pulmonary ADJ2 hypertensi*)
5. (1 OR 2 OR 3 OR 4)
6. (T1 ADJ3 (map* OR value OR time* OR native OR contrast))
7. (SASHA OR MOLLI)
8. (recovery AND ((shot AND saturation) OR (modified AND look)))
9. (ECV OR (extracellular ADJ1 volume))
10. 6 OR 7 OR 8 OR 9
11. (5 AND 10)

Statistical analysis and data synthesis

We used Review Manager 5.4 (The Cochrane Collaboration, 2020) to perform a meta-analysis of the mean differences (MD) and produce the forest-plot. A random-effect model with 95% confidence intervals (CI) was used in the analyses. The available data allowed us to calculate the mean difference for 1.5 T field strength only. We pooled the means and standard deviations (SD) of T_1 values for the PAH and control groups when they were measured using the same MRI system and field strength, which was only possible for 1.5 T Siemens systems. The non-weighted means of T_1 values with their 95% CI for PAH patients and controls were presented on a forest plot using GraphPad Prism version 8.3 (GraphPad Software, La Jolla CA, USA). If the T_1 and ECV values for the RV insertion points were measured both at the superior and inferior insertion points, data for the inferior insertion points were chosen, as this was the case with the majority of the studies. A funnel plot to assess publication bias was not performed as the number of included studies in each meta-analysis were too low to identify real asymmetry [290].

7.3 Results

7.3.1 Results of the search

Our comprehensive search identified a total of 12 studies that were included in the meta-analysis. Nine studies reported T_1 mapping values in PAH [67, 84, 86–90, 291, 292] including one conference abstract [293] and five studies reported ECV values [58, 67, 89, 90, 294]. The details of the literature search are presented in the PRISMA flow diagram Figure 7.1.

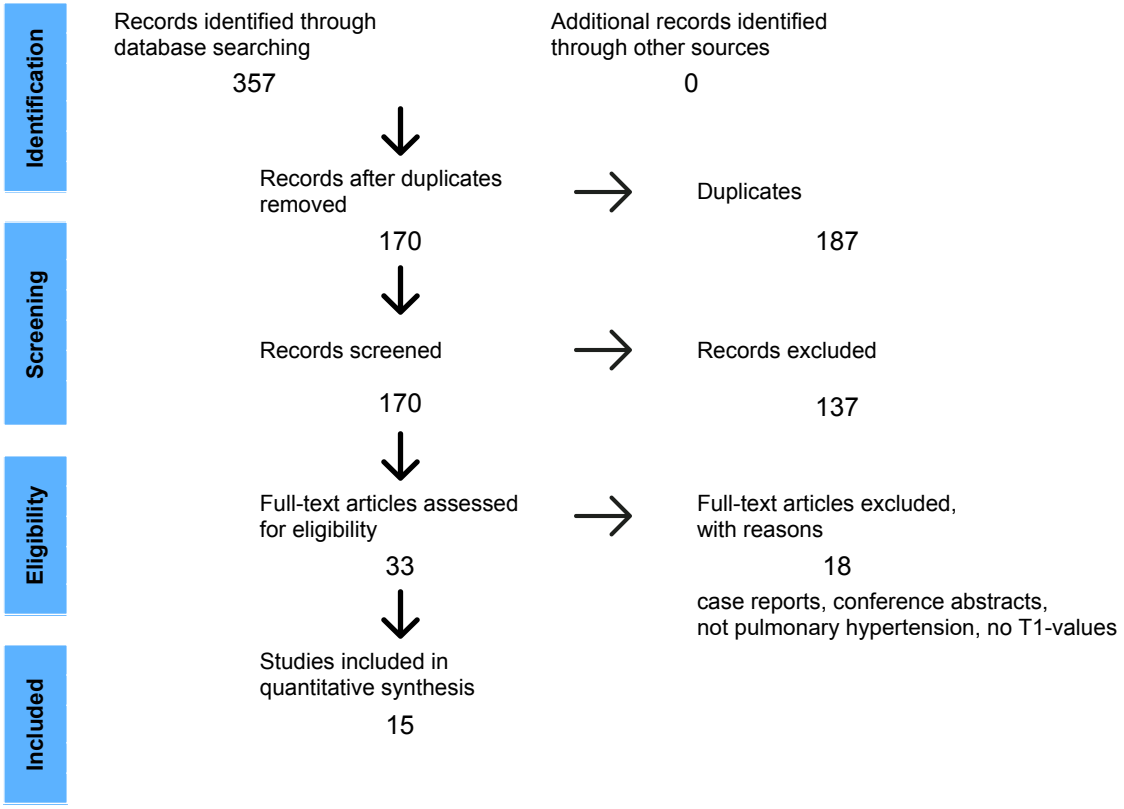


Figure 7.1: PRISMA Flow Chart

7.3.2 Description of included studies

Study Design

The review includes nine case-control studies published and three case-series. The studies were published between 2015 and 2020. Prospective recruitment was performed in seven studies and retrospective recruitment in five studies. Only three studies had a sample size >60. The largest study was Saunders 2018 with 223 PAH patients and 24 controls.

Population

The studies were conducted in 8 different countries; 6 studies were conducted in Europe, 2 in the USA and 4 in Japan and China. The studies included 554 PH patients of which 513 (93%) had PAH, 32 (6%) had chronic thromboembolic pulmonary hypertension (CTEPH) and 9 (1%) had PH secondary to lung disease. The control group included 120 people, of whom 97 were healthy and 23 were non-PH patients. The age of patients with PAH was 53 ± 15 years and 64% were women with a pooled average mPAP of 48 ± 15 mmHg, and RVEF of 42 ± 14%. The control group had a pooled average age of 49 ± 7 years and 52% were women. The pooled average of RVEF was 55 ± 5%. Details of included studies are

presented in Table 7.1.

MRI systems and T_1 techniques

The majority of studies were performed using 1.5 Tesla field-strength MRI systems. A 3 Tesla system was used in Asano 2018, Dong 2018 and Reiter 2017 [58, 87, 293]. Siemens MRI scanners were used in most studies apart from Saunders 2018 and Homsi 2017 who used a GE and Philips MRI system, respectively [67, 84]. A modified look-locker inversion recovery (MOLLI) sequence was used in all studies apart from Mehta 2015 who used accelerated and navigator-gated look-locker imaging for cardiac T_1 estimation (ANGIE) [90]. T_1 mapping values were measured on short-axis images on a single mid-chamber slice [58, 84, 88, 291, 294], single basal slice [292], average of two mid-chamber slices [90], average of a basal and a mid-chamber slice [86, 89] or averaged over basal, mid-chamber and apical slices [67, 87]. T_1 values were measured in end-diastole, apart from Mehta 2015 and Reiter 2017 who measured T_1 values in systole [87, 90]. None of the included studies used stress MRI to assess T_1 values. Details of the MRI systems, techniques and sequences used are provided in Table 2.

7.3.3 Results of the Meta-analysis of T_1 values and ECV

The mean difference of T_1 value in PAH and controls

Seven studies compared myocardial T_1 values at 1.5 Tesla between 375 PAH patients and 87 healthy controls. The T_1 values in PAH are significantly larger than the T_1 values of the control group. The largest difference was reported at the RVIP (MD 108 ms, 95% CI 89 to 128), followed by the RV free wall (MD 91 ms, 95% CI 56 to 126). The mean difference at the mid septum and LV lateral wall were relatively smaller (MD 56 ms, 95% CI 41 to 72) and (MD 36 ms, 95% CI 14 to 58), respectively. The forest plot of the meta-analysis of mean differences is presented in Figure 7.2.

Two studies reported myocardial T_1 values at 3 Tesla; Reiter 2017 and Asano 2018 [87, 293]. However, their results could not be pooled as they reported T_1 values at different regions of the myocardium. In addition, Reiter 2017 included non-PH patients as the control group, whereas Asano 2018 included healthy people. The highest calculated mean difference of the T_1 values in Reiter 2017 was 105 ms, 95% CI (76 to 133) at the RVIP. The mean difference at the septum and LV lateral wall were smaller 63 ms, 95% CI (40 to

Table 7.1: T_1 mapping Study Characteristics

	Study Author Year	Country	Design	Study period	MRI system	Software used	Technique	Group	Size	Sex F%	Age	mPAP	RVEF	PH Type
PH and Control	Asano 2018	Japan	RCC	2015 - 2017	3 T Siemens, Magnetom Verio	N.R	MOLLI	PH	30	N.R	N.R	41±13	N.R	100% PAH
								Control	10	N.R	N.R	N.R	N.R	Healthy
	Chen 2017	China	PCC	2015 - 2016	1.5 T Siemens, Magnetom Aera	Argus, Siemens	MOLLI	PH	22	73%	40 ± 13	60 ± 18	35 ± 11	PAH 82%, CTEPH 18%
								Control	10	60%	38 ± 14	N.R	56 ± 6	Healthy
	Homsi 2017	Germany	PCC	2014 - 2015	1.5 T Philips, Ingenia	Segment, Medviso	MOLLI	PH	17	47%	64 ± 14	N.R	40 ± 13	PAH 100%
								Control	20	50%	63 ± 11	N.R	55 ± 3	Healthy
	Mehta 2015	USA	RCC	N.R	1.5 T Siemens, Magnetom Avanto	Argus, Siemens	ANGIE	PH	12	67%	61 ± 19	N.R	36 ± 10	PAH 83%, CTEPH 17%
								Control	10	80%	24 ± 3	N.R	55 ± 5	
	Pagano 2016	Canada	RCC	2010 - 2013	1.5 T Siemens, Avanto or Sonata	N.R	SASHA	PH	6	83%	50 ± 18	N.R	38 ± 7	PAH 100%
								Control	21	52%	41 ± 16	N.R	58 ± 7	Healthy
	Patel 2020	USA	PCC	2014 - 2016	1.5 T Siemens, Magnetom Aera	MASS, Medis	MOLLI	PH	13	77%	59 ± 16	48 ± 8	44 ± 11	PAH 100%
								Control	8	38%	71 ± 4	N.R	53 ± 4	Healthy
	Reiter 2017	Austria	PCC	2012 - 2015	3 T Siemens, Magnetom Trio	Argus, Siemens	MOLLI	PH	35	57%	64 ± 16	44 ± 13	41 ± 14	PAH 51%, CTEPH 23%, Lung disease 20%, Other 6%
								Control	23	70%	59 ± 12	18 ± 4	54 ± 7	Non-PH patients
	Roller 2017	Germany	PCC	2014 - 2015	1.5 T Siemens, Magnetom Avanto	Argus, Siemens	MOLLI	PH	24	71%	63 ± 11	39 ± 11	40 ± 15	CTEPH 100%
								Control	24	50%	61 ± 12	N.R	65 ± 7	Healthy
	Saunders 2018	UK	RCC	2015	1.5 T GE, HDx	GE Advantage Workstation	MOLLI	PH	369	66%	58 ± 15	44 ± 13	43 ± 14	PAH 60%, CTEPH 25%, Lung disease 11%, Other 4%
								Control	24	51%	58 ± 4	N.R	N.R	Healthy
	Spruijt 2016	Holland	RCC	2011 - 2014	1.5 T Siemens, Magnetom Avanto	N.R	MOLLI	PH	70	73%	54 ± 16	47 ± 13	42 ± 16	PAH 86% , CTEPH 14%
								Control	10	40%	20 ± 1	N.R	N.R	Healthy
	Wang 2018	China	PCC	N.R	1.5 T Siemens, Magnetom Avanto	MASS, Medis	MOLLI	PH	18	44%	61 ± 12	44 ± 13	38 ± 11	PAH 100%
								Control	5	60%	37 ± 13	N.R	55 ± 3	Healthy
PH only	Dong 2018	China	RCS	2016 - 2017	3 T Siemens, Magnetom Trim Trio	MASS, Medis	MOLLI		60	65%	35 ± 15	59 ± 20	40 ± 12	PAH 100%
	Habert 2020	France	PCS	2012 - 2013	1.5 T Siemens, Magnetom Symphony	Argus, Siemens	MOLLI		12	50%	50 ± 16	44 ± 12	40 ± 18	PAH 83%, 17% CTEPH
	Jankowich 2019	USA	RCS	N.R	3 T Siemens, Magnetom Verio	Circle, Cardio-vascular Imaging	MOLLI		16	6%	70 ± 9	31 ± 7	46 ± 12	PAH 13%, Left heart and Lung disease 38% each, CTEPH 6%
	Nitsche 2019	Austria	PCS	2012 - 2017	1.5 T Siemens, Magnetom Avanto	Circle, Cardio-vascular Imaging	MOLLI		167	71%	71.9 (66.9–76.5)	32.0 (26.0–39.0)	52.0 (45.0–61.0)	78.1% PH, 21.9% non-PH HFpEF
	Tello 2019	Germany	PCS	2016 - 2018	1.5 T Siemens, Magnetom Avanto	Argus, Siemens	MOLLI		42	52%	56 ± 13	43 ± 14	38 ± 13	PAH 86%, CTEPH 14%

ANGIE, Accelerated and navigator-gated look-locker imaging; CTEPH, chronic thromboembolic pulmonary hypertension; HFpEF, Heart failure with preserved ejection fraction; MOLLI, Modified Look-Locker inversion recovery; mPAP, mean pulmonary artery pressure; n, number; N.R, not reported; PAH, pulmonary arterial hypertension; PH, pulmonary hypertension; PCC, prospective case-control; PCS, prospective case series; RCC, retrospective case-control; RCS, retrospective case series; RVEF, right ventricle ejection fraction; SASHA, Saturation recovery single-shot acquisition; T, tesla)

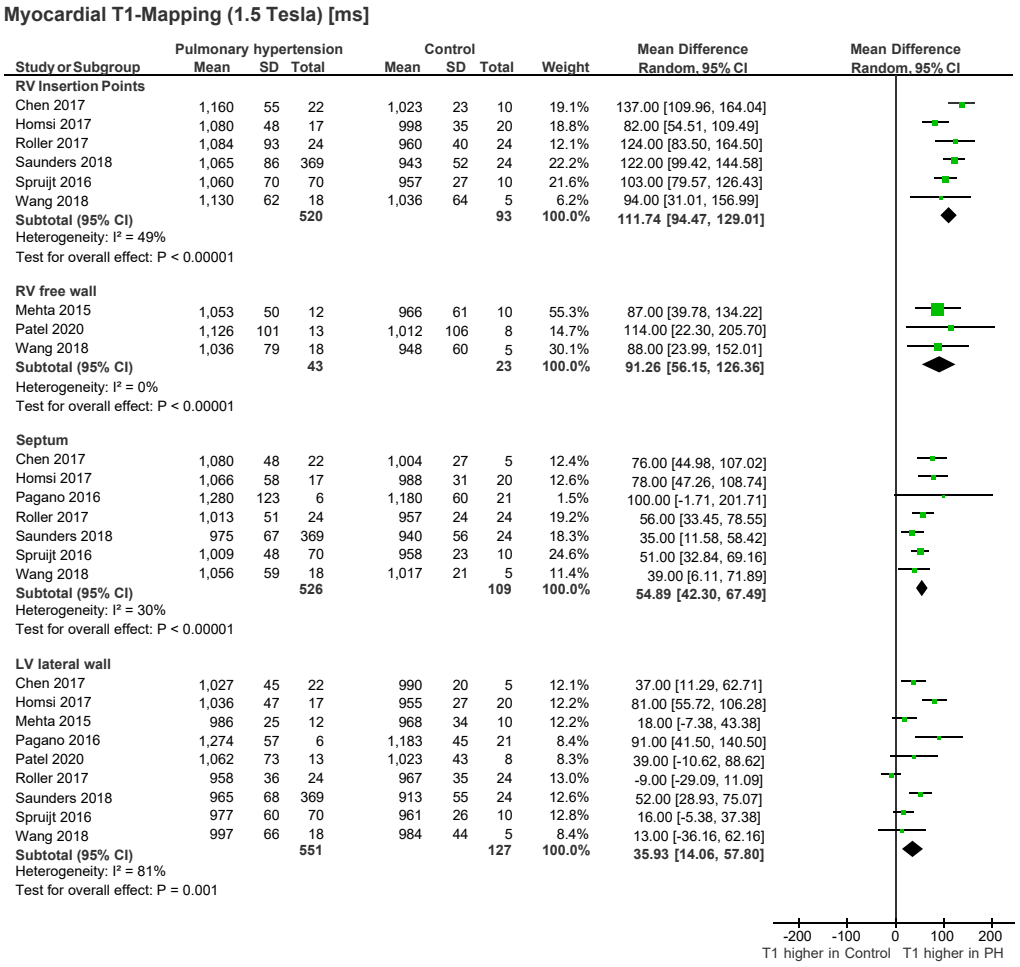


Figure 7.2: The mean difference of T_1 values in PH and controls.
CI, confidence intervals; ECV, extracellular volume; ms, millisecond; LV, left ventricle; PH, pulmonary hypertension; RV, right ventricle; SD, standard deviation

86) and 40 ms, 95% CI (20 to 60), respectively. In Asano 2018 the calculated MD at the RV free wall was 178 ms, 95% CI (134 to 222).

T_1 values in different myocardial regions

Comparing the mean difference in the T_1 values in different myocardial regions measured within PAH patients showed that the T_1 values at the RVIP is on average 6% higher than the septum (MD 62 ms, 95% CI 49 to 74) and 9% higher than the LV lateral wall (MD 102 ms, 95% CI 82 to 121). The pooled mean value of T_1 with a 1.5 T Siemens system for PAH patients and healthy controls are shown in Table 3 and their ranges are illustrated in (Figure 7.3). The T_1 values for PAH were pooled including values reported in the case series Tello 2019 and Habert 2020. Excluding the results of Mehta 2015, who used ANGIE sequences, did not significantly change the pooled values.

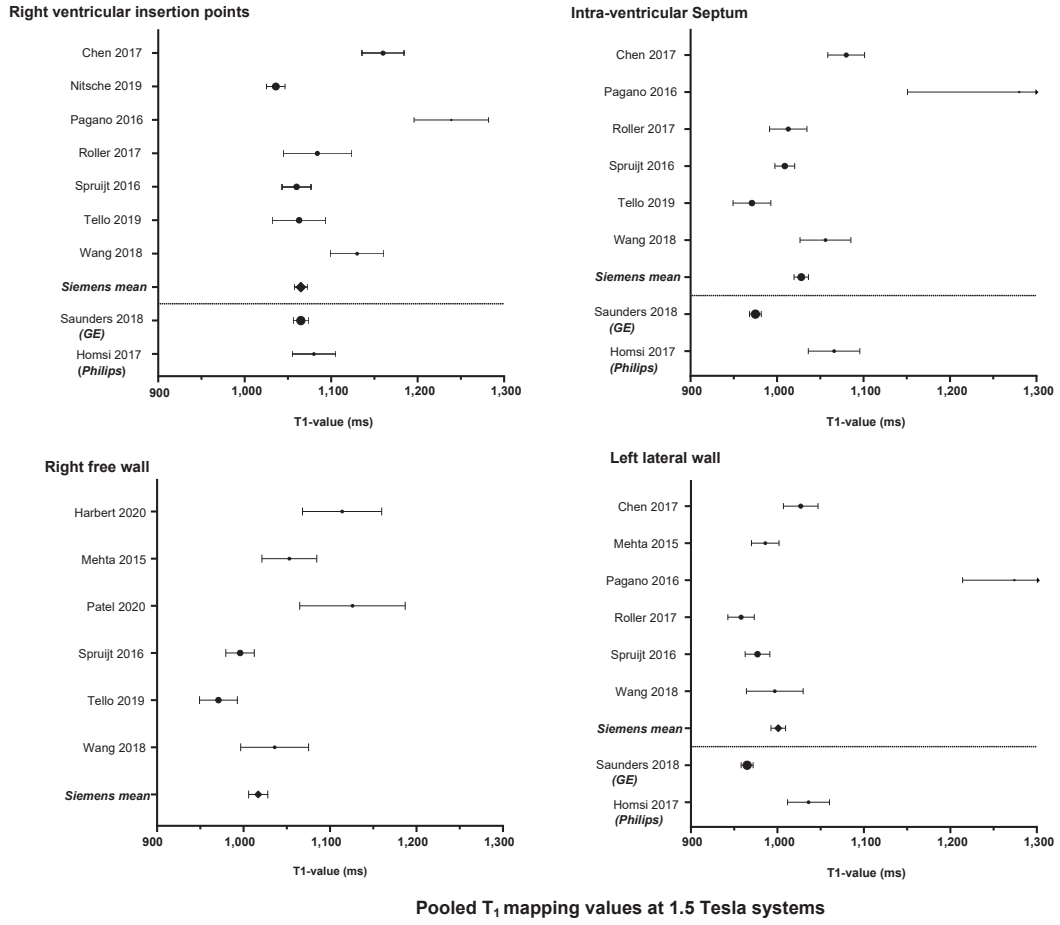


Figure 7.3: Pooled T_1 values in PAH.
For abbreviation list see legend for Figure 7.2.

T_1 values in different PAH subgroups

Three studies compared T_1 values in 140 patients with idiopathic PAH to T_1 values in 118 patients with PAH associated with CTD or CHD [67, 84, 291]. These studies showed no significant differences in the T_1 mapping values between the different PAH subgroups (MD 4 ms, 95% CI -18 to 26).

ECV mean difference

Five studies reported ECV values in PAH. Three used a 1.5 T Siemens and one a 1.5 [89, 90, 294], one a 1.5 T Philips system [67] and one a 3 T Siemens scanner [58]. Three studies compared the value of myocardial ECV at 1.5 T between PAH and healthy controls [67, 89, 90]. Due to the limited data pooling the mean differences of ECV values was only possible for the RV free wall and LV lateral wall and was significantly higher in PAH compared to the control group at both sites with a mean difference of 7.5%, 95% CI (5.9 to 9.1) at the RV free wall and 4.8%, 95% CI (1.6 to 8.1) at the LV free wall. The mean difference at the RVIP was reported in Homsi 2017 as 5.8%, 95% CI (2.2 to 9.4) and at the septum as 5.7%, 95% CI (3 to 8.4).The forest plot of the meta-analysis of ECV mean differences is presented in Figure 7.4.

ECV values in different myocardial regions

Limited data reporting ECV values was available and only ECV values measured at the RV and LV free walls using a 1.5 T Siemens scanner could be pooled (Table 3 and (Figure 7.3)). Dong 2018 evaluated ECV values with a 3 T Siemens system and reported a value of 29.3% $\pm$ 4.9 at the septum and 38.5% $\pm$ 3.9 at the RVIP.

7.4 Discussion

In this systematic review and meta-analysis, we demonstrate a significant rise in myocardial T_1 values in patients with PAH when compared to healthy controls scanned under the same conditions, confirming potential diagnostic value in measuring T_1 mapping in PAH. The myocardial region with the largest difference between PAH and healthy controls was the RVIP with a mean difference of 108 ms, 95% (CI 89 to 128). The mean difference at the septum was 63 ms, 95% CI (40 to 86) and at the LV free wall 40 ms, 95% CI (20 to 60). The pooled T_1 value at the RVIP on 1.5 Tesla Siemens system and MOLLI sequence was

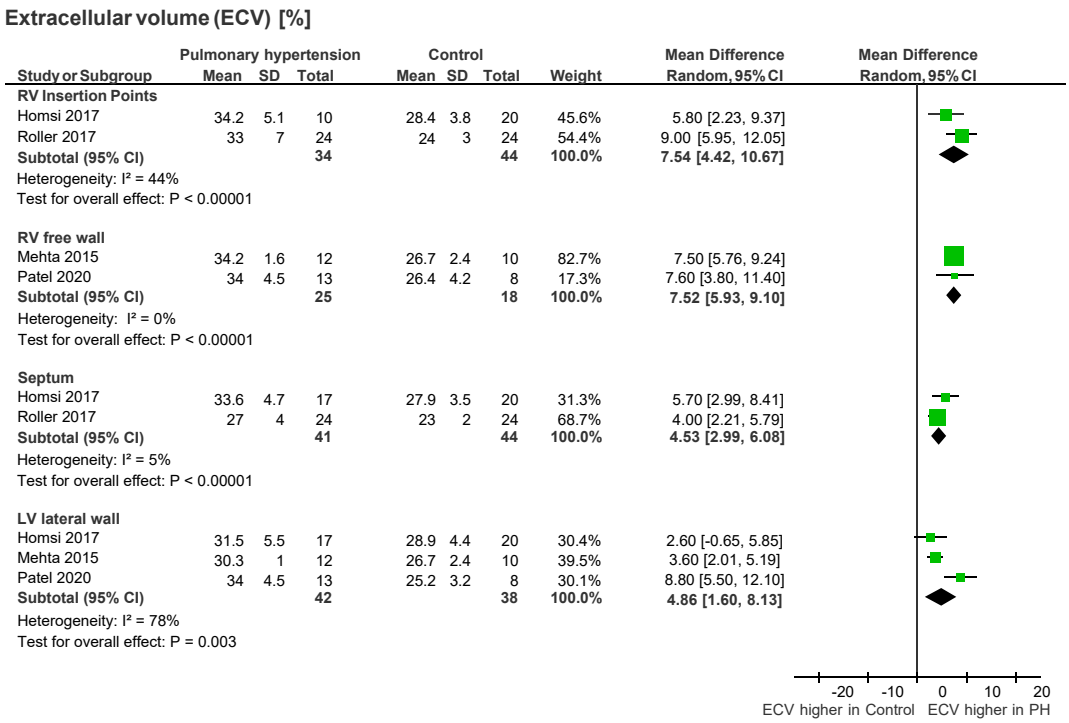


Figure 7.4: The mean difference of ECV in PAH and controls.
CI, confidence intervals; ECV, extracellular volume; ms, millisecond; LV, left ventricle; PH, pulmonary hypertension; RV, right ventricle; SD, standard deviation

1,084 ms, 95% CI (1,071 to 1,097), which is on average 9% higher than in healthy people. Therefore the RVIP should be used to measure T_1 mapping values in suspected PAH. The pooled normal T_1 values at the septum on 1.5T Siemens scanners in our meta-analysis were 988 ms, 95% CI (978 to 998) which are similar to the values reported in a large meta-analysis of normal T_1 values of 977 ms, 95% CI (969 to 985) [286]. Limited data exists on ECV values in PAH. The highest mean difference of ECV in PAH compared to healthy people was at the RV free wall and measured 7.5%, 95% CI (5.9 to 9.1) on 1.5 Tesla Siemens systems.

The T_1 values in PAH and the mean difference in T_1 values between PH and healthy controls is higher than what is reported in cardiomyopathies [288, 289]. The mean difference in septal T_1 values were 45 ms, 95% CI (31 to 60) in dilated cardiomyopathies and 47 ms, 95% CI (33 to 62) in hypertrophic cardiomyopathies compared to healthy controls [288]. Reiter 2017 assessed the difference between PH patients and patients with cardiomyopathies which is more realistically seen in a clinical setting [87]. They found that at 3T there remained a significant difference between the patient groups (105 ms) at

the insertion points. Saunders 2018 found the differences smaller on a 1.5T scanner. The T_1 values at the RVIP were 1065 ± 86 ms in PH patients compared to 1017 ± 69 ms in non-PH patients and 943 ± 52 ms in healthy volunteers [84]. Myocardial T_1 mapping might therefore not be able to differentiate between PAH and other cardiac abnormalities at 1.5T but it would indicate an underlying pathological process increasing the mechanical strain on the RV. The elevated T_1 values at the insertion point in particular might represent engorgement of extracellular spaces in the early phases [281] or fibrosis in more advanced stages of PAH [76].

We pooled the results of different subtypes of PAH including idiopathic PAH, PAH secondary to connective tissue disease (CTD) and congenital heart disease (CHD). CTD and CHD are known to generate fibrosis in the RV and septum the current studies [295–298] and particularly CTD has shown elevated T_1 values [299]. However, three included studies reporting different subtypes of PAH showed no significant differences with the T_1 mapping values seen in idiopathic PAH compared to PAH secondary to CTD or CHD [67, 84, 291]. Therefore, pooling the results of the subgroups of PAH was considered appropriate.

T_1 values [84, 86] and ECV [90, 294] significantly correlated to RV function and volumes. This is in agreement with similar findings in nonischaemic cardiomyopathies, which showed an association between elevated T_1 at the RVIP and RV dysfunction [300]. Deterioration in RV function is associated with a poor prognosis in PAH [4, 46]. However, Saunders 2018 found that myocardial T_1 mapping did not predict mortality. During a median follow-up of 27 months, they reported 59 deaths from 369 included patients. Their univariate Cox regression showed that RV ejection fraction, end-diastolic and end-systolic RV volumes, RV mass index and the septal angle were prognostic markers, but not T_1 values [84].

Areas with late gadolinium enhancement (LGE) showed a high T_1 value and ECV in Homsi 2017 and were associated with a significantly impaired RV function [67]. LGE at the RVIP is linked to focal fibrosis and more severe disease [76, 77, 296, 301–303]. T_1 values and ECV remained significantly higher in PAH compared to controls even in the absence of LGE, which may indicate that myocardial T_1 mapping is more sensitive than LGE and could serve as an early marker for fibrosis [67]. This observation is in keeping

with findings in patients with PH secondary to CTEPH [85] and findings of a porcine model of chronic PH, where T_1 values and ECV were elevated in areas of fibrosis before the onset of LGE [304].

Almost all included studies used the MOLLI myocardial T_1 mapping sequences, while Mehta 2015 used ANGIE sequences [90]. However, even within MOLLI sequences there are a multitude of sequence parameters. The accuracy and precision of MOLLI sequences are highly dependent on several factors including flip angle, inversion times, recovery times, numbers of inversions, off-resonance, heart rate, spatial resolution, parallel imaging and field strength to mention some [305]. The primary meta-analysis compared PAH patients with controls imaged under the same conditions and therefore any differences in image acquisition does not affect the result of the meta-analysis. However, patients with PAH are more likely to have a higher heart rate than controls. When stroke volume decreases cardiac output is compensated by increased heart rate. The MOLLI sequence is heart rate sensitive owing to the time between inversion and the influence of the readout during each inversion recovery. The effect is that when 'normal heart rate' MOLLI is used in a high heart rate situation the T_1-values can be expected to be falsely lower. Many of the studies pooled in the meta-analysis have T_1-values within the higher limit of normal range, even in the insertion point area. These values could be falsely normal if 'normal heart rate' MOLLI sequences were used.

Repeatability of T_1-values was high in Saunders 2018 and Chen 2017 and was highest at the septum followed by the RVIP. Intraobserver T_1-values varied by up to 20 ms in Chen 2017. The high agreement between readers, confirms that there is good repeatability of T_1 measurement when the same system, parameters and sequences are used [305, 306]. A standardised method of measuring T_1 values might support the myocardial T_1 mapping technique becoming more reliable in the assessment of PH [307]. Deep learning methods for the automated quantification of T_1 mapping and ECV are being developed and have shown good performance compared to manual assessment [308]. However, the repeatability of T_1 mapping between different scanners can vary considerably even when using the same sequences and field strength [309] which can limit its utility as a follow-up tool.

MRI Harmonisation techniques have been proposed to reduce inter-scan and inter-site variability by adjusting MRI data to a correction factor [310]. Several methods have been suggested for estimating correction factors such as using data from matched controls or travelling subjects between sites [311]. The correction covariate applied at a voxel level can bring T_1 values from different sequences into a unified space for a more accurate quantitative comparison of percent differences in T_1 and ECV values between normal and PAH subjects.

RV free wall T_1 mapping was assessed in five studies [88–90, 291, 293]. T_1 mapping at the RV free was significantly higher compared to the LV lateral wall in the same patient and compared to the RV free wall in healthy control. The increase of T_1 values in a thickened RV free wall compared to the LV free wall may reflect the increased pressure and afterload on the RV in PH [300]. However, despite this, assessing the T_1 value at the RV free wall remains challenging. The partial volume effects from the adjacent blood pool or epicardial fat is truly an issue particularly as image slices are currently thicker than the RV wall itself. Sequences with higher spatial resolution might allow better assessment of the thin RV free wall such as the ANGIE sequences used by Mehta 2015 employed [90]. However, technical problems eliciting T_1 values can also be caused by the relative asymmetrical shape of the RV with its curved wall at midventricular level and its myocardial trabeculation [300]. Mehta 2015 tried to overcome this by measuring T_1 values of the RV free wall in end-systole when the RV is at its thickest. While this might facilitate obtaining a T_1 value reading, it is not as accurate or reliable as diastolic T_1 maps and more likely to cause false values. In systole the likelihood of including RV trabeculations and hence blood is increased when the borders of trabeculations/compact myocardium are less distinguishable. Assessing the T_1 value in a thin RV free wall in healthy people is particularly difficult [84, 291]. In Spruijt 2016 it was not feasible to draw regions of interest in the RV free wall in the majority of PAH patients and the healthy controls because the RV free wall was too thin and the partial volume effects too substantial, emphasising the concerns and limited benefit of measuring T_1 values at the RV free wall.

Limitations

The strength of the meta-analysis is the extensive literature search that identified all reported myocardial T_1 mapping and ECV values reported in PAH. The main limitation is that the analysis is based on mainly small retrospective studies. The variation in MRI systems and field strength limited the number of meta-analyses possible. The main analysis pooled the differences between T_1 mapping values in PAH and healthy volunteers measured under the same condition to reduce the effect of variations between T_1 mapping estimation methods across studies. However, pooling the actual T_1 mapping values across studies is prone to heterogeneity from several sources including imaging sites, MRI systems, sequences, techniques and slice selection. ECV was only reported in five studies using different techniques and a mean ECV value could not be calculated. PH is a group of very diverse diseases that vary in the mechanisms of affecting pulmonary artery pressure, resistance and morphology and have different RV remodeling responses. We included studies reporting T_1 mapping mainly in PAH and excluding other PH groups to limit disease heterogeneity. However, the included studies included 6% of patients with CTEPH and 1% of other types of PH which might have introduced some heterogeneity.

7.5 Conclusion

T_1 mapping values in PAH patients are on average 9% higher than healthy controls when assessed under the same conditions including the same MRI system, magnetic field strength or sequence used for acquisition. The highest T_1 and ECV values are at the RV insertion points. T_1 values and ECV in PH are higher than the values reported in cardiomyopathies and were associated with poor RV function and RV dilatation.

7.6 Author Contributions

S.A. and A.J.S. conceived the idea and need for the systematic review and contributed to the study conception and design. Protocol registration in PROSPERO was performed by S.A.. S.A. created the search strategy and performed the literature search. S.A. performed screening and eligibility assessments. S.A. evaluated the included studies and collected relevant data. Material preparation and analysis were performed by S.A.. The manuscript,

figures and tables were drafted by S.A.. All authors contributed to the interpretation of data. The first draft of the manuscript was written by S.A. and all authors commented on previous versions of the manuscript. The final draft was written by S.A., taking into account comments and suggestions from peer reviewers and editors. All authors read and approved the final manuscript. All authors took part in the critical review and drafting of the manuscript and have read and approved the final manuscript.

Discussion

7.7 Knowledge Dissemination and Impact of Thesis

A total of 17 articles and 15 abstracts related to this thesis have been published in peer-reviewed journals and cited over 90 times in the past two years (Appendix A1). The most cited aspect is the meta-analysis of cardiac MRI measurements predicting prognosis in PAH described in Chapter 2, which was included in the European Society of Cardiology (ESC) and the European Respiratory Society (ERS) guidelines for the diagnosis and treatment of pulmonary hypertension [312]. Word underlying this thesis has been presented at 13 international and six national conferences, including the major cardiac imaging and radiology conferences (Appendix A2 and A3). At the Radiological Society of North America (RSNA) and the Royal College of Radiologists (RCR) Global conference, the research was featured in oral presentations and received the prestigious RSNA Trainee Research and the RCR Global Oral Abstract awards, respectively. Additionally, the AI cardiac MRI segmentation tool developed and validated in this research received the Medipex NHS Innovation, the Yorkshire & Humber School of Radiology and the Professor Ronald Grainger awards. The tool has also been featured in national and regional media, such as The Daily Mail, The Star, and the largest medical imaging forum AuntMinnie [313–315]. The research was also invited to be presented on the RSNA Radiology podcast, "Validation of AI Cardiac MRI Measurements." [316]. Finally, work presented in this thesis (Chapter 5) was awarded the British Thoracic Society 2022 Conference Award.

7.8 Future Research Directions

Additional cardiac MRI measurements

Based on the developed segmentation algorithms of the short-axis, four-chamber and long-axis views, there are a multitude of measurements that can be obtained, including cardiac chamber lengths, diameters, areas and myocardial thickness. From these measurements,

parameters such as linear myocardial strain and atrial volumes can be obtained. Linear strain is the length of the chamber in diastole minus its length in systole [317]. Atrial volumes are calculated from the biplanar area-length formula (0.85 x atrial area / atrial length) [318]. Where the area and length are obtained from the four-chamber and two-chamber views. Manual estimates of these measurements have shown prognostic impact in the ability to detect impaired longitudinal shortening post-myocardial infarction [317] and the identification of raised left ventricular filling pressure [319]. A potential research study is to assess the prognostic impact of left and right atrial longitudinal strain in a large cohort of PAH patients.

Right ventricle (RV) myocardial thickness assessment is difficult due to the thin configuration of the normal RV. Chapter 3 has identified no study that has reported automated RV myocardial thickness assessment. Despite the challenges of measuring the RV thickness, it has a role in identifying RV hypertrophy in PAH, which is an important marker of remodelling and prognosis [214, 215]. Our current short-axis segmentation model (Chapter 4) can automatically assess RV thickness throughout the cardiac cycle (Figure 7.5). A study to identify clinically significant thresholds of RV thickness to monitor remodelling in PAH would be of interest.

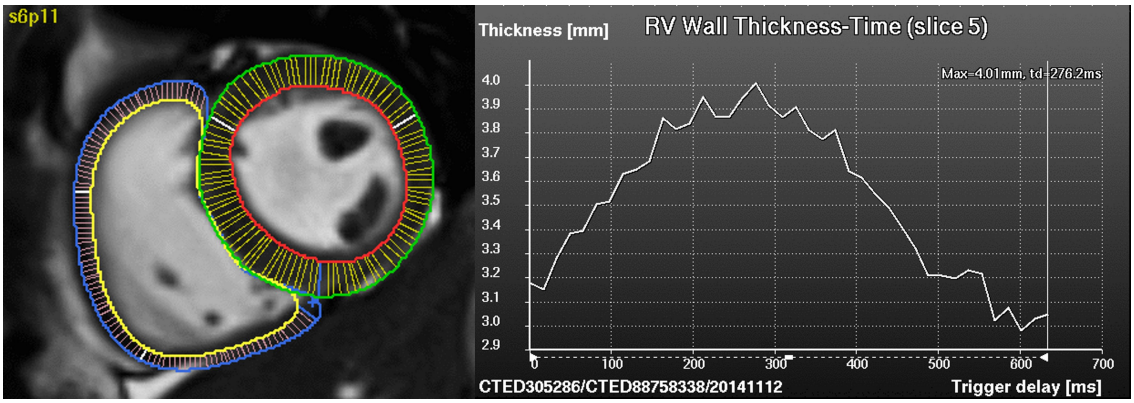


Figure 7.5: RV myocardial thickness assessment (left panel) with measurements performed throughout the cardiac cycle (right panel)

While RV mass is another potential measurement of RV hypertrophy, it is more challenging to perform due to the presence of ventricular papillary muscles. Our current segmentation analysis does not separate these papillary muscles, also called trabeculations,

from the blood pool. While this is an acceptable method of calculating the mass and blood volumes [320], trabecular mass itself might have diagnostic and prognostic implications in pulmonary hypertension [321]. We are currently developing and testing an automated method to detect trabeculations using signal thresholding (Figure 7.6). My aim is to examine the added prognostic value from including trabecular mass to RV mass in patients with PAH. This could be integrated with the automated RV thickness to identify a combined marker of RV hypertrophy.

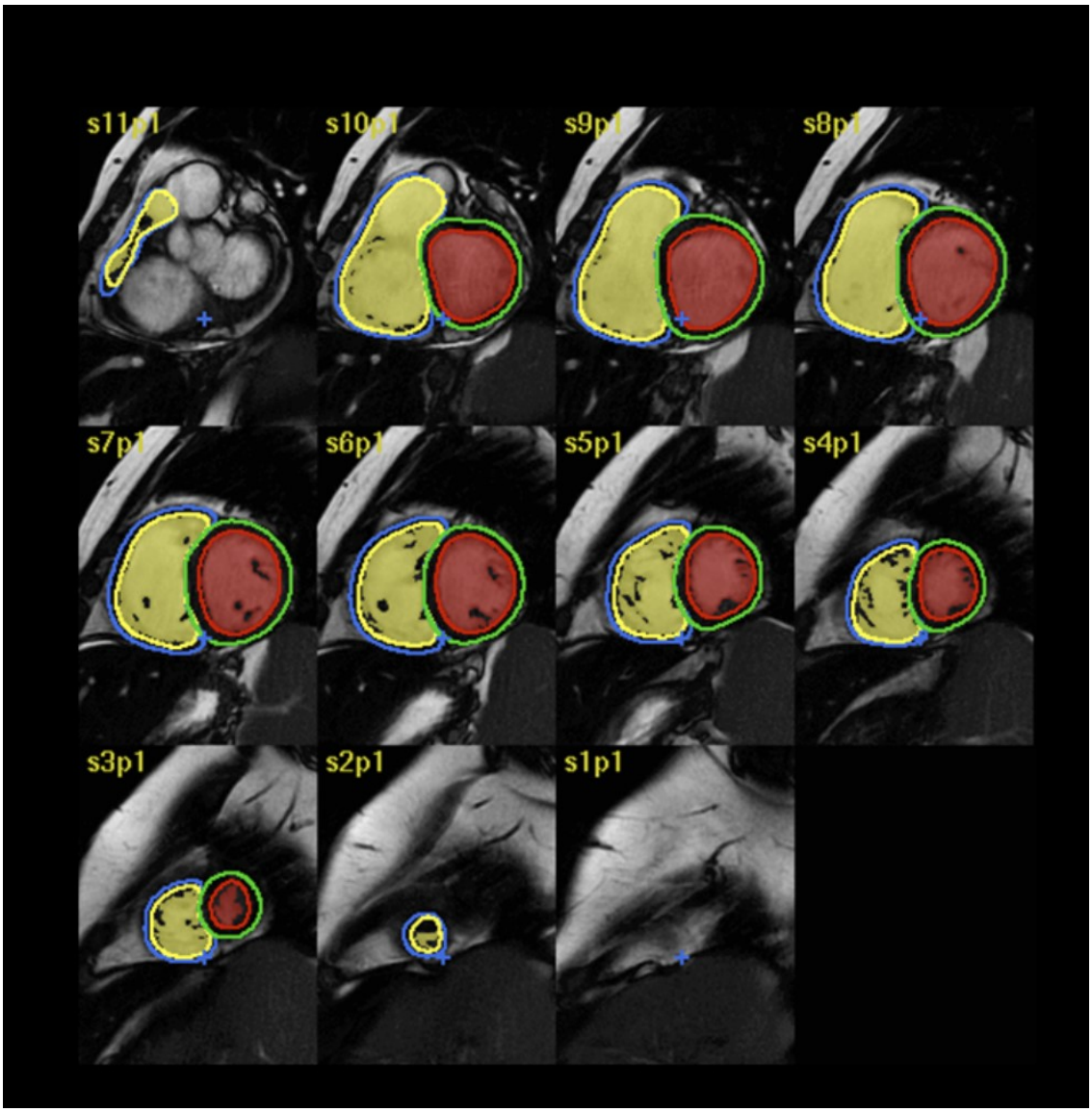


Figure 7.6: Automated segmentation excluding trabeculation from the highlighted right ventricular (yellow) and left ventricular (red) blood pool.

Automation of native myocardial T_1-mapping assessment

Chapter 7 discusses the value of myocardial T_1 mapping in PH. One of the main challenges with T_1 mapping is the high interobserver variability and artefact from the blood pool. In addition, the RV wall remains difficult to assess clinically but is significantly elevated in PH [10]. An automated, standardised and reproducible method that minimises blood pool artefact could improve the utility of T_1-mapping in clinical practice and add value to the risk stratification of patients with PH. Future research developing a model to automatically measure the RV myocardial T_1-values might provide diagnostic and prognostic value. Tensor machine learning applications The multilinear principal component analysis (MPCA) machine learning approach has demonstrated its effectiveness in both diagnostic [9] and prognostic [322] applications in PAH. One of the key advantages of this tensor-based method is its ability to provide clear explanations of the underlying linear machine-learning processes, without the need for segmentation. Future applications of the MPCA machine learning include follow-up and treatment response assessment. For example, in a baseline-follow-up cohort, an MPCA-based mortality prediction algorithm can be used to identify changes in prognostic features after treatment, which may provide valuable insights into which areas of the heart are most responsive to treatment and how this impacts overall outcome.

Other applications that would benefit from the strong classification ability of the MPCA method are predicting haemodynamics on short-axis and four-chamber cine imaging and altered pulmonary artery flow on phase contrast flow images. In a large cohort of PH patients with same or next-day right heart catheterisation (RHC), the MPCA model will be trained to predict normal and elevated pulmonary capillary wedge pressure (PCWP). The added value of the MPCA method can be compared to using validated equations for non-invasive heart pressure prediction [319].

Prospective assessment of AI cardiac MRI measurements

Following retrospective validation, AI interventions must be assessed in prospective clinical studies [323]. Our large retrospective multi-centre and multi-vendor validation of the short-axis [7] and four-chamber [324] cardiac MRI segmentation has shown the accuracy,

repeatability and prognostic impact of automatic measurements. I plan to perform a phase II study using a prospective single (or multi-centre) assessment of the AI segmentation tool in a clinical setting to confirm the performance in real-world scenarios [325]. This pilot data will be used to plan a randomised multi-centre and multi-vendor study. In this phase III study, the diagnostic impact of a battery of automated measurements will be assessed. Currently, a diagnosis of pulmonary hypertension (PH) relies on right heart catheterisation (RHC) and multi-disciplinary team opinion. Cardiac MRI has shown good accuracy in predicting the presence of PH [207] and estimating cardiac haemodynamics [319, 326]. The phase III study will aim to assess whether automated cardiac MRI measurements can replace the need for invasive diagnosis. Patients will be randomised to cardiac MRI assessment with and without AI [327]. Randomisation will be performed in cluster time periods in which scan analysis is performed in alternating turns with and without AI assistance [327]. Previous studies of MRI diagnostic accuracy and confidence have shown an impact on patient outcome and management [328]. Patients with suspected PH will have their cardiac MRI analysed with or without AI measurement and will have a subsequent RHC and MDT. Cardiac MRI readers will rate their diagnostic confidence of a PH diagnosis on five points likert scale, ranging from unsure to confident. The outcome reference diagnosis (ORD) will be made by clinicians blinded to the MRI assessment and based on RHC performed within 24 hours and MDT discussion. Diagnostic accuracy, including sensitivity, specificity, area under the receiver operating characteristic curve and negative and positive predictive values of the cardiac MRI will be compared to the outcome reference diagnosis. The cardiac MRI parameters assessed will include routine volumetric and mass measurements, septal deviation, and pulmonary artery flow [35]. In addition, patient convenience and overall costs will be compared.

Natural language processing creation of radiology reports based on AI measurements

Medical imaging analysis through machine learning can generate a vast amount of data from segmentation alone, including measurements such as length, area, volume, angle, mass, and density for every segmented structure. These numbers can be overwhelming for radiologists and clinicians to process. My proposal is to develop a natural language processing (NLP) algorithm that can analyse combinations of these measurements and

provide a more meaningful interpretation. For example, instead of presenting numerical values, the NLP algorithm would report a cardiac chamber as ”dilated.” To evaluate the performance of the algorithm, scans will be analysed by both a radiologist with NLP support and by a radiologist alone. The NLP algorithm will be randomly assigned to each scan to minimise performance bias. The outcomes of the study will include time savings for radiologists, increased accuracy of diagnosis, and improved report clarity, as evaluated by blinded referring clinicians.

Natural language processing summarising serial radiology reports of multiple modalities.

Reviewing previous imaging is crucial for radiologists to understand a patient's background and changes in their condition. However, with the increasing demand for imaging, including cardiac MRI (Figure 7.7), patients often have multiple serial imaging studies from various modalities. While this trend provides more information for radiologists, it also increases the workload of reviewing prior investigation reports.

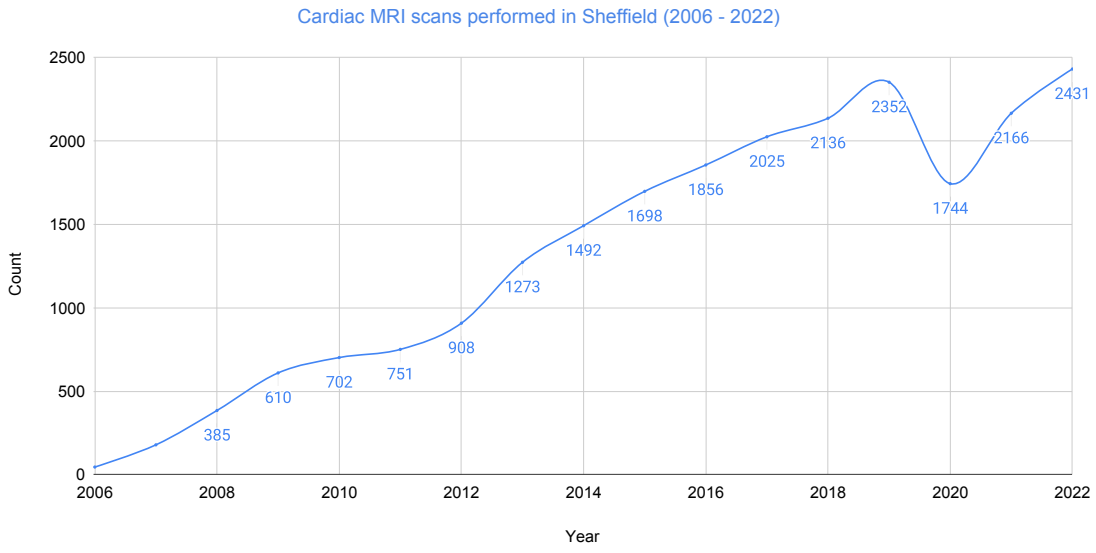


Figure 7.7: The rise in the number of patients requiring cardiac MRI between 2006 and 2022.

Natural language processing (NLP) can be used to analyse large amounts of text and provide summaries in tabular or graphical format. My goal is to develop and validate an NLP algorithm that summarises longitudinal health data from imaging reports.

The algorithm will highlight any previously identified abnormalities and suggest disease diagnoses based on a constellation of findings or commonly associated findings with the previously identified abnormalities. For example, when a patient undergoes a CT pulmonary angiogram, the NLP algorithm can gather important findings from previous radiographs, ultrasound, CT, and MRI scans and create a summary. The NLP program could highlight the presence of previously identified lung nodules and pulmonary emboli, which would alert the radiologist to scrutinise these findings and comment on any changes. The NLP algorithm could also suggest looking for cardiac or vascular changes commonly associated with previous findings, such as chronic pulmonary embolism and right heart abnormalities as a sequela of acute pulmonary embolism. Similar to the previous NLP application, the algorithm will be assessed by two radiologists assessing the scan, one with the help of the NLP algorithm and one without. An additional outcome will be the number of additional findings detected.

In summary, cardiac MRI is becoming increasingly important in pulmonary hypertension diagnosis and prognosis, and AI segmentation can streamline the process by providing multiple measurements at once. Future research will focus on using these measurements to automate imaging reports through natural language processing.

Appendices

APPENDIX A

Publications, Presentations and Awards

A.1 List of Publications

Publications Directly Related To This Thesis

[1] Maiter A., Salehi M., Swift A., and **Alabed** S. *How should studies using AI be reported? Lessons from a systematic review in cardiac MRI*. Vol. 3. 2023.

[2] **Alabed** S., Alandejani F., Dwivedi K., Karunasaagarar K., Sharkey M., Garg P., Koning P. J. H. de, Tóth A., Shahin Y., Johns C., et al. *Validation of Artificial Intelligence Cardiac MRI Measurements: Relationship to Heart Catheterization and Mortality Prediction*. Radiology, 2022.

[3] **Alabed** S., Uthoff J., Zhou S., Garg P., Dwivedi K., Alandejani F., Gosling R., Schobs L., Brook M., Capener D., et al. *Machine Learning cardiac-MRI features predict mortality in newly diagnosed pulmonary arterial hypertension*. European Heart Journal - Digital Health, May 2022.

[4] **Alabed** S., Maiter A., Salehi M., Mahmood A., Daniel S., Jenkins S., Goodlad M., Sharkey M., Mamalakis M., Dwivedi K., et al. *Quality of reporting in AI cardiac MRI segmentation studies - a systematic review and recommendations for future studies*. Vol. 9. Frontiers in Cardiovascular Medicine, 2022.

[5] **Alabed** S., Shahin Y., Alandejani F., Johns C., Lewis R., Condliffe R., Wild J., Kiely D., and Swift A. *Cardiac-MRI Predicts Clinical Worsening and Mortality in Pulmonary Arterial Hypertension: A Systematic Review and Meta-Analysis*. JACC Cardiovascular Imaging, 2021.

[6] **Alabed** S., Saunders L., Garg P., Shahin Y., Rolf A., Puntmann V., Nagel E., Wild J., Kiely D., and Swift A. *Myocardial T1-mapping and extracellular volume in pulmonary arterial hypertension: A systematic review and meta-analysis*. Vol. 79. Magnetic Resonance Imaging, 2021, pp. 66–75.

[7] **Alabed** S., Garg P., Johns C. S., Alandejani F., Shahin Y., Dwivedi K., Wild J., Kiely D., and Swift A. *Cardiac Magnetic Resonance in Pulmonary Hypertension-an Update*. Vol. 13. 12. Current Cardiovascular Imaging Reports, 2020.

[8] Alandejani F., **Alabed** S., Garg P., Goh Z. M., Karunasaagarar K., Sharkey M., Salehi M., Aldabbagh Z., Dwivedi K., Mamalakis M., et al. *Training and clinical testing of artificial intelligence derived right atrial cardiovascular magnetic resonance measurements*. Vol. 24. 1. Journal of Cardiovascular Magnetic Resonance, 2022, p. 25.

[9] Shahin Y., **Alabed** S., Lewis R. A., Johns C., Garg P., Wild J. M., Condliffe R., Swift A. J., Kiely D. G., and al. et. *CMR Measures of Left Atrial Volume Index and Right Ventricular Function Have Prognostic Value in Chronic Thromboembolic Pulmonary Hypertension*. Vol. 9. Frontiers in Medicine, 2022.

[10] Assadi H., **Alabed** S., Maiter A., Salehi M., Li R., Ripley D. P., Van der Geest R. J., Zhong Y., Zhong L., Swift A. J., and Garg P. *The Role of Artificial Intelligence in Predicting Outcomes by Cardiovascular Magnetic Resonance: A Comprehensive Systematic Review*. Vol. 58. 8. Medicina, 2022.

[11] Alandejani F., Hameed A., Tubman E., **Alabed** S., Shahin Y., Lewis R. A., Dwivedi K., Mahmood A., Middleton J., Watson L., et al. *Imaging and Risk Stratification in Pulmonary Arterial Hypertension: Time to Include Right Ventricular Assessment*. Vol. 9. Frontiers in Cardiovascular Medicine, 2022.

[12] Garg P., Gosling R., Swoboda P., Jones R., Rothman A., Wild J. M., Kiely D. G., Condliffe R., **Alabed** S., and Swift A. J. *Cardiac magnetic resonance identifies raised left ventricular filling pressure: prognostic implications*. European Heart Journal, May 2022.

[13] Goh Z. M., Balasubramanian N., **Alabed** S., Dwivedi K., Shahin Y., Rothman A. M. K., Garg P., Lawrie A., Capener D., Thompson A. A. R., et al. *Right ventricular remodelling in pulmonary arterial hypertension predicts treatment response.* Heart, 2022.

[14] Goh Z., **Alabed** S., Rothman A., Garg P., Lawrie A., Thompson R., Condliffe R., Wild J., Kiely D., Swift A., and al. et. *Right Ventricular Adaptation Assessed Using Cardiac Magnetic Resonance Predicts Survival in Pulmonary Arterial Hypertension.* JACC: Cardiovascular Imaging, 2020.

[15] Saunders L., Hughes P., **Alabed** S., Capener D., Marshall H., Vogel-Claussen J., Beek E. R. van, Kiely D., Swift A., and Wild J. *Integrated Cardiopulmonary MRI Assessment of Pulmonary Hypertension.* Journal of Magnetic Resonance Imaging, 2021.

[16] Swift A. J., Wilson F., Cogliano M., Kendall L., Alandejani F., **Alabed** S., Hughes P., Shahin Y., Saunders L., Oram C., et al. *Repeatability and sensitivity to change of non-invasive end points in PAH: the RESPIRE study.* BMJ Thorax, 2021.

[17] Swift A. J., Lu H., Uthoff J., Garg P., Cogliano M., Taylor J., Metherall P., Zhou S., Johns C. S., **Alabed** S., Condliffe R. A., Lawrie A., Wild J. M., and Kiely D. G. *A machine learning cardiac magnetic resonance approach to extract disease features and automate pulmonary arterial hypertension diagnosis.* Vol. 22. European Heart Journal Cardiovascular Imaging.

Publications Not Related To This Thesis

[18] Macdonald A., Salehi M., **Alabed** S., Maiter A., Goh Z. M., Dwivedi K., Johns C., Cogliano M., Alandejani F., Condliffe R., Wild J. M., Kiely D. G., Garg P., and Swift A. J. "Semi-automatic thresholding of RV trabeculation improves repeatability and diagnostic value in suspected pulmonary hypertension". In: *Frontiers in cardiovascular medicine* 9 (Jan. 2023).

[19] Sharkey M. J., Taylor J. C., **Alabed** S., Dwivedi K., Karunasaagarar K., Johns C. S., Rajaram S., Garg P., Alkhanfar D., Metherall P., et al. "Fully automatic cardiac four chamber and great vessel segmentation on CT pulmonary angiography using deep learning". In: *Frontiers in Cardiovascular Medicine* 9 (2022).

[20] Shahin Y., **Alabed** S., Alkhanfar D., Tschirren J., Rothman A. M. K., Condliffe R., Wild J. M., Kiely D. G., and Swift A. J. "Quantitative CT Evaluation of Small Pulmonary Vessels Has Functional and Prognostic Value in Pulmonary Hypertension". In: *Radiology* (2022).

[21] Al Said S., **Alabed** S., Kaier K., Tan A., Bode C., Meerpohl J., and Duerschmied D. "Non-vitamin K antagonist oral anticoagulants (NOACs) post-percutaneous coronary intervention: a network meta-analysis". In: *Cochrane* 12.12 (2019).

[22] Al Said S., Katus H. A., and **Alabed** S. "Cochrane corner: NOACs in atrial fibrillation patients post-percutaneous coronary intervention". In: *Heart* 106.17 (2020), pp. 1293–1295.

[23] Al Said S., Garg P., Jenkins S., Ahmad M., Qintar M., Kyriacou A., Verma N., Providencia R., Camm J., and **Alabed** S. "Catheter ablation for atrial fibrillation (Protocol)". In: *Cochrane Database of Systematic Reviews* 1 (2022).

[24] Njoku P., Grafton-Clarke C., Assadi H., Gosling R., Archer G., Swift A. J., Morris P., **Alabed** S., Flather M., Cameron D., et al. "Validation of time-resolved, automated peak trans-mitral velocity tracking: Two center four-dimensional flow cardiovascular magnetic resonance study". In: *International Journal of Cardiology* (June 2022).

[25] Alkhanfar D., Shahin Y., Alandejani F., Dwivedi K., **Alabed** S., Johns C., Lawrie A., Thompson A. R., Rothman A. M., Tschirren J., et al. "Severe pulmonary hyper-

tension associated with lung disease is characterised by a loss of small pulmonary vessels on quantitative computed tomography". In: *European Respiratory Jornal Open Research* 8.2 (2022).

[26] Dwivedi K., Condliffe R., Sharkey M., Lewis R., **Alabed** S., Rajaram S., Hill C., Saunders L., Metherall P., Lu H., Wild J., Kiely D., and Swift A. "Computed tomography lung parenchymal descriptions in routine radiological reporting have diagnostic and prognostic utility in patients with idiopathic pulmonary arterial hypertension and pulmonary hypertension associated with lung disease". In: *European Respiratory Jornal Open Research* 8.1 (2022).

[27] Lanham S., Maiter A., Swift A. J., Dwivedi K., **Alabed** S., Evans O., Sharkey M. J., Matthews S., and Johns C. S. "The reproducibility of manual RV/LV ratio measurement on CT pulmonary angiography". In: *British Journal of Radiology—Open* eprint (2022). DOI: 10.1259/bjro.20220041.

[28] Dwivedi K., Sharkey M., Condliffe R., Uthoff J., **Alabed** S., Metherall P., Lu H., Wild J., Hoffman E., Swift A., and Kiely D. "Pulmonary Hypertension in Association with Lung Disease: Quantitative CT and Artificial Intelligence to the Rescue? State-of-the-Art Review". In: *Diagnostics* 11.4 (2021).

[29] Jenkins S., **Alabed** S., Swift A., Marques G., Ryding A., Sawh C., Wardley J., Shah B., Swoboda P., Senior R., Nijveldt R., Vassiliou V., and Garg P. "Diagnostic accuracy of handheld cardiac ultrasound device for assessment of left ventricular structure and function: systematic review and meta-analysis". In: *Heart* (2021).

[30] Hocking K., Alhun U., Balian V., Kabuli M., Tse G., Chopra A., Kotnis N., Connelly D., and **Alabed** S. "Acute haemorrhage rate in 28,000 Out-of-Hours CT heads". In: *The British Journal of Radiology* 12 (2021).

[31] Kaur H., Assadi H., **Alabed** S., Vassiliou V. S., Westenberg J. J. M., Geest R. van der, Swift A. J., and Garg P. "Left Ventricular Blood Flow Kinetic Energy Assessment by 4D Flow Cardiovascular Magnetic Resonance: A Systematic Review". In: *Journal of Cardiovascular Development and Disease* 7.3 (2020).

[32] Jones R., Varian F., **Alabed** S., Morris P., Rothman A., Swift A., Wild J., and Garg P. "Meta-analysis of echocardiographic quantification of left ventricular filling pressure". In: *ESC Heart Failure* (2020).

[33] **Alabed** S., Sabouni A., Providencia R., Atallah E., Qintar M., and Chico T. "Adenosine versus intravenous calcium channel antagonists for supraventricular tachycardia". In: *Emergencias* 32.1 (2020).

[34] Davis H., **Alabed** S., and Chico T. "Effect of sports massage on performance and recovery: a systematic review and meta-analysis". In: *BMJ Open Sport Exercise Medicine* 6.1 (2020).

Conference Proceedings

[35] **Alabed** S., Garg P., Dwivedi K., Maiter A., Karunasaagarar K., Rajaram S., Hill C., Thomas S., Wild J., Swift A., and Kiely D. "Establishing minimally important differences for cardiac MRI endpoints in pulmonary arterial hypertension". In: *Thorax*. Vol. 77. 2022.

[36] **Alabed** S., Garg P., Dwivedi K., Maiter A., Karunasaagarar K., Rajaram S., Hill C., Thomas S., Wild J., Swift A., and Kiely D. "Prediction of outcome with cardiac MRI measurements in patients with pulmonary arterial hypertension". In: *The International Journal of Cardiovascular Imaging*. Vol. A-264. 2022.

[37] **Alabed** S., Dwivedi K., Durrington C., Alandajani F., Condliffe R., Elliot C., Charalampopoulos A., Hameed A., Thompson R., Rothman A., Armstrong I., Swift A., and Kiely D. "Correlation of emPHasis-10 with clinical tests: insights from the ASPIRE registry". In: *Thorax*. Vol. 77. 2022.

[38] **Alabed** S., Maiter A., Mahmood A., Daniel S., Salehi M., Jenkins S., Sharkey M., Dwivedi K., Mamalakis M., Assadi H., Garg P., and Swift A. "Quality of reporting of artificial intelligence studies: Lessons learnt from a systematic review of the literature". In: *Clinical Radiology*. Vol. 77. 2022.

[39] **Alabed** S., Maiter A., Mahmood A., Daniel S., Salehi M., Jenkins S., Sharkey M., Rakocevic V., Dwivedi K., Asaadi H., Mamalakis M., O'regan D. P., Garg P., Van Der Geest R., and Swift A. J. "The quality of reporting in cardiac MRI artificial intelligence segmentation studies - a systematic review". In: *European Heart Journal - Cardiovascular Imaging*. Vol. 23. 2022.

[40] **Alabed** S., Salehi M., Mohammad D., Ochieng L., Maiter A., Dwivedi K., Johns C., Hill C., Rajaram S., Tuner D., Thomas S., Swift A. J., and Karunasaagarar K. "Cardiac findings on body CT: a review of 275,000 CT reports over the past 14 years". In: *Heart*. Vol. 108. Suppl 2. 2022.

[41] Alandajani F., **Alabed** S., Garg P., Goh Z., Karunasaagarar K., Sharkey M., Salehi M., Aldabbagh Z., Dwivedi K., Metherall P., et al. "Training and Clinical Validation of Artificial Intelligence Derived Right Atrial Cardiovascular Magnetic Resonance Measurements". In: *American Thoracic Society*. 2022, A2285–A2285.

[42] Gosling R., **Alabed** S., Swoboda P., Nagueh S. F., Jones R., Rothman A., Wild J. M., Kiely D. G., Condliffe R., Swift A. J., and Garg P. "Cardiac magnetic resonance to identify raised left ventricular filling pressure". In: *Heart*. Vol. 107. Suppl 3. 2021.

[43] **Alabed** S., Karunasaagarar K., Garg P., Lu H., Wild J., Kiely D., Van Der Geest R., and Swift A. "A fully automated cardiac magnetic resonance (CMR) assessment improves the evaluation of patients with pulmonary arterial hypertension (PAH)". In: vol. 58. suppl 65. 2021.

[44] **Alabed** S., Karunasaagarar K., Garg P., Uthoff J., Metherall P., Sharkey M., Lu H., Wild J., Kiely D., Van Der Geest R., and Swift A. "Fully automated CMR derived stroke volume correlates with right heart catheter measurements in patients with suspected pulmonary hypertension". In: *European Heart Journal - Cardiovascular Imaging*. Vol. 22. 2021, suppl 2.

[45] **Alabed** S., Karunasaagarar K., Garg P., Uthoff J., Metherall P., Sharkey M., Lu H., Wild J., Kiely D., Van Der Geest R., and Swift A. "High interstudy repeatability of automatic deep learnt biventricular CMR measurements". In: *European Heart Journal - Cardiovascular Imaging*. 2021, suppl 2.

[46] **Alabed** S., Metherall P., Sharkey M., and Swift A. "Deep Learning derived T1-mapping values compared to manual assessment". In: *European Congress of Radiology*. 2021.

[47] **Alabed** S., Wild J., Kiely D., and Swift A. "Cardiac MRI for Prognosis in Pulmonary Arterial Hypertension: A Systematic Review and Meta-Analysis". In: *European Respiratory Journal*. Vol. 56. suppl 64. 2020.

[48] Goh Z., Balasubramanian N., **Alabed** S., Dwivedi K., Shahin Y., Rothman A., Garg P., Lawrie A., Wild J., Johns C., Kiely D., and Swift A. "Right ventricular remodelling assessed using cardiac magnetic resonance predicts survival and treatment response in pulmonary arterial hypertension". In: *Thorax*. Vol. 77. 2022.

[49] Uthoff J., **Alabed** S., Lawrie A., Kiely D., Lu H., and Swift A. "Sex bias exists in diagnosing pulmonary arterial hypertension via machine learning". In: *European Respiratory Journal*. Vol. 56. suppl 64. 2020.

[50] Uthoff J., **Alabed** S., Swift A., and Lu H. "Geodesically Smoothed Tensor Features for Pulmonary Hypertension Prognosis using the Heart and Surrounding Tissues". In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020*. 2020.

[51] **Alabed** S., Hocking K., Alhun U., Wright C., Hughes F., Balian V., Kabuli M., Kotnis N., and Connolly D. "Natural language processing to audit CT Head reports". In: *European Congress of Radiology*. 2021.

[52] Dwivedi K., Lewis R., Condliffe R., Sharkey M., Mamalakis M., **Alabed** S., Wild J., Swift A., and Kiely D. "Computed Tomography (CT) Features Are of Diagnostic Utility in Pre Diagnosis Idiopathic Pulmonary Arterial Hypertension (IPAH): A Case Controlled Study". In: *American Thoracic Society*. 2022, A3605–A3605.

[53] Abdulaal L., Dwivedi K., Sharkey M., **Alabed** S., Mamalakis M., Alkhanfar D., Condliffe R., Kiely D., and Swift A. "CT lung parenchymal appearances in chronic thromboembolic pulmonary hypertension (CTEPH)". In: *Thorax*. Vol. 77. 2022.

[54] Mamalakis M., Dwivedi K., Sharkey M., **Alabed** S., Metherall P., Kiely D., and Swift A. "Deep Learning Approaches to Classify Lung Parenchymal Disease on CT Images". In: *American Thoracic Society*. 2022, A5431–A5431.

[55] Sharkey M., Karunasaagarar K., Johns C., Rajaram S., Alkhanfar D., Dwivedi K., **Alabed** S., Metherall P., Geest R. V. D., Mamalakis M., Kiely D., and Swift A. "Fully Automatic Cardiac and Great Vessel Segmentation on CT Pulmonary Angiography (CTPA) Using Deep Learning". In: *American Thoracic Society*. 2022, A5437–A5437.

[56] Alkhanfar D., Shahin Y., Alandejani F., **Alabed** S., Johns C., Rothman A., Garg P., Condliffe R., Quadery R., Kiely D., Wild J., and Swift A. "Serial cardiac MRI for assessment of cardiac morphology and function in CTEPH patients after PEA or vasodilator therapy". In: vol. 56. suppl 64. European Respiratory Society, 2020.

[57] Alandejani F., Tubman E., Shahin Y., Lewis R., Dwivedi K., Alkhanfar D., **Alabed** S., Johns C., Garg P., Condliffe R., Lawrie A., Kiely D., Wild J., and Swift A. "Cardiac MRI right atrial area measurement thresholds for risk stratification in patients with PAH". In: vol. 56. suppl 64. European Respiratory Society, 2020.

A.2 Oral Presentations

1. **Title:** AI cardiac MRI segmentation studies - Lessons learnt from the literature
   Meeting: Royal College of Radiologists Global Conference, 2022
   Location: Dubai, UAE

2. **Title:** AI Cardiac MRI Measurements Clinical Benchmarking
   Meeting: Radiological Society of North America, 2021
   Location: Chicago, USA

3. **Title:** Using AI to improve pulmonary hypertension assessment
   Meeting: National Pulmonary Hypertension Research Forum, 2021
   Location: London

4. **Title:** The future of cardiac imaging
   Meeting: Yorkshire & Humber Chest and Cardiac Regional Study Day, 2021
   Location: Sheffield

5. **Title:** Clinical Validation of Cardiac MRI AI Segmentation
   Meeting: School of Radiology - Yorkshire & Humber annual meeting
   Meeting: Professor Ronald Grainger Memorial Meeting
   Meeting: Clinical Imaging Clinical Research meeting
   Meeting: Department Research in Progress Meeting
   Meeting: Sheffield Medical School Research Conference
   Location: Sheffield

6. **Title:** Natural Language Processing in Radiology Audit
   Meeting: Yorkshire School of Radiology Annual Conference, 2021
   Location: Virtual

A.3 International Poster Presentations

1. **Title:** Prediction of outcome using cardiac MRI in pulmonary arterial hypertension
   Meeting: European Society of Cardiovascular Radiology 22
   Location: Rome, Italy

2. **Title:** Correlation of emPHasis-10 with cardiac MRI and clinical tests
   Meeting: European Respiratory Society 22
   Location: Barcelona, Spain

3. **Title:** Time-resolved cardiac MRI prgonostic feature extraction
   Meeting: Society for Cardiovascular Magnetic Resonance (SCMR) 22
   Location: Virtual

4. **Title:** Quality of reporting in AI cardiac MRI segmentation studies
   Meeting: SCMR/EACVI 22
   Location: London

5. **Title:** Reporting cardiac findings on body CT over the last decade
   Meeting: European Congress of Radiology 22
   Location: Virtual

6. **Title:** Automated CMR assessment in pulmonary hypertension
   Meeting: European Respiratory Society 21
   Location: Virtual

7. **Title:** High repeatability of deep learnt CMR measurements
   Meeting: EuroCMR - European Association of CardioVascular Imaging 21
   Location: Virtual

8. **Title:** Automated CMR correlates with right heart catheter

Meeting: EuroCMR - European Association of CardioVascular Imaging 21

Location: Virtual

9. **Title:** Deep Learning derived T1-mapping values

   Meeting: European Congress of Radiology 21

   Location: Virtual

10. **Title:** Machine Learning in Cardiac MRI Predicts Mortality

    Meeting: SCMR 21

    Location: Virtual

11. **Title:** Meta-analysis of T_1-mapping in Pulmonary Hypertension

    Meeting: RSNA 20

    Location: Virtual

12. **Title:** Cardiac MRI predicts prognosis in pulmonary hypertension

    Meeting: European Respiratory Society 20

    Location: Virtual

A.4 Awards

1. **Award:** NHS Medipex Innovation Award
   Awarding body: Medipex NHS, 2022
   Research: Clinical implementation of AI cardiac MRI segmentation

2. **Award:** RCR Global Oral Abstract Award
   Awarding body: Royal College of Radiologists, 2022
   Research: AI cardiac MRI segmentation studies - Lessons learnt from the literature

3. **Award:** British Thoracic Society Conference Award
   Awarding body: British Thoracic Society, 2022
   Research: Cardiac MRI thresholds benchmarked to quality of life in pulmonary arterial hypertension

4. **Award:** Research Prize for Cardiac Imaging
   Awarding body: Radiological Society of North America, 2021
   Research: AI Cardiac MRI Measurements Clinical Benchmarking
   Significance: Abstract chosen from 10,000 studies submitted to the largest radiological conference the world. First UK trainee to win award since 2013.

5. **Award:** 1st Prize for Best Oral Presentation
   Awarding body: School of Radiology - Yorkshire & Humber, 2022
   Research: AI Cardiac MRI Measurements Clinical Benchmarking

6. **Award:** 1st Prize - Professor Ronald Grainger Memorial Meeting
   Awarding body: Sheffield Teaching Hospitals, 2021
   Research: AI Cardiac MRI Measurements Clinical Benchmarking

7. **Award:** 1st Prize - Professor Ronald Grainger Memorial Meeting
   Awarding body: Sheffield Teaching Hospitals, 2020

Research: Natural Language Processing in Radiology Audit

8. **Award:** European Congress of Radiology Travel Award

   Awarding body: Royal College of Radiologists, 2022

   Research: Reporting cardiac findings on body CT over the last decade

9. **Award:** The Sir Ernest Finch Travelling Fellowship

   Awarding body: Sheffield Teaching Hospitals, 2021

   Research: Deep Learning derived T1-mapping values

Bibliography

[1] OŔegan D. P. "Putting machine learning into motion: applications in cardiovascular imaging". In: *Clinical radiology* 75.1 (Jan. 2020), pp. 33–37.

[2] **Alabed** S., Garg P., Johns C. S., Alandejani F., Shahin Y., Dwivedi K., Wild J., Kiely D., and Swift A. *Cardiac Magnetic Resonance in Pulmonary Hypertension-an Update*. Vol. 13. 12. Current Cardiovascular Imaging Reports, 2020.

[3] Clarke M., Hopewell S., and Chalmers I. "Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report". In: *Journal of the Royal Society of Medicine* 100.4 (Apr. 2007), pp. 187–190.

[4] **Alabed** S., Shahin Y., Alandejani F., Johns C., Lewis R., Condliffe R., Wild J., Kiely D., and Swift A. *Cardiac-MRI Predicts Clinical Worsening and Mortality in Pulmonary Arterial Hypertension: A Systematic Review and Meta-Analysis*. JACC Cardiovascular Imaging, 2021.

[5] **Alabed** S., Maiter A., Salehi M., Mahmood A., Daniel S., Jenkins S., Goodlad M., Sharkey M., Mamalakis M., Dwivedi K., et al. *Quality of reporting in AI cardiac MRI segmentation studies - a systematic review and recommendations for future studies*. Vol. 9. Frontiers in Cardiovascular Medicine, 2022.

[6] Cooper N. J., Jones D. R., and Sutton A. J. "The use of systematic reviews when designing studies". In: *Clinical Trials* 2.3 (2005), pp. 260–264.

[7] **Alabed** S., Alandejani F., Dwivedi K., Karunasaagarar K., Sharkey M., Garg P., Koning P. J. H. de, Tóth A., Shahin Y., Johns C., et al. *Validation of Artificial Intelligence Cardiac MRI Measurements: Relationship to Heart Catheterization and Mortality Prediction*. Radiology, 2022.

[8] Saunders L., Hughes P., **Alabed** S., Capener D., Marshall H., Vogel-Claussen J., Beek E. R. van, Kiely D., Swift A., and Wild J. *Integrated Cardiopulmonary MRI Assessment of Pulmonary Hypertension*. Journal of Magnetic Resonance Imaging, 2021.

[9] Swift A. J., Lu H., Uthoff J., Garg P., Cogliano M., Taylor J., Metherall P., Zhou S., Johns C. S., **Alabed** S., Condliffe R. A., Lawrie A., Wild J. M., and Kiely D. G. *A machine learning cardiac magnetic resonance approach to extract disease features and automate pulmonary arterial hypertension diagnosis*. Vol. 22. European Heart Journal Cardiovascular Imaging.

[10] **Alabed** S., Saunders L., Garg P., Shahin Y., Rolf A., Puntmann V., Nagel E., Wild J., Kiely D., and Swift A. *Myocardial T1-mapping and extracellular volume in pulmonary arterial hypertension: A systematic review and meta-analysis*. Vol. 79. Magnetic Resonance Imaging, 2021, pp. 66–75.

[11] **Alabed** S., Metherall P., Sharkey M., and Swift A. "Deep Learning derived T1-mapping values compared to manual assessment". In: *European Congress of Radiology*. 2021.

[12] Kiely D. G., Levin D., Hassoun P., Ivy D. D., Jone P.-N., Bwika J., Kawut S. M., Lordan J., Lungu A., Mazurek J., et al. "EXPRESS: Statement on imaging and pulmonary hypertension from the Pulmonary Vascular Research Institute (PVRI)". In: *Pulmonary circulation* 9.3 (Mar. 2019), p. 2045894019841990.

[13] Galiè N., Humbert M., Vachiery J.-L., Gibbs S., Lang I., Torbicki A., Simonneau G., Peacock A., Vonk Noordegraaf A., Beghetti M., et al. "2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT)". In: *The European Respiratory Journal* 46.4 (Jan. 2016), pp. 903–975.

[14] Wijeratne D. T., Lajkosz K., Brogly S. B., Lougheed M. D., Jiang L., Housin A., Barber D., Johnson A., Doliszny K. M., and Archer S. L. "Increasing Incidence

and Prevalence of World Health Organization Groups 1 to 4 Pulmonary Hypertension: A Population-Based Cohort Study in Ontario, Canada". In: *Circulation. Cardiovascular quality and outcomes* 11.2 (Feb. 2018), e003973.

[15] Mathai S. C. and Ryan J. J. "The Growing Burden of Pulmonary Hypertension in the Modern Era: A Zebra No More?" In: *Circulation. Cardiovascular quality and outcomes* 11.2 (Feb. 2018), e004536.

[16] Hoeper M. M., Humbert M., Souza R., Idrees M., Kawut S. M., Sliwa-Hahnle K., Jing Z.-C., and Gibbs J. S. R. "A global view of pulmonary hypertension". In: *The Lancet. Respiratory medicine* 4.4 (Apr. 2016), pp. 306–322.

[17] Ling Y., Johnson M. K., Kiely D. G., Condliffe R., Elliot C. A., Gibbs J. S. R., Howard L. S., Pepke-Zaba J., Sheares K. K. K., Corris P. A., et al. "Changing demographics, epidemiology, and survival of incident pulmonary arterial hypertension: results from the pulmonary hypertension registry of the United Kingdom and Ireland". In: *American journal of respiratory and critical care medicine* 186.8 (Oct. 2012), pp. 790–796.

[18] Rich S., Haworth S. G., Hassoun P. M., and Yacoub M. H. "Pulmonary hypertension: the unaddressed global health burden". In: *The Lancet. Respiratory medicine* 6.8 (Aug. 2018), pp. 577–579.

[19] Gidwani S. and Nair A. "The burden of pulmonary hypertension in resource-limited settings". In: *Global heart* 9.3 (Sept. 2014), pp. 297–310.

[20] Simonneau G., Montani D., Celermajer D. S., Denton C. P., Gatzoulis M. A., Krowka M., Williams P. G., and Souza R. "Haemodynamic definitions and updated clinical classification of pulmonary hypertension". In: *The European Respiratory Journal* 53.1 (Jan. 2019).

[21] Cournand A., Lauson H. D., Bloomfield R. A., Breed E. S., and Baldwin E. D. F. "Recording of Right Heart Pressures in Man". In: *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine* 55.1 (Jan. 1944), pp. 34–36.

[22] Heath D. and Whitaker W. “Hypertensive pulmonary vascular disease”. In: *Circulation* 14.3 (Sept. 1956), pp. 323–343.

[23] Thompson A. A. R., Wilkins M. R., Wild J. M., Kiely D. G., and Lawrie A. “Editorial: Pulmonary Hypertension: Mechanisms and Management, History and Future”. In: *Frontiers of medicine* 7 (Apr. 2020), p. 125.

[24] Hatano S., Strasser T., World Health Organization U., et al. “Primary pulmonary hypertension: report on a WHO meeting, Geneva, 15-17 October 1973”. In: (1975).

[25] Newman J. H. “Pulmonary hypertension”. In: ().

[26] Elliot C. and Kiely D. G. “Pulmonary hypertension: diagnosis and treatment”. In: *Clinical medicine* 4.3 (May 2004), pp. 211–215.

[27] Benza R. L., Miller D. P., Barst R. J., Badesch D. B., Frost A. E., and McGoon M. D. “An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL Registry”. In: *Chest* 142.2 (Aug. 2012), pp. 448–456.

[28] Schulthess G. K. von, Fisher M. R., and Higgins C. B. “Pathologic blood flow in pulmonary vascular disease as shown by gated magnetic resonance imaging”. In: *Annals of internal medicine* 103.3 (Sept. 1985), pp. 317–323.

[29] “National Audit of Pulmonary Hypertension, 10th Annual Report - NHS Digital”. In: (). Accessed: 2020-7-9.

[30] Kiely D. G., Elliot C. A., Sabroe I., and Condliffe R. “Pulmonary hypertension: diagnosis and management”. In: *BMJ* 346 (Apr. 2013), f2028.

[31] Sun W. and Chan S. Y. “Pulmonary Arterial Stiffness: An Early and Pervasive Driver of Pulmonary Arterial Hypertension”. In: *Frontiers of medicine* 5 (July 2018), p. 204.

[32] McGoon M. D., Benza R. L., Escribano-Subias P., Jiang X., Miller D. P., Peacock A. J., Pepke-Zaba J., Pulido T., Rich S., Rosenkranz S., Suissa S., and Humbert M. “Pulmonary arterial hypertension: epidemiology and registries”. In: *Journal of the American College of Cardiology* 62.25 Suppl (Dec. 2013), pp. D51–9.

[33] Johns C. S., Wild J. M., Rajaram S., Swift A. J., and Kiely D. G. "Current and emerging imaging techniques in the diagnosis and assessment of pulmonary hypertension". In: *Expert review of respiratory medicine* 12.2 (Feb. 2018), pp. 145–160.

[34] Hur D. J. and Sugeng L. "Non-invasive Multimodality Cardiovascular Imaging of the Right Heart and Pulmonary Circulation in Pulmonary Hypertension". In: *Frontiers in cardiovascular medicine* 6 (Mar. 2019), p. 24.

[35] Swift A. J., Wild J. M., Nagle S. K., Roldán-Alzate A., François C. J., Fain S., Johnson K., Capener D., Beek E. J. R. van, Kiely D. G., Wang K., and Schiebler M. L. "Quantitative magnetic resonance imaging of pulmonary hypertension: a practical approach to the current state of the art". In: *Journal of thoracic imaging* 29.2 (Mar. 2014), pp. 68–79.

[36] Johns C. S., Kiely D. G., Rajaram S., Hill C., Thomas S., Karunasaagarar K., Garg P., Hamilton N., Solanki R., Capener D. A., et al. "Diagnosis of Pulmonary Hypertension with Cardiac MRI: Derivation and Validation of Regression Models". In: *Radiology* 290.1 (Jan. 2019), pp. 61–68.

[37] Whitfield A. J., Solanki R., Johns C. S., Kiely D., Wild J., and Swift A. J. "MRI Prediction of Precapillary Pulmonary Hypertension according to the Sixth World Symposium on Pulmonary Hypertension". In: *Radiology* 294.2 (Feb. 2020), p. 482.

[38] Johns C. S., Rajaram S., Capener D. A., Oram C., Elliot C., Condliffe R., Kiely D. G., Wild J. M., and Swift A. J. "Non-invasive methods for estimating mPAP in COPD using cardiovascular magnetic resonance imaging". In: *European Radiology* 28.4 (Apr. 2018), pp. 1438–1448.

[39] Chin M., Johns C., Currie B. J., Weatherley N., Hill C., Elliot C., Rajaram S., Wild J. M., Condliffe R., Bianchi S., Kiely D. G., and Swift A. J. "Pulmonary Artery Size in Interstitial Lung Disease and Pulmonary Hypertension: Association with Interstitial Lung Disease Severity and Diagnostic Utility". In: *Frontiers in cardiovascular medicine* 5 (June 2018), p. 53.

[40] Meyer G. M. B., Spilimbergo F. B., Altmayer S., Pacini G. S., Zanon M., Watte G., Marchiori E., and Hochhegger B. "Correction to: Multiparametric Magnetic Resonance Imaging in the Assessment of Pulmonary Hypertension: Initial Experience of a One-Stop Study". In: *Lung* 196.4 (Aug. 2018), p. 497.

[41] Knight D. S., Kotecha T., Martinez-Naharro A., Brown J. T., Bertelli M., Fontana M., Muthurangu V., and Coghlan J. G. "Cardiovascular magnetic resonance-guided right heart catheterization in a conventional CMR environment – predictors of procedure success and duration in pulmonary artery hypertension". In: *Journal of cardiovascular magnetic resonance* 21.1 (Sept. 2019), p. 57.

[42] Rogers T., Ratnayaka K., Khan J. M., Stine A., Schenke W. H., Grant L. P., Mazal J. R., Grant E. K., Campbell-Washburn A., Hansen M. S., et al. "CMR fluoroscopy right heart catheterization for cardiac output and pulmonary vascular resistance: results in 102 patients". In: *Journal of cardiovascular magnetic resonance* 19.1 (July 2017), p. 54.

[43] Ratnayaka K., Kanter J. P., Faranesh A. Z., Grant E. K., Olivieri L. J., Cross R. R., Cronin I. F., Hamann K. S., Campbell-Washburn A. E., OBérien K. J., Rogers T., Hansen M. S., and Lederman R. J. "Radiation-free CMR diagnostic heart catheterization in children". In: *Journal of cardiovascular magnetic resonance* 19.1 (Sept. 2017), p. 65.

[44] Lau E. M. T., Giannoulatou E., Celermajer D. S., and Humbert M. "Epidemiology and treatment of pulmonary arterial hypertension". In: *Nature reviews. Cardiology* 14.10 (Oct. 2017), pp. 603–614.

[45] Ling Y., Johnson M., Kiely D. G., Condliffe R., Elliot C. A., Gibbs S., Howard L., Pepke-Zaba J., Sheares K., Corris P., et al. "Survival Of Incident Cases Of Idiopathic, Heritable And Anorexigen-Associated Pulmonary Arterial Hypertension. Results From Pulmonary Hypertension Registry Of The United Kingdom And Ireland (Proud Registry)". In: *D34. Pulmonary Hypertension: Endpoints And Outcomes*. American Thoracic Society International Conference Abstracts. May 2011, A5952–a5952.

[46] Swift A. J., Capener D., Johns C., Hamilton N., Rothman A., Elliot C., Condliffe R., Charalampopoulos A., Rajaram S., Lawrie A., Campbell M. J., Wild J. M., and Kiely D. G. "Magnetic Resonance Imaging in the Prognostic Evaluation of Patients with Pulmonary Arterial Hypertension". In: *American journal of respiratory and critical care medicine* 196.2 (July 2017), pp. 228–239.

[47] Veerdonk M. C. van de, Kind T., Marcus J. T., Mauritz G.-J., Heymans M. W., Bogaard H.-J., Boonstra A., Marques K. M. J., Westerhof N., and Vonk-Noordegraaf A. "Progressive right ventricular dysfunction in patients with pulmonary arterial hypertension responding to therapy". In: *Journal of the American College of Cardiology* 58.24 (Dec. 2011), pp. 2511–2519.

[48] Veerdonk M. C. van de, Marcus J. T., Westerhof N., Man F. S. de, Boonstra A., Heymans M. W., Bogaard H.-J., and Vonk Noordegraaf A. "Signs of right ventricular deterioration in clinically stable patients with pulmonary arterial hypertension". In: *Chest* 147.4 (Apr. 2015), pp. 1063–1071.

[49] Ray J. C., Burger C., Mergo P., Safford R., Blackshear J., Austin C., Fairweather D., Heckman M. G., Zeiger T., Dubin M., and Shapiro B. "Pulmonary arterial stiffness assessed by cardiovascular magnetic resonance imaging is a predictor of mild pulmonary arterial hypertension". In: *The international journal of cardiovascular imaging* 35.10 (Oct. 2019), pp. 1881–1892.

[50] Blyth K. G., Bellofiore A., Jayasekera G., Foster J. E., Steedman T., Chesler N. C., and Peacock A. J. "Dobutamine stress MRI in pulmonary hypertension: relationships between stress pulmonary artery relative area change, RV performance, and 10-year survival". In: *Pulmonary circulation* 7.2 (Apr. 2017), pp. 465–475.

[51] Abe N., Kato M., Kono M., Fujieda Y., Ohira H., Tsujino I., Oyama-Manabe N., Oku K., Bohgaki T., Yasuda S., and Atsumi T. "Right ventricular dimension index by cardiac magnetic resonance for prognostication in connective tissue diseases and pulmonary hypertension". In: *Rheumatology* 59.3 (Mar. 2020), pp. 622–633.

[52] Simpson C. E., Damico R. L., Kolb T. M., Mathai S. C., Khair R. M., Sato T., Bourji K., Tedford R. J., Zimmerman S. L., and Hassoun P. M. "Ventricular mass

as a prognostic imaging biomarker in incident pulmonary arterial hypertension". In: *The European Respiratory Journal* 53.4 (Apr. 2019).

[53] Brewis M. J., Bellofiore A., Vanderpool R. R., Chesler N. C., Johnson M. K., Naeije R., and Peacock A. J. "Imaging right ventricular function to predict outcome in pulmonary arterial hypertension". In: *International journal of cardiology* 218 (Sept. 2016), pp. 206–211.

[54] Mercurio V., Mukherjee M., Tedford R. J., Zamanian R. T., Khair R. M., Sato T., Minai O. A., Torres F., Girgis R. E., Chin K., Damico R., Kolb T. M., Mathai S. C., and Hassoun P. M. "Improvement in Right Ventricular Strain with Ambrisentan and Tadalafil Upfront Therapy in Scleroderma-associated Pulmonary Arterial Hypertension". In: *American journal of respiratory and critical care medicine* 197.3 (Feb. 2018), pp. 388–391.

[55] Hassoun P. M., Zamanian R. T., Damico R., Lechtzin N., Khair R., Kolb T. M., Tedford R. J., Hulme O. L., Housten T., Pisanello C., et al. "Ambrisentan and Tadalafil Up-front Combination Therapy in Scleroderma-associated Pulmonary Arterial Hypertension". In: *American journal of respiratory and critical care medicine* 192.9 (Nov. 2015), pp. 1102–1110.

[56] Johns C. S., Wild J. M., Rajaram S., Tubman E., Capener D., Elliot C., Condliffe R., Charalampopoulos A., Kiely D. G., and Swift A. J. "Identifying At-Risk Patients with Combined Pre- and Postcapillary Pulmonary Hypertension Using Interventricular Septal Angle at Cardiac MRI". In: *Radiology* 289.1 (Oct. 2018), pp. 61–68.

[57] Karakus G., Kammerlander A. A., Aschauer S., Marzluf B. A., Zotter-Tufaro C., Bachmann A., Degirmencioglu A., Duca F., Babayev J., Pfaffenberger S., Bonderman D., and Mascherbauer J. "Pulmonary artery to aorta ratio for the detection of pulmonary hypertension: cardiovascular magnetic resonance and invasive hemodynamics in heart failure with preserved ejection fraction". In: *Journal of cardiovascular magnetic resonance* 17.79 (Aug. 2015), p. 79.

[58] Dong Y., Sun J., Yang D., He J., Cheng W., Wan K., Liu H., Greiser A., Zhou X., Han Y., and Chen Y. "Right ventricular septomarginal trabeculation hypertrophy is

associated with disease severity in patients with pulmonary arterial hypertension". In: *The international journal of cardiovascular imaging* 34.9 (Sept. 2018), pp. 1439–1449.

[59] Lewis R. A., Johns C. S., Cogliano M., Capener D., Tubman E., Elliot C. A., Charalampopoulos A., Sabroe I., Thompson A. A. R., Billings C. G., et al. "Identification of Cardiac Magnetic Resonance Imaging Thresholds for Risk Stratification in Pulmonary Arterial Hypertension". In: *American journal of respiratory and critical care medicine* 201.4 (Feb. 2020), pp. 458–468.

[60] Lewis R. A., Thompson A. A. R., Billings C. G., Charalampopoulos A., Elliot C. A., Hamilton N., Hill C., Hurdman J., Rajaram S., Sabroe I., Swift A. J., Kiely D. G., and Condliffe R. "Mild parenchymal lung disease and/or low diffusion capacity impacts survival and treatment response in patients diagnosed with idiopathic pulmonary arterial hypertension". In: *The European Respiratory Journal* 55.6 (June 2020).

[61] Peacock A. J., Ling Y., Johnson M. K., Kiely D. G., Condliffe R., Elliot C. A., Gibbs J. S. R., Howard L. S., Pepke-Zaba J., Sheares K. K. K., et al. "Idiopathic pulmonary arterial hypertension and co-existing lung disease: is this a new phenotype?" In: *Pulmonary circulation* 10.1 (Jan. 2020), p. 2045894020914851.

[62] Almutairi H. M., Boubertakh R., Miquel M. E., and Petersen S. E. "Myocardial deformation assessment using cardiovascular magnetic resonance-feature tracking technique". In: *The British journal of radiology* 90.1080 (Dec. 2017), p. 20170072.

[63] Pedrizzetti G., Claus P., Kilner P. J., and Nagel E. "Principles of cardiovascular magnetic resonance feature tracking and echocardiographic speckle tracking for informed clinical use". In: *Journal of cardiovascular magnetic resonance* 18.1 (Aug. 2016), p. 51.

[64] Vo H. Q., Marwick T. H., and Negishi K. "MRI-Derived Myocardial Strain Measures in Normal Subjects". In: *JACC. Cardiovascular imaging* 11.2 Pt 1 (Feb. 2018), pp. 196–205.

[65] Lin A. C. W., Seale H., Hamilton-Craig C., Morris N. R., and Strugnell W. "Quantification of biventricular strain and assessment of ventriculo-ventricular interaction in pulmonary arterial hypertension using exercise cardiac magnetic resonance imaging and myocardial feature tracking". In: *Journal of magnetic resonance imaging: JMRI* 49.5 (May 2019), pp. 1427–1436.

[66] Kallianos K., Brooks G. C., Mukai K., Seguro de Carvalho F., Liu J., Naeger D. M., De Marco T., and Ordovas K. G. "Cardiac Magnetic Resonance Evaluation of Left Ventricular Myocardial Strain in Pulmonary Hypertension". In: *Academic radiology* 25.1 (Jan. 2018), pp. 129–135.

[67] Homsi R., Luetkens J. A., Skowasch D., Pizarro C., Sprinkart A. M., Gieseke J., Meyer Zur Heide Gen Meyer-Arend J., Schild H. H., and Naehle C. P. "Left Ventricular Myocardial Fibrosis, Atrophy, and Impaired Contractility in Patients With Pulmonary Arterial Hypertension and a Preserved Left Ventricular Function: A Cardiac Magnetic Resonance Study". In: *Journal of thoracic imaging* 32.1 (Jan. 2017), pp. 36–42.

[68] Siqueira M. E. M. de, Pozo E., Fernandes V. R., Sengupta P. P., Modesto K., Gupta S. S., Barbeito-Caamaño C., Narula J., Fuster V., Caixeta A., and Sanz J. "Characterization and clinical significance of right ventricular mechanics in pulmonary hypertension evaluated with cardiovascular magnetic resonance feature tracking". In: *Journal of cardiovascular magnetic resonance* 18.1 (June 2016), p. 39.

[69] Padervinskienė L., Krivickienė A., Hoppenot D., Miliauskas S., Basevičius A., Nedzelskienė I., Jankauskas A., Šimkus P., and Ereminienė E. "Prognostic Value of Left Ventricular Function and Mechanics in Pulmonary Hypertension: A Pilot Cardiovascular Magnetic Resonance Feature Tracking Study". In: *Medicina* 55.3 (Mar. 2019), p. 73.

[70] Leng S., Dong Y., Wu Y., Zhao X., Ruan W., Zhang G., Allen J. C., Koh A. S., Tan R.-S., Yip J. W., Tan J. L., Chen Y., and Zhong L. "Impaired Cardiovascular Magnetic Resonance-Derived Rapid Semiautomated Right Atrial Longitudinal Strain Is Associated With Decompensated Hemodynamics in Pulmonary Arterial Hypertension". In: *Circulation. Cardiovascular imaging* 12.5 (May 2019), e008582.

[71] Tello K., Dalmer A., Vanderpool R., Ghofrani H. A., Naeije R., Roller F., Seeger W., Wiegand M., Gall H., and Richter M. J. "Right ventricular function correlates of right atrial strain in pulmonary hypertension: a combined cardiac magnetic resonance and conductance catheter study". In: *American journal of physiology. Heart and circulatory physiology* 318.1 (Jan. 2020), H156–H164.

[72] Ferdian E., Suinesiaputra A., Fung K., Aung N., Lukaschuk E., Barutcu A., Maclean E., Paiva J., Piechnik S. K., Neubauer S., Petersen S. E., and Young A. A. "Fully Automated Myocardial Strain Estimation from Cardiovascular MRI–tagged Images Using a Deep Learning Framework in the UK Biobank". In: *Radiology. Cardiothoracic imaging* 2.1 (Feb. 2020), e190032.

[73] Croisille P., Revel D., and Saeed M. "Contrast agents and cardiac MR imaging of myocardial ischemia: from bench to bedside". In: *European Radiology* 16.9 (Sept. 2006), pp. 1951–1963.

[74] Gulati A., Jabbour A., Ismail T. F., Guha K., Khwaja J., Raza S., Morarji K., Brown T. D. H., Ismail N. A., Dweck M. R., et al. "Association of fibrosis with mortality and sudden cardiac death in patients with nonischemic dilated cardiomyopathy". In: *JAMA: the journal of the American Medical Association* 309.9 (Mar. 2013), pp. 896–908.

[75] Abouelnour A. E., Doyle M., Thompson D. V., Yamrozik J., Williams R. B., Shah M. B., Soma S. K., Murali S., Benza R. L., and Biederman R. W. "Does Late Gadolinium Enhancement still have Value? Right Ventricular Internal Mechanical Work, Ea/Emax and Late Gadolinium Enhancement as Prognostic Markers in Patients with Advanced Pulmonary Hypertension via Cardiac MRI". In: *Cardiology research and cardiovascular medicine* 2017.1 (Jan. 2017).

[76] Swift A. J., Rajaram S., Capener D., Elliot C., Condliffe R., Wild J. M., and Kiely D. G. "LGE patterns in pulmonary hypertension do not impact overall mortality". In: *JACC. Cardiovascular imaging* 7.12 (Dec. 2014), pp. 1209–1217.

[77] Freed B. H., Gomberg-Maitland M., Chandra S., Mor-Avi V., Rich S., Archer S. L., Jamison Jr E. B., Lang R. M., and Patel A. R. "Late gadolinium enhancement cardiovascular magnetic resonance predicts clinical worsening in patients with

pulmonary hypertension". In: *Journal of cardiovascular magnetic resonance* 14.11 (Feb. 2012), p. 11.

[78] Haaf P., Garg P., Messroghli D. R., Broadbent D. A., Greenwood J. P., and Plein S. "Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: a comprehensive review". In: *Journal of cardiovascular magnetic resonance* 18.1 (Nov. 2016), p. 89.

[79] Puntmann V. O., Peker E., Chandrashekhar Y., and Nagel E. "T1 Mapping in Characterizing Myocardial Disease: A Comprehensive Review". In: *Circulation research* 119.2 (July 2016), pp. 277–299.

[80] Messroghli D. R., Plein S., Higgins D. M., Walters K., Jones T. R., Ridgway J. P., and Sivananthan M. U. "Human myocardium: single-breath-hold MR T1 mapping with high spatial resolution–reproducibility study". In: *Radiology* 238.3 (Mar. 2006), pp. 1004–1012.

[81] Garg P., Saunders L. C., Swift A. J., Wild J. M., and Plein S. "Role of cardiac T1 mapping and extracellular volume in the assessment of myocardial infarction". In: *Anatolian journal of cardiology* 19.6 (June 2018), pp. 404–411.

[82] Bull S., White S. K., Piechnik S. K., Flett A. S., Ferreira V. M., Loudon M., Francis J. M., Karamitsos T. D., Prendergast B. D., Robson M. D., Neubauer S., Moon J. C., and Myerson S. G. "Human non-contrast T1 values and correlation with histology in diffuse fibrosis". In: *Heart* 99.13 (July 2013), pp. 932–937.

[83] Kuruvilla S., Janardhanan R., Antkowiak P., Keeley E. C., Adenaw N., Brooks J., Epstein F. H., Kramer C. M., and Salerno M. "Increased extracellular volume and altered mechanics are associated with LVH in hypertensive heart disease, not hypertension alone". In: *JACC. Cardiovascular imaging* 8.2 (Feb. 2015), pp. 172–180.

[84] Saunders L. C., Johns C. S., Stewart N. J., Oram C. J. E., Capener D. A., Puntmann V. O., Elliot C. A., Condliffe R. C., Kiely D. G., Graves M. J., Wild J. M., and Swift A. J. "Diagnostic and prognostic significance of cardiovascular magnetic

resonance native myocardial T1 mapping in patients with pulmonary hypertension". In: *Journal of cardiovascular magnetic resonance* 20.1 (Dec. 2018), p. 78.

[85] Roller F. C., Wiedenroth C., Breithecker A., Liebetrau C., Mayer E., Schneider C., Rolf A., Hamm C., and Krombach G. A. "Native T1 mapping and extracellular volume fraction measurement for assessment of right ventricular insertion point and septal fibrosis in chronic thromboembolic pulmonary hypertension". In: *European Radiology* 27.5 (May 2017), pp. 1980–1991.

[86] Chen Y. Y., Yun H., Jin H., Kong D. H., Long Y. L., Fu C. X., Yang S., and Zeng M. S. "Association of native T1 times with biventricular function and hemodynamics in precapillary pulmonary hypertension". In: *The international journal of cardiovascular imaging* 33.8 (Aug. 2017), pp. 1179–1189.

[87] Reiter U., Reiter G., Kovacs G., Adelsmayr G., Greiser A., Olschewski H., and Fuchsjäger M. "Native myocardial T1 mapping in pulmonary hypertension: correlations with cardiac function and hemodynamics". In: *European Radiology* 27.1 (Jan. 2017), pp. 157–166.

[88] Wang J., Zhao H., Wang Y., Herrmann H. C., Witschey W. R. T., and Han Y. "Native T1 and T2 mapping by cardiovascular magnetic resonance imaging in pressure overloaded left and right heart diseases". In: *Journal of thoracic disease* 10.5 (May 2018), pp. 2968–2975.

[89] Patel R. B., Li E., Benefield B. C., Swat S. A., Polsinelli V. B., Carr J. C., Shah S. J., Markl M., Collins J. D., and Freed B. H. "Diffuse right ventricular fibrosis in heart failure with preserved ejection fraction and pulmonary hypertension". In: *ESC heart failure* 7.1 (Feb. 2020), pp. 253–263.

[90] Mehta B. B., Auger D. A., Gonzalez J. A., Workman V., Chen X., Chow K., Stump C. J., Mazimba S., Kennedy J. L. W., Gay E., Salerno M., Kramer C. M., Epstein F. H., and Bilchick K. C. "Detection of elevated right ventricular extracellular volume in pulmonary hypertension using Accelerated and Navigator-Gated Look-Locker Imaging for Cardiac T1 Estimation (ANGIE) cardiovascular magnetic resonance". In: *Journal of cardiovascular magnetic resonance* 17.110 (Dec. 2015), p. 110.

[91] Schoenfeld C., Hinrichs J. B., Olsson K. M., Kuettner M.-A., Renne J., Kaireit T., Czerner C., Wacker F., Hoeper M. M., Meyer B. C., and Vogel-Claussen J. "Cardio-pulmonary MRI for detection of treatment response after a single BPA treatment session in CTEPH patients". In: *European Radiology* 29.4 (Apr. 2019), pp. 1693–1702.

[92] Maschke S. K., Schoenfeld C. O., Kaireit T. F., Cebotari S., Olsson K., Hoeper M., Wacker F., and Vogel-Claussen J. "MRI-derived Regional Biventricular Function in Patients with Chronic Thromboembolic Pulmonary Hypertension Before and After Pulmonary Endarterectomy". In: *Academic radiology* 25.12 (Dec. 2018), pp. 1540–1547.

[93] Roller F. C., Kriechbaum S., Breithecker A., Liebetrau C., Haas M., Schneider C., Rolf A., Guth S., Mayer E., Hamm C., Krombach G. A., and Wiedenroth C. B. "Correlation of native T1 mapping with right ventricular function and pulmonary haemodynamics in patients with chronic thromboembolic pulmonary hypertension before and after balloon pulmonary angioplasty". In: *European Radiology* 29.3 (Mar. 2019), pp. 1565–1573.

[94] Johns C. S., Swift A. J., Hughes P. J. C., Ohno Y., Schiebler M., and Wild J. M. "Pulmonary MR angiography and perfusion imaging-A review of methods and applications". In: *European Journal of Radiology* 86 (Jan. 2017), pp. 361–370.

[95] Johns C. S., Swift A. J., Rajaram S., Hughes P. J. C., Capener D. J., Kiely D. G., and Wild J. M. "Lung perfusion: MRI vs. SPECT for screening in suspected chronic thromboembolic pulmonary hypertension". In: *Journal of magnetic resonance imaging: JMRI* 46.6 (Dec. 2017), pp. 1693–1697.

[96] Pöhler G. H., Klimes F., Voskrebenzev A., Behrendt L., Czerner C., Gutberlet M., Cebotari S., Ius F., Fegbeutel C., Schoenfeld C., et al. "Chronic Thromboembolic Pulmonary Hypertension Perioperative Monitoring Using Phase-Resolved Functional Lung (PREFUL)-MRI". In: *Journal of magnetic resonance imaging* (2020).

[97] Geest R. J. van der and Garg P. “Advanced Analysis Techniques for Intra-cardiac Flow Evaluation from 4D Flow MRI”. In: *Current radiology reports* 4 (May 2016), p. 38.

[98] Reiter G., Reiter U., Kovacs G., Kainz B., Schmidt K., Maier R., Olschewski H., and Rienmueller R. “Magnetic resonance-derived 3-dimensional blood flow patterns in the main pulmonary artery as a marker of pulmonary hypertension and a measure of elevated mean pulmonary arterial pressure”. In: *Circulation. Cardiovascular imaging* 1.1 (July 2008), pp. 23–30.

[99] Sieren M. M., Berlin C., Oechtering T. H., Hunold P., Drömann D., Barkhausen J., and Frydrychowicz A. “Comparison of 4D Flow MRI to 2D Flow MRI in the pulmonary arteries in healthy volunteers and patients with pulmonary hypertension”. In: *PloS one* 14.10 (Oct. 2019), e0224121.

[100] Reiter U., Reiter G., Kovacs G., Stalder A. F., Gulsun M. A., Greiser A., Olschewski H., and Fuchsjäger M. “Evaluation of elevated mean pulmonary arterial pressure based on magnetic resonance 4D velocity mapping: comparison of visualization techniques”. In: *PloS one* 8.12 (Dec. 2013), e82212.

[101] Reiter G., Reiter U., Kovacs G., Olschewski H., and Fuchsjäger M. “Blood flow vortices along the main pulmonary artery measured with MR imaging for diagnosis of pulmonary hypertension”. In: *Radiology* 275.1 (Apr. 2015), pp. 71–79.

[102] Wang Z., Lakes R. S., Golob M., Eickhoff J. C., and Chesler N. C. “Changes in large pulmonary arterial viscoelasticity in chronic pulmonary hypertension”. In: *PloS one* 8.11 (Nov. 2013), e78569.

[103] Barker A. J., Roldán-Alzate A., Entezari P., Shah S. J., Chesler N. C., Wieben O., Markl M., and François C. J. “Four-dimensional flow assessment of pulmonary artery flow and wall shear stress in adult pulmonary arterial hypertension: results from two institutions”. In: *Magnetic resonance in medicine* 73.5 (May 2015), pp. 1904–1913.

[104] Feneis J. F., Kyubwa E., Atianzar K., Cheng J. Y., Alley M. T., Vasanawala S. S., Demaria A. N., and Hsiao A. “4D flow MRI quantification of mitral and tricuspid

regurgitation: Reproducibility and consistency relative to conventional MRI". In: *Journal of magnetic resonance imaging: JMRI* 48.4 (Oct. 2018), pp. 1147–1158.

[105] Driessen M. M. P., Schings M. A., Sieswerda G. T., Doevendans P. A., Hulzebos E. H., Post M. C., Snijder R. J., Westenberg J. J. M., Dijk A. P. J. van, Meijboom F. J., and Leiner T. "Tricuspid flow and regurgitation in congenital heart disease and pulmonary hypertension: comparison of 4D flow cardiovascular magnetic resonance and echocardiography". In: *Journal of cardiovascular magnetic resonance* 20.1 (Jan. 2018), p. 5.

[106] Fenster B. E., Browning J., Schroeder J. D., Schafer M., Podgorski C. A., Smyser J., Silveira L. J., Buckner J. K., and Hertzberg J. R. "Vorticity is a marker of right ventricular diastolic dysfunction". In: *American journal of physiology. Heart and circulatory physiology* 309.6 (Sept. 2015), H1087–93.

[107] Barker N., Fidock B., Johns C. S., Kaur H., Archer G., Rajaram S., Hill C., Thomas S., Karunasaagarar K., Capener D., et al. "A Systematic Review of Right Ventricular Diastolic Assessment by 4D Flow CMR". In: *BioMed research international* 2019 (Mar. 2019), p. 6074984.

[108] Grossfeld B. "Deep learning vs machine learning:" in: (Jan. 2020). Accessed: 2020-4-9.

[109] Leiner T., Rueckert D., Suinesiaputra A., Baeßler B., Nezafat R., Išgum I., and Young A. A. "Machine learning in cardiovascular magnetic resonance: basic concepts and applications". In: *Journal of cardiovascular magnetic resonance* 21.1 (Oct. 2019), p. 61.

[110] LeCun Y., Bengio Y., and Hinton G. "Deep learning". In: *Nature* 521.7553 (May 2015), pp. 436–444.

[111] Chartrand G., Cheng P. M., Vorontsov E., Drozdzal M., Turcotte S., Pal C. J., Kadoury S., and Tang A. "Deep Learning: A Primer for Radiologists". In: *Radiographics: a review publication of the Radiological Society of North America, Inc* 37.7 (Nov. 2017), pp. 2113–2131.

[112] Avendi M. R., Kheradvar A., and Jafarkhani H. “Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach”. In: *Magnetic resonance in medicine* 78.6 (Dec. 2017), pp. 2439–2448.

[113] Bai W., Sinclair M., Tarroni G., Oktay O., Rajchl M., Vaillant G., Lee A. M., Aung N., Lukaschuk E., Sanghvi M. M., et al. “Automated cardiovascular magnetic resonance image analysis with fully convolutional networks”. In: *Journal of cardiovascular magnetic resonance* 20.1 (Sept. 2018), p. 65.

[114] Duan J., Bello G., Schlemper J., Bai W., Dawes T. J. W., Biffi C., Marvao A. de, Doumoud G., OŔegan D. P., and Rueckert D. “Automatic 3D Bi-Ventricular Segmentation of Cardiac Images by a Shape-Refined Multi- Task Deep Learning Approach”. In: *IEEE transactions on medical imaging* 38.9 (Sept. 2019), pp. 2151–2164.

[115] Lungu A., Swift A. J., Capener D., Kiely D., Hose R., and Wild J. M. “Diagnosis of pulmonary hypertension from magnetic resonance imaging-based computational models and decision tree analysis”. In: *Pulmonary circulation* 6.2 (June 2016), pp. 181–190.

[116] Dawes T. J. W., Marvao A. de, Shi W., Fletcher T., Watson G. M. J., Wharton J., Rhodes C. J., Howard L. S. G. E., Gibbs J. S. R., Rueckert D., Cook S. A., Wilkins M. R., and O’Regan D. P. “Machine Learning of Three-dimensional Right Ventricular Motion Enables Outcome Prediction in Pulmonary Hypertension: A Cardiac MR Imaging Study”. In: *Radiology* 283.2 (May 2017), pp. 381–390.

[117] Dawes T. J. W., Cai J., Quinlan M., Marvao A. de, Ostrowski P. J., Tokarczuk P. F., Watson G. M. J., Wharton J., Howard L. S. G. E., Gibbs J. S. R., Cook S. A., Wilkins M. R., and OŔegan D. P. “Fractal Analysis of Right Ventricular Trabeculae in Pulmonary Hypertension”. In: *Radiology* 288.2 (Aug. 2018), pp. 386–395.

[118] Swift A. J., Cogliano M., Oram C., Kendall L., Capener D., Garg P., Johns C., Macdonald A., Wilson F., Cahn T., Lawrie A., Condliffe R., Kiely D., and Wild J. “Repeatability and Sensitivity to change of right ventricular analysis methods using cardiac magnetic resonance imaging in PAH: results from the RESPIRE Study”. In: *The European Respiratory Journal* 54.suppl 63 (Sept. 2019).

[119] Noordegraaf A. V., Channick R., Cottreel E., Kiely D., Martin N., Moiseeva O., Peacock A., Tawakol A., Torbicki A., Rosenkranz S., and Galiè N. "Results from the REPAIR Study Final Analysis: Effects of Macitentan on Right Ventricular (RV) Remodelling in Pulmonary Arterial Hypertension (PAH)". In: *The Journal of heart and lung transplantation* 39.4, Supplement (Apr. 2020), S16–S17.

[120] Thenappan T. "Beta-blockers in Pulmonary Arterial Hypertension". In: ().

[121] Solomon M. A. "Spironolactone for Pulmonary Arterial Hypertension". In: ().

[122] Danoff T. "PRIMEx - A Study of 2 Doses of Oral CXA-10 in Pulmonary Arterial Hypertension". In: ().

[123] Ventetuolo C. "Effects of DHEA in Pulmonary Hypertension (EDIPHY)". In: ().

[124] Yacoub M. H. "Effects of Treprostinil on Right Ventricular Structure and Function in Patients With Pulmonary Arterial Hypertension". In: ().

[125] Chaouat A., Cherifi A., Sitbon O., Girerd N., Zysman M., Faure M., Mandry D., Mercy M., Guillaumot A., Fay R., Marie P.-Y., and Chabot F. "[Evaluation of cardiac MRI in the follow up assessment of patients with pulmonary arterial hypertension]". In: *Revue des maladies respiratoires* 35.7 (Sept. 2018), pp. 749–758.

[126] Vogel-Claussen J. "CTEPH DIAGNOSIS Europe - MRI". In: ().

[127] Galiè N., Barberà J. A., Frost A. E., Ghofrani H.-A., Hoeper M. M., McLaughlin V. V., Peacock A. J., Simonneau G., Vachiery J.-L., Grünig E., et al. "Initial Use of Ambrisentan plus Tadalafil in Pulmonary Arterial Hypertension". In: *The New England journal of medicine* 373.9 (Aug. 2015), pp. 834–844.

[128] Baggen V. J. M., Leiner T., Post M. C., Dijk A. P. van, Roos-Hesselink J. W., Boersma E., Habets J., and Sieswerda G. T. "Cardiac magnetic resonance findings predicting mortality in patients with pulmonary arterial hypertension: a systematic review and meta-analysis". In: *European Radiology* 26.11 (Nov. 2016), pp. 3771–3780.

[129] Swift A. J., Rajaram S., Campbell M. J., Hurdman J., Thomas S., Capener D., Elliot C., Condliffe R., Wild J. M., and Kiely D. G. "Prognostic value of cardiovascular magnetic resonance imaging measurements corrected for age and sex in idiopathic

pulmonary arterial hypertension". In: *Circulation. Cardiovascular imaging* 7.1 (Jan. 2014), pp. 100–106.

[130] Moher D., Liberati A., Tetzlaff J., Altman D. G., and PRISMA Group. "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement". In: *International journal of surgery* 6.7 (July 2009), e1000097.

[131] Hayden J. A., Côté P., and Bombardier C. "Evaluation of the quality of prognosis studies in systematic reviews". In: *Annals of internal medicine* 144.6 (Mar. 2006), pp. 427–437.

[132] Wan X., Wang W., Liu J., and Tong T. "Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range". In: *BMC medical research methodology* 14.135 (Dec. 2014), p. 135.

[133] Deeks J. J., Higgins J. P. T., and Altman D. G. "Analysing Data and Undertaking Meta-Analyses". In: *Cochrane Handbook for Systematic Reviews of Interventions*. Sept. 2008, pp. 243–296.

[134] Badagliacca R., Poscia R., Pezzuto B., Papa S., Pesce F., Manzi G., Giannetta E., Raineri C., Schina M., Sciomer S., Parola D., Francone M., Carbone I., Fedele F., and Vizza C. D. "Right ventricular concentric hypertrophy and clinical worsening in idiopathic pulmonary arterial hypertension". In: *The Journal of heart and lung transplantation* 35.11 (Nov. 2016), pp. 1321–1329.

[135] Bredfelt A., Radegran G., Hesselstrand R., Arheden H., and Ostenfeld E. "Increased right atrial volume measured with cardiac magnetic resonance is associated with worse clinical outcome in patients with pre-capillary pulmonary hypertension." In: *ESC heart failure* 5.5 (Oct. 2018), pp. 864–875.

[136] Gan C. T.-J., Lankhaar J.-W., Westerhof N., Marcus J. T., Becker A., Twisk J. W. R., Boonstra A., Postmus P. E., and Vonk-Noordegraaf A. "Noninvasively assessed pulmonary artery stiffness predicts mortality in pulmonary arterial hypertension". In: *Chest* 132.6 (Dec. 2007), pp. 1906–1912.

[137] Grapsa J., Tan T. C., Nunes M. C. P., OŔegan D. P., Durighel G., Howard L. S. G. E., Gibbs J. S. R., and Nihoyannopoulos P. "Prognostic impact of right ventricular

mass change in patients with idiopathic pulmonary arterial hypertension". In: *International journal of cardiology* 304 (Apr. 2020), pp. 172–174.

[138] Jose A., Kher A., ODéonnell R. E., and Elwing J. M. "Cardiac magnetic resonance imaging as a prognostic biomarker in treatment-naıęve pulmonary hypertension". In: *European Journal of Radiology* 123 (Feb. 2020), p. 108784.

[139] Kang K.-W., Chang H.-J., Yoo Y. P., Yoon H. S., Kim Y.-J., Choi B.-W., Shim C.-Y., Ha J., and Chung N. "Cardiac magnetic resonance-derived right ventricular outflow tract systolic flow acceleration: a novel index of right ventricular function and prognosis in patients with pulmonary arterial hypertension". In: *The international journal of cardiovascular imaging* 29.8 (Dec. 2013), pp. 1759–1767.

[140] Knight D. S., Steeden J. A., Moledina S., Jones A., Coghlan J. G., and Muthurangu V. "Left ventricular diastolic dysfunction in pulmonary hypertension predicts functional capacity and clinical worsening: a tissue phase mapping study". In: *Journal of cardiovascular magnetic resonance* 17.116 (Dec. 2015), p. 116.

[141] Li W., Yang T., Zhang Y., Gu Q., Liu Z.-H., Ni X.-H., Luo Q., Xiong C.-M., and He J.-G. "Prognostic value of right ventricular ejection/filling parameters in IPAH using cardiac magnetic resonance: A prospective pilot study". In: *Respirology* 22.1 (Jan. 2017), pp. 172–178.

[142] Mouratoglou S. A., Kallifatidis A., Pitsiou G., Grosomanidis V., Kamperidis V., Chalikias G., Kristo D., Tziakas D., Konstantinides S., Hadjimiltiades S., Karvounis H., and Giannakoulas G. "Duration of interventricular septal shift toward the left ventricle is associated with poor clinical outcome in precapillary pulmonary hypertension: A cardiac magnetic resonance study". In: *Hellenic journal of cardiology: HJC = Hellenike kardiologike epitheorese* 61.2 (Mar. 2020), pp. 112–117.

[143] Sato T., Tsujino I., Ohira H., Oyama-Manabe N., Ito Y. M., Yamada A., Ikeda D., Watanabe T., and Nishimura M. "Right atrial volume and reservoir function are novel independent predictors of clinical worsening in patients with pulmonary hypertension". In: *The Journal of heart and lung transplantation* 34.3 (Mar. 2015), pp. 414–423.

[144] Wang L., Chen X., Wan K., Gong C., Li W., Xu Y., Wang J., He J., Wen B., Han Y., Zeng R., and Chen Y. "Diagnostic and prognostic value of right ventricular eccentricity index in pulmonary artery hypertension". In: *Pulmonary circulation* 10.2 (Apr. 2020), p. 2045894019899778.

[145] Wolferen S. A. van, Marcus J. T., Boonstra A., Marques K. M. J., Bronzwaer J. G. F., Spreeuwenberg M. D., Postmus P. E., and Vonk-Noordegraaf A. "Prognostic value of right ventricular mass, volume, and function in idiopathic pulmonary arterial hypertension". In: *European heart journal* 28.10 (May 2007), pp. 1250–1257.

[146] Yamada Y., Okuda S., Kataoka M., Tanimoto A., Tamura Y., Abe T., Okamura T., Fukuda K., Satoh T., and Kuribayashi S. "Prognostic value of cardiac magnetic resonance imaging for idiopathic pulmonary arterial hypertension before initiating intravenous prostacyclin therapy". In: *Circulation journal* 76.7 (Apr. 2012), pp. 1737–1743.

[147] Sica G. T. "Bias in research studies". In: *Radiology* 238.3 (Mar. 2006), pp. 780–789.

[148] Hagger D., Condliffe R., Woodhouse N., Elliot C. A., Armstrong I. J., Davies C., Hill C., Akil M., Wild J. M., and Kiely D. G. "Ventricular mass index correlates with pulmonary artery pressure and predicts survival in suspected systemic sclerosis-associated pulmonary arterial hypertension". In: *Rheumatology* 48.9 (Sept. 2009), pp. 1137–1142.

[149] McLaughlin V. V., Hoeper M. M., Channick R. N., Chin K. M., Delcroix M., Gaine S., Ghofrani H.-A., Jansa P., Lang I. M., Mehta S., et al. "Pulmonary Arterial Hypertension-Related Morbidity Is Prognostic for Mortality". In: *Journal of the American College of Cardiology* 71.7 (Feb. 2018), pp. 752–763.

[150] Sitbon O., Gomberg-Maitland M., Granton J., Lewis M. I., Mathai S. C., Rainisio M., Stockbridge N. L., Wilkins M. R., Zamanian R. T., and Rubin L. J. "Clinical trial design and new therapies for pulmonary arterial hypertension". In: *The European Respiratory Journal* 53.1 (Jan. 2019).

[151] Galiè N., Simonneau G., Barst R. J., Badesch D., and Rubin L. “Clinical worsening in trials of pulmonary arterial hypertension: results and implications”. In: *Current opinion in pulmonary medicine* 16 Suppl 1 (May 2010), S11–9.

[152] Frost A. E., Badesch D. B., Miller D. P., Benza R. L., Meltzer L. A., and McGoon M. D. “Evaluation of the predictive value of a clinical worsening definition using 2-year outcomes in patients with pulmonary arterial hypertension: a REVEAL Registry analysis”. In: *Chest* 144.5 (Nov. 2013), pp. 1521–1529.

[153] Vonk Noordegraaf A., Westerhof B. E., and Westerhof N. “The Relationship Between the Right Ventricle and its Load in Pulmonary Hypertension”. In: *Journal of the American College of Cardiology* 69.2 (Jan. 2017), pp. 236–243.

[154] Badagliacca R., Poscia R., Pezzuto B., Nocioni M., Mezzapesa M., Francone M., Giannetta E., Papa S., Gambardella C., Sciomer S., Volterrani M., Fedele F., and Dario Vizza C. “Right ventricular remodeling in idiopathic pulmonary arterial hypertension: adaptive versus maladaptive morphology”. In: *The Journal of heart and lung transplantation* 34.3 (Mar. 2015), pp. 395–403.

[155] Darsaklis K., Dickson M. E., Cornwell 3rd W., Ayers C. R., Torres F., Chin K. M., and Matulevicius S. “Right atrial emptying fraction non-invasively predicts mortality in pulmonary hypertension”. In: *The international journal of cardiovascular imaging* 32.7 (July 2016), pp. 1121–1130.

[156] Benza R. L., Miller D. P., Gomberg-Maitland M., Frantz R. P., Foreman A. J., Coffey C. S., Frost A., Barst R. J., Badesch D. B., Elliott C. G., Liou T. G., and McGoon M. D. “Predicting survival in pulmonary arterial hypertension: insights from the Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management (REVEAL)”. In: *Circulation* 122.2 (July 2010), pp. 164–172.

[157] Humbert M., Sitbon O., Yaıęci A., Montani D., OĆallaghan D. S., Jaıęs X., Parent F., Savale L., Natali D., Günther S., et al. “Survival in incident and prevalent cohorts of patients with pulmonary arterial hypertension”. In: *The European Respiratory Journal* 36.3 (Sept. 2010), pp. 549–555.

[158] Reardon S. “Rise of Robot Radiologists”. In: *Nature* 576.7787 (Dec. 2019), S54–S58.

[159] Chen, Chen C., Qin C., Qiu H., Tarroni G., Duan J., Bai W., and Rueckert D. "Deep Learning for Cardiac Image Segmentation: A Review". In: *Frontiers in cardiovascular medicine* 7 (2020).

[160] Hosny A., Parmar C., Quackenbush J., Schwartz L. H., and Aerts H. J. W. L. "Artificial intelligence in radiology". In: *Nature reviews. Cancer* 18.8 (Aug. 2018), pp. 500–510.

[161] Mongan J., Moy L., and Kahn Jr C. E. "Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers". In: *Radiology. Artificial intelligence* 2.2 (Mar. 2020), e200029.

[162] Page M. J., McKenzie J. E., Bossuyt P. M., Boutron I., Hoffmann T. C., Mulrow C. D., Shamseer L., Tetzlaff J. M., Akl E. A., Brennan S. E., et al. "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews". In: *BMJ* 372 (Mar. 2021), n71.

[163] Alandejani F., **Alabed** S., Garg P., Goh Z. M., Karunasaagarar K., Sharkey M., Salehi M., Aldabbagh Z., Dwivedi K., Mamalakis M., et al. "Training and clinical testing of artificial intelligence derived right atrial cardiovascular magnetic resonance measurements". In: *Journal of Cardiovascular Magnetic Resonance* 24.1 (2022).

[164] Fonseca C. G., Backhaus M., Bluemke D. A., Britten R. D., Chung J. D., Cowan B. R., Dinov I. D., Finn J. P., Hunter P. J., Kadish A. H., et al. "The Cardiac Atlas Project—an imaging database for computational modeling and statistical atlases of the heart". In: *Bioinformatics* 27.16 (July 2011), pp. 2288–2295.

[165] Chalmers I. and Glasziou P. "Avoidable waste in the production and reporting of research evidence". In: *Obstetrics and gynecology* 114.6 (Dec. 2009), pp. 1341–1345.

[166] Glasziou P. and Chalmers I. "Research waste is still a scandal—an essay by Paul Glasziou and Iain Chalmers". In: *BMJ* 363 (Nov. 2018).

[167] Andreopoulos A. and Tsotsos J. K. "Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI". In: *Medical Image Analysis* 12.3 (June 2008), pp. 335–357.

[168] Radau P., Lu Y., Connelly K., Paul G., Dick A., and Wright G. "Evaluation framework for algorithms segmenting short axis cardiac MRI". In: *The MIDAS Journal-Cardiac MR Left Ventricle Segmentation Challenge* 49 (2009).

[169] Suinesiaputra A., Cowan B. R., Al-Agamy A. O., Elattar M. A., Ayache N., Fahmy A. S., Khalifa A. M., Medrano-Gracia P., Jolly M.-P., Kadish A. H., Lee D. C., Margeta J., Warfield S. K., and Young A. A. "A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images". In: *Medical Image Analysis* 18.1 (Jan. 2014), pp. 50–62.

[170] Karim R., Bhagirath P., Claus P., James Housden R., Chen Z., Karimaghaloo Z., Sohn H.-M., Lara Rodrıéguez L., Vera S., Albà X., et al. "Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images". In: *Medical Image Analysis* 30 (May 2016), pp. 95–107.

[171] Petitjean C., Zuluaga M. A., Bai W., Dacher J.-N., Grosgeorge D., Caudron J., Ruan S., Ayed I. B., Cardoso M. J., Chen H.-C., et al. "Right ventricle segmentation from cardiac MRI: a collation study". In: *Medical Image Analysis* 19.1 (Jan. 2015), pp. 187–202.

[172] Karim R., James Housden R., Balasubramaniam M., Chen Z., Perry D., Uddin A., Al-Beyatti Y., Palkhi E., Acheampong P., Obom S., et al. "Evaluation of current algorithms for segmentation of scar tissue from late Gadolinium enhancement cardiovascular magnetic resonance of the left atrium: an open-access grand challenge". In: *Journal of Cardiovascular Magnetic Resonance* 15.1 (2013).

[173] Tobon-Gomez C., Geers A. J., Peters J., Weese J., Pinto K., Karim R., Ammar M., Daoudi A., Margeta J., Sandoval Z., et al. "Benchmark for Algorithms Segmenting the Left Atrium From 3D CT and MRI Datasets". In: *IEEE transactions on medical imaging* 34.7 (July 2015), pp. 1460–1473.

[174] Karim R., Blake L.-E., Inoue J., Tao Q., Jia S., James Housden R., Bhagirath P., Duval J.-L., Varela M., Behar J. M., et al. "Algorithms for left atrial wall segmentation and thickness – Evaluation on an open-source CT and MRI image database". In: *Medical Image Analysis* 50 (2018), pp. 36–53.

[175] Pace D. F., Dalca A. V., Geva T., Powell A. J., Moghari M. H., and Golland P. "Interactive Whole-Heart Segmentation in Congenital Heart Disease". In: *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 9351 (Oct. 2015), pp. 80–88.

[176] Zhuang X. and Shen J. "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI". In: *Medical Image Analysis* 31 (July 2016), pp. 77–87.

[177] Bernard O., Lalande A., Zotti C., Cervenansky F., Yang X., Heng P.-A., Cetin I., Lekadir K., Camara O., Gonzalez Ballester M. A., et al. "Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?" In: *IEEE transactions on medical imaging* 37.11 (Nov. 2018), pp. 2514–2525.

[178] Xiong Z., Xia Q., Hu Z., Huang N., Bian C., Zheng Y., Vesal S., Ravikumar N., Maier A., Yang X., et al. "A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging". In: *Medical Image Analysis* 67 (Jan. 2021), p. 101832.

[179] Xue W., Li J., Hu Z., Kerfoot E., Clough J., Oksuz I., Xu H., Grau V., Guo F., Ng M., et al. "Left Ventricle Quantification Challenge: A Comprehensive Comparison and Evaluation of Segmentation and Regression for Mid-Ventricular Short-Axis Cardiac MR Data". In: *IEEE Journal of Biomedical and Health Informatics* 25.9 (2021), pp. 3541–3553.

[180] Corral Acero J., Xu H., Zacur E., Schneider J. E., Lamata P., Bueno-Orovio A., and Grau V. "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*. Springer International Publishing, 2020, pp. 384–394.

[181] Zhuang X. "Multivariate Mixture Model for Myocardial Segmentation Combining Multi-Source Images". In: *IEEE transactions on pattern analysis and machine intelligence* 41.12 (Dec. 2019), pp. 2933–2946.

[182] Chen C., Liu Y., Schniter P., Tong M., Zareba K., Simonetti O., Potter L., and Ahmad R. "OCMR (v1.0)–Open-Access Multi-Coil k-Space Dataset for Cardiovascular Magnetic Resonance Imaging". In: (Aug. 2020).

[183] Lalande A., Chen Z., Decourselle T., Qayyum A., Pommier T., Lorgis L., Rosa E. de la, Cochet A., Cottin Y., Ginhac D., Salomon M., Couturier R., and Meriaudeau F. "Emidec: A Database Usable for the Automatic Evaluation of Myocardial Infarction from Delayed-Enhancement Cardiac MRI". In: *Brown University digest of addiction theory and application: DATA* 5.4 (Sept. 2020), p. 89.

[184] Campello V. M., Gkontra P., Izquierdo C., Martin-Isla C., Sojoudi A., Full P. M., Maier-Hein K., Zhang Y., He Z., Ma J., et al. "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge". In: *IEEE transactions on medical imaging* 40.12 (Dec. 2021), pp. 3543–3554.

[185] Moghari M. H., Pace D. F., Contreras H., Ghelani S., Olyaee E., and Arasteh S. T. "Whole-heart and Great Vessel Segmentation from 3D Cardiovascular Magnetic Resonance Images in Congenital Heart Disease (Part II)". In: (Mar. 2021).

[186] Zhuang X., Li L., Wang S., and Wu F. "Left Atrial and Scar Quantification & Segmentation Challenge 2022". In: (Mar. 2022).

[187] Puyol-Antón E., Ruijsink B., Mariscal Harana J., Piechnik S. K., Neubauer S., Petersen S. E., Razavi R., Chowienczyk P., and King A. P. "Fairness in Cardiac Magnetic Resonance Imaging: Assessing Sex and Racial Bias in Deep Learning-Based Segmentation". In: *Frontiers in cardiovascular medicine* 9 (Apr. 2022), p. 859310.

[188] Larrazabal A. J., Nieto N., Peterson V., Milone D. H., and Ferrante E. "Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis". In: *Proceedings of the National Academy of Sciences of the United States of America* 117.23 (June 2020), pp. 12592–12594.

[189] Roberts M., Driggs D., Thorpe M., Gilbey J., Yeung M., Ursprung S., Aviles-Rivero A. I., Etmann C., McCague C., Beer L., et al. "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-

19 using chest radiographs and CT scans". In: *Nature Machine Intelligence* 3.3 (Mar. 2021), pp. 199–217.

[190] Tao Q., Yan W., Wang Y., Paiman E. H. M., Shamonin D. P., Garg P., Plein S., Huang L., Xia L., Sramko M., Tintera J., Roos A. de, Lamb H. J., and Geest R. J. van der. "Deep Learning-based Method for Fully Automatic Quantification of Left Ventricle Function from Cine MR Images: A Multivendor, Multicenter Study". In: *Radiology* 290.1 (Jan. 2019), pp. 81–88.

[191] Hurdman J., Condliffe R., Elliot C. A., Davies C., Hill C., Wild J. M., Capener D., Sephton P., Hamilton N., Armstrong I. J., et al. "ASPIRE registry: assessing the Spectrum of Pulmonary hypertension Identified at a REferral centre". In: *The European Respiratory Journal* 39.4 (Apr. 2012), pp. 945–955.

[192] Swift A. J., Wilson F., Cogliano M., Kendall L., Alandejani F., **Alabed** S., Hughes P., Shahin Y., Saunders L., Oram C., et al. *Repeatability and sensitivity to change of non-invasive end points in PAH: the RESPIRE study.* BMJ Thorax, 2021.

[193] Ronneberger O., Fischer P., and Brox T. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015.* 2015, pp. 234–241.

[194] Gibson E., Li W., Sudre C., Fidon L., Shakir D. I., Wang G., Eaton-Rosen Z., Gray R., Doel T., Hu Y., et al. "NiftyNet: a deep-learning platform for medical imaging". In: *Computer methods and programs in biomedicine* 158 (May 2018), pp. 113–122.

[195] Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G. S., Davis A., Dean J., Devin M., et al. "TensorFlow: Large-scale machine learning on heterogeneous distributed systems". In: (Mar. 2016).

[196] Schroeder W., Martin K., and Lorensen B. "The visualization toolkit, 4th edn. Kitware". In: *New York* (2006).

[197] Maceira A. M., Prasad S. K., Khan M., and Pennell D. J. "Normalized left ventricular systolic and diastolic function by steady state free precession cardiovascular magnetic

resonance". In: *Journal of cardiovascular magnetic resonance* 8.3 (2006), pp. 417–426.

[198] Maceira A. M., Prasad S. K., Khan M., and Pennell D. J. "Reference right ventricular systolic and diastolic function normalized to age, gender and body surface area from steady-state free precession cardiovascular magnetic resonance". In: *European heart journal* 27.23 (Dec. 2006), pp. 2879–2888.

[199] Steiger J. H. "Tests for comparing elements of a correlation matrix". In: *Psychological bulletin* 87.2 (1980), pp. 245–251.

[200] Vallat R. "Pingouin: statistics in Python". In: *Journal of open source software* 3.31 (Nov. 2018), p. 1026.

[201] Davidson-Pilon C., Kalderstam J., Jacobson N., Reed S., Kuhn B., Zivich P., Williamson M., AbdealiJK, Datta D., Fiore-Gartland A., et al. "CamDavidson-Pilon/lifelines: v0.25.9". In: (Feb. 2021).

[202] Hunter J. D. "Matplotlib: A 2D Graphics Environment". In: *Computing in science & engineering* 9.3 (2007), pp. 90–95.

[203] Budai A., Suhai F. I., Csorba K., Toth A., Szabo L., Vago H., and Merkely B. "Fully automatic segmentation of right and left ventricle on short-axis cardiac MRI images". In: *Computerized medical imaging and graphics* 85 (Oct. 2020), p. 101786.

[204] Chen H., Xiang B., Zeng J., Luo H., and Yang Q. "The feasibility in estimating pulmonary vascular resistance by cardiovascular magnetic resonance in pulmonary hypertension: A systematic review and meta-analysis". In: *European Journal of Radiology* 114 (May 2019), pp. 137–145.

[205] Rajaram S., Swift A. J., Capener D., Elliot C. A., Condliffe R., Davies C., Hill C., Hurdman J., Kidling R., Akil M., Wild J. M., and Kiely D. G. "Comparison of the Diagnostic Utility of Cardiac Magnetic Resonance Imaging, Computed Tomography, and Echocardiography in Assessment of Suspected Pulmonary Arterial Hypertension in Patients with Connective Tissue Disease". In: *The Journal of Rheumatology* 39.6 (2012), pp. 1265–1274.

[206] Ullah W., Minalyan A., Saleem S., Nadeem N., Abdullah H. M., Abdalla A., Chan V., Saeed R., Khan M., Collins S., et al. "Comparative accuracy of non-invasive imaging versus right heart catheterization for the diagnosis of pulmonary hypertension: A systematic review and meta-analysis". In: *International journal of cardiology. Heart & vasculature* 29 (2020), p. 100568.

[207] Swift A. J., Rajaram S., Condliffe R., Capener D., Hurdman J., Elliot C. A., Wild J. M., and Kiely D. G. "Diagnostic accuracy of cardiovascular magnetic resonance imaging of right ventricular morphology and function in the assessment of suspected pulmonary hypertension results from the ASPIRE registry". In: *Journal of cardiovascular magnetic resonance* 14 (June 2012), p. 40.

[208] Alunni J.-P., Degano B., Arnaud C., Tétu L., Blot-Soulétie N., Didier A., Otal P., Rousseau H., and Chabbert V. "Cardiac MRI in pulmonary artery hypertension: correlations between morphological and functional parameters and invasive measurements". In: *European Radiology* 20.5 (May 2010), pp. 1149–1159.

[209] Zhang Z., Wang M., Yang Z., Yang F., Li D., Yu T., and Zhang N. "Noninvasive prediction of pulmonary artery pressure and vascular resistance by using cardiac magnetic resonance indices". In: *International journal of cardiology* 227 (Jan. 2017), pp. 915–922.

[210] Ali E. R. and Mohamad A. M. "Diagnostic accuracy of cardiovascular magnetic resonance imaging for assessment of right ventricular morphology and function in pulmonary artery hypertension". In: *The Egyptian journal of chest diseases and tuberculosis* 66.3 (July 2017), pp. 477–486.

[211] Klem I., Shah D. J., White R. D., Pennell D. J., Rossum A. C. van, Regenfus M., Sechtem U., Schvartzman P. R., Hunold P., Croisille P., Parker M., Judd R. M., and Kim R. J. "Prognostic value of routine cardiac magnetic resonance assessment of left ventricular ejection fraction and myocardial damage: an international, multicenter study". In: *Circulation. Cardiovascular imaging* 4.6 (Nov. 2011), pp. 610–619.

[212] Mordi I., Bezerra H., Carrick D., and Tzemos N. "The Combined Incremental Prognostic Value of LVEF, Late Gadolinium Enhancement, and Global Circumferential

Strain Assessed by CMR". In: *JACC. Cardiovascular imaging* 8.5 (May 2015), pp. 540–549.

[213] Rodriguez-Palomares J. F., Gavara J., Ferreira-González I., Valente F., Rios C., Rodrıéguez-Garcıéa J., Bonanad C., Garcıéa Del Blanco B., Miñana G., Mutuberria M., Nuñez J., Barrabés J., Evangelista A., Bodié V., and Garcıéa-Dorado D. "Prognostic Value of Initial Left Ventricular Remodeling in Patients With Reperfused STEMI". In: *JACC. Cardiovascular imaging* 12.12 (Dec. 2019), pp. 2445–2456.

[214] Goh Z. M., **Alabed** S., Shahin Y., Rothman A. M. K., Garg P., Lawrie A., Capener D., Thompson A. A. R., Alandejani F. A. A., Johns C. S., et al. "Right Ventricular Adaptation Assessed Using Cardiac Magnetic Resonance Predicts Survival in Pulmonary Arterial Hypertension". In: *JACC Cardiovascular Imaging* 14.6 (June 2021), pp. 1271–1272.

[215] Goh Z. M., Balasubramanian N., **Alabed** S., Dwivedi K., Shahin Y., Rothman A. M. K., Garg P., Lawrie A., Capener D., Thompson A. A. R., et al. "Right ventricular remodelling in pulmonary arterial hypertension predicts treatment response". In: *Heart* (May 2022).

[216] Bradlow W. M., Hughes M. L., Keenan N. G., Bucciarelli-Ducci C., Assomull R., Gibbs J. S. R., and Mohiaddin R. H. "Measuring the heart in pulmonary arterial hypertension (PAH): implications for trial study size". In: *Journal of magnetic resonance imaging: JMRI* 31.1 (Jan. 2010), pp. 117–124.

[217] Benza R. L., Raina A., Gupta H., Murali S., Burden A., Zastrow M. S., Park M. H., and Simon M. A. "Bosentan-based, treat-to-target therapy in patients with pulmonary arterial hypertension: results from the COMPASS-3 study". In: *Pulmonary Circulation* 8.1 (Jan. 2018), p. 2045893217741480.

[218] Vonk Noordegraaf A., Channick R., Cottreel E., Kiely D. G., Marcus J. T., Martin N., Moiseeva O., Peacock A., Swift A. J., Tawakol A., Torbicki A., Rosenkranz S., and Galie N. "The REPAIR Study: Effects of Macitentan on RV Structure and Function in Pulmonary Arterial Hypertension". In: *JACC. Cardiovascular imaging* 15.2 (Feb. 2022), pp. 240–253.

[219] Wolferen S. A. van, Veerdonk M. C. van de, Mauritz G.-J., Jacobs W., Marcus J. T., Marques K. M. J., Bronzwaer J. G. F., Heymans M. W., Boonstra A., Postmus P. E., Westerhof N., and Vonk Noordegraaf A. "Clinically significant change in stroke volume in pulmonary hypertension". In: *Chest* 139.5 (May 2011), pp. 1003–1009.

[220] Davies R. H., Augusto J. B., Bhuva A., Xue H., Treibel T. A., Ye Y., Hughes R. K., Bai W., Lau C., Shiwani H., et al. "Precision measurement of cardiac structure and function in cardiovascular magnetic resonance using machine learning". In: *Journal of cardiovascular magnetic resonance: official journal of the Society for Cardiovascular Magnetic Resonance* 24.1 (Mar. 2022), p. 16.

[221] Yorke J., Corris P., Gaine S., Gibbs J. S. R., Kiely D. G., Harries C., Pollock V., and Armstrong I. "emPHasis-10: development of a health-related quality of life measure in pulmonary hypertension". In: *The European Respiratory Journal* 43.4 (Apr. 2014), pp. 1106–1113.

[222] Singh S. J., Morgan M. D., Scott S., Walters D., and Hardman A. E. "Development of a shuttle walking test of disability in patients with chronic airways obstruction". In: *Thorax* 47.12 (1992), pp. 1019–1024.

[223] Benza R. L., Gomberg-Maitland M., Miller D. P., Frost A., Frantz R. P., Foreman A. J., Badesch D. B., and McGoon M. D. "The REVEAL Registry risk score calculator in patients newly diagnosed with pulmonary arterial hypertension". In: *Chest* 141.2 (Feb. 2012), pp. 354–362.

[224] Lewis R. A., Billings C. G., Hurdman J. A., Smith I. A., Austin M., Armstrong I. J., Middleton J., Rothman A. M. K., Harrington J., Hamilton N., et al. "Maximal Exercise Testing Using the Incremental Shuttle Walking Test Can Be Used to Risk-Stratify Patients with Pulmonary Arterial Hypertension". In: *Annals of the American Thoracic Society* 18.1 (Jan. 2021), pp. 34–43.

[225] Lewis R., Armstrong I., Bergbaum C., Brewis M. J., Cannon J., Charalampopoulos A., Colin Church A., Gerry Coghlan J., Davies R., Dimopoulos K., et al. "EmPHasis-10 health-related quality of life score predicts outcomes in patients with idiopathic and connective tissue disease-associated pulmonary arterial hypertension: results from a UK multi-centre study". In: *Pulmonary hypertension* (2020).

[226] Borgese M., Badesch D., Bull T., Chakinala M., DeMarco T., Feldman J., Ford H. J., Grinnan D., Klinger J. R., Bolivar L., et al. “EmPHasis-10 as a measure of health-related quality of life in pulmonary arterial hypertension: data from PHAR”. In: *The European Respiratory Journal* 57.2 (Feb. 2021).

[227] Singh S. J., Puhan M. A., Andrianopoulos V., Hernandes N. A., Mitchell K. E., Hill C. J., Lee A. L., Camillo C. A., Troosters T., Spruit M. A., et al. “An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease”. In: *The European Respiratory Journal* 44.6 (Dec. 2014), pp. 1447–1478.

[228] Singh S. J., Jones P. W., Evans R., and Morgan M. D. L. “Minimum clinically important improvement for the incremental shuttle walking test”. In: *Thorax* 63.9 (Sept. 2008), pp. 775–777.

[229] Salas Apaza J. A., Franco J. V. A., Meza N., Madrid E., Loezar C., and Garegnani L. “Minimal clinically important difference: The basics”. In: *Medwave* 21.3 (Apr. 2021), e8149.

[230] Guyatt G. H., Osoba D., Wu A. W., Wyrwich K. W., Norman G. R., and Clinical Significance Consensus Meeting Group. “Methods to explain the clinical significance of health status measures”. In: *Mayo Clinic proceedings. Mayo Clinic* 77.4 (Apr. 2002), pp. 371–383.

[231] Revicki D., Hays R. D., Cella D., and Sloan J. “Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes”. In: *Journal of Clinical Epidemiology* 61.2 (Feb. 2008), pp. 102–109.

[232] DeLong E. R., DeLong D. M., and Clarke-Pearson D. L. “Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach”. In: *Biometrics* 44.3 (1988), p. 837.

[233] Michelakis E. D., Tymchak W., Noga M., Webster L., Wu X.-C., Lien D., Wang S.-H., Modry D., and Archer S. L. “Long-term treatment with oral sildenafil is safe and improves functional capacity and hemodynamics in patients with pulmonary arterial hypertension”. In: *Circulation* 108.17 (Oct. 2003), pp. 2066–2069.

[234] Roeleveld R. J., Vonk-Noordegraaf A., Marcus J. T., Bronzwaer J. G. F., Marques K. M. J., Postmus P. E., and Boonstra A. “Effects of epoprostenol on right ventricular hypertrophy and dilatation in pulmonary hypertension”. In: *Chest* 125.2 (Feb. 2004), pp. 572–579.

[235] Veerdonk M. C. van de, Huis In T Veld A. E., Marcus J. T., Westerhof N., Heymans M. W., Bogaard H.-J., and Vonk-Noordegraaf A. “Upfront combination therapy reduces right ventricular volumes in pulmonary arterial hypertension”. In: *The European Respiratory Journal* 49.6 (June 2017).

[236] Wolferen S. A. van, Boonstra A., Marcus J. T., Marques K. M. J., Bronzwaer J. G. F., Postmus P. E., and Vonk-Noordegraaf A. “Right ventricular reverse remodelling after sildenafil in pulmonary arterial hypertension”. In: *Heart* 92.12 (Dec. 2006), pp. 1860–1861.

[237] Wilkins M. R., Paul G. A., Strange J. W., Tunariu N., Gin-Sing W., Banya W. A., Westwood M. A., Stefanidis A., Ng L. L., Pennell D. J., Mohiaddin R. H., Nihoyannopoulos P., and Gibbs J. S. R. “Sildenafil versus Endothelin Receptor Antagonist for Pulmonary Hypertension (SERAPH) study”. In: *American journal of respiratory and critical care medicine* 171.11 (June 2005), pp. 1292–1297.

[238] Billings C. G., Lewis R., Hurdman J. A., Condliffe R., Elliot C. A., Roger Thompson A. A., Smith I. A., Austin M., Armstrong I. J., Hamilton N., et al. “The incremental shuttle walk test predicts mortality in non-group 1 pulmonary hypertension: results from the ASPIRE Registry”. In: *Pulmonary Circulation* 9.2 (2019), pp. 1–9.

[239] Richeldi L., Ryerson C. J., Lee J. S., Wolters P. J., Koth L. L., Ley B., Elicker B. M., Jones K. D., King Jr T. E., Ryu J. H., and Collard H. R. “Relative versus absolute change in forced vital capacity in idiopathic pulmonary fibrosis”. In: *Thorax* 67.5 (May 2012), pp. 407–411.

[240] Mannil M., Eberhard M., Spiczak J. von, Heindel W., Alkadhi H., and Baessler B. “Artificial Intelligence and Texture Analysis in Cardiac Imaging”. In: *Current cardiology reports* 22.11 (Sept. 2020), p. 131.

[241] Luo C., Shi C., Li X., and Gao D. “Cardiac MR segmentation based on sequence propagation by deep learning”. In: *PloS one* 15.4 (Apr. 2020), e0230415.

[242] Song X., Meng L., Shi Q., and Lu H. “Learning Tensor-Based Features for Whole-Brain fMRI Classification”. In: *Lecture notes in computer science* (2015), pp. 613–620.

[243] Collins G. S., Reitsma J. B., Altman D. G., and Moons K. G. M. “Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement”. In: *BMJ* 350 (Jan. 2015), g7594.

[244] Uthoff J., **Alabed** S., Swift A., and Lu H. “Geodesically smoothed tensor features for pulmonary hypertension prognosis using the heart and surrounding tissues”. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Lecture Notes in Computer Science, June 2020, vol 12262.

[245] Lu H., Plataniotis K. N. K., and Venetsanopoulos A. N. “MPCA: Multilinear Principal Component Analysis of Tensor Objects”. In: *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council* 19.1 (Jan. 2008), pp. 18–39.

[246] Farha S., Laskowski D., George D., Park M. M., Tang W. H. W., Dweik R. A., and Erzurum S. C. “Loss of alveolar membrane diffusing capacity and pulmonary capillary blood volume in pulmonary arterial hypertension”. In: *Respiratory research* 14 (Jan. 2013), p. 6.

[247] Low A. T., Medford A. R. L., Millar A. B., and Tulloh R. M. R. “Lung function in pulmonary hypertension”. In: *Respiratory Medicine* 109.10 (2015), pp. 1244–1249.

[248] Szturmowicz M., Kacprzak A., Franczuk M., Burakowska B., Kurzyna M., Fijałkowska A., Skoczylas A., Wesołowski S., Kuś J., and Torbicki A. “Low DLCO in idiopathic pulmonary arterial hypertension - clinical correlates and prognostic significance”. In: *Pneumonologia i alergologia polska* 84.2 (2016), pp. 87–94.

[249] Quanjer P. H., Stanojevic S., Cole T. J., Baur X., Hall G. L., Culver B. H., Enright P. L., Hankinson J. L., Ip M. S. M., Zheng J., Stocks J., and the ERS Global Lung Function Initiative. “Multi-ethnic reference values for spirometry for the 3–95-yr

age range: the global lung function 2012 equations". In: *European Respiratory Journal* 40.6 (2012), pp. 1324–1343.

[250] Stanojevic S., Graham B. L., Cooper B. G., Thompson B. R., Carter K. W., Francis R. W., Hall G. L., Global Lung Function Initiative TLCO working group, and Global Lung Function Initiative (GLI) TLCO. "Official ERS technical standards: Global Lung Function Initiative reference values for the carbon monoxide transfer factor for Caucasians". In: *The European Respiratory Journal* 50.3 (Sept. 2017).

[251] Raymond R. J., Hinderliter A. L., Willis P. W., Ralph D., Caldwell E. J., Williams W., Ettinger N. A., Hill N. S., Summer W. R., Boisblanc B. de, et al. "Echocardiographic predictors of adverse outcomes in primary pulmonary hypertension". In: *Journal of the American College of Cardiology* 39.7 (Apr. 2002), pp. 1214–1219.

[252] Fenstad E. R., Le R. J., Sinak L. J., Maradit-Kremers H., Ammash N. M., Ayalew A. M., Villarraga H. R., Oh J. K., Frantz R. P., McCully R. B., McGoon M. D., and Kane G. C. "Pericardial effusions in pulmonary arterial hypertension: characteristics, prognosis, and role of drainage". In: *Chest* 144.5 (Nov. 2013), pp. 1530–1538.

[253] Humbert M., Sitbon O., Chaouat A., Bertocchi M., Habib G., Gressin V., Yaıęci A., Weitzenblum E., Cordier J.-F., Chabot F., et al. "Survival in patients with idiopathic, familial, and anorexigen-associated pulmonary arterial hypertension in the modern management era". In: *Circulation* 122.2 (July 2010), pp. 156–163.

[254] Groepenhoff H., Vonk-Noordegraaf A., Boonstra A., Spreeuwenberg M. D., Postmus P. E., and Bogaard H. J. "Exercise Testing to Estimate Survival in Pulmonary Hypertension". In: *Medicine & Science in Sports & Exercise* 40.10 (2008), pp. 1725–1732.

[255] Proficiency Standards for Clinical Pulmonary Function Laboratories A. C. on. "ATS statement: guidelines for the six-minute walk test". In: *American journal of respiratory and critical care medicine* 166.1 (July 2002), pp. 111–117.

[256] Smeden M. van, Moons K. G., Groot J. A. de, Collins G. S., Altman D. G., Eijkemans M. J., and Reitsma J. B. "Sample size for binary logistic prediction

models: Beyond events per variable criteria". In: *Statistical methods in medical research* 28.8 (Aug. 2019), pp. 2455–2474.

[257] Dayton C. M. and Mitchell Dayton C. "Model Comparisons Using Information Measures". In: *Journal of Modern Applied Statistical Methods* 2.2 (2003), pp. 281–292.

[258] Rudin C. "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead". In: *Nature Machine Intelligence* 1.5 (2019), pp. 206–215.

[259] Kim B. S., Heo R., Shin J., Lim Y.-H., and Park J.-K. "E'E and D-shaped Left Ventricle Severity in Patients with Increased Pulmonary Artery Pressure". In: *Journal of Cardiovascular Imaging* 26.2 (2018), p. 85.

[260] Tello K., Dalmer A., Vanderpool R., Ghofrani H. A., Naeije R., Roller F., Seeger W., Wilhelm J., Gall H., and Richter M. J. "Cardiac Magnetic Resonance Imaging-Based Right Ventricular Strain Analysis for Assessment of Coupling and Diastolic Function in Pulmonary Hypertension". In: *JACC. Cardiovascular imaging* 12.11 Pt 1 (Nov. 2019), pp. 2155–2164.

[261] Schobs L., Zhou S., Cogliano M., Swift A. J., and Lu H. "Confidence-Quantifying Landmark Localisation For Cardiac MRI". In: *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)* (2021).

[262] Bello G. A., Dawes T. J. W., Duan J., Biffi C., Marvao A. de, Howard L. S. G., Gibbs J. S. R., Wilkins M. R., Cook S. A., Rueckert D., and OŔegan D. P. "Deep-learning cardiac motion analysis for human survival prediction". In: *Nature Machine Intelligence* 1.2 (2019), pp. 95–104.

[263] Ferreira V. M., Piechnik S. K., Dall'Armellina E., Karamitsos T. D., Francis J. M., Choudhury R. P., Friedrich M. G., Robson M. D., and Neubauer S. "Non-contrast T1-mapping detects acute myocardial edema with high diagnostic accuracy: a comparison to T2-weighted cardiovascular magnetic resonance". In: *Journal of cardiovascular magnetic resonance* 14.42 (June 2012), p. 42.

[264] Karamitsos T. D., Piechnik S. K., Banypersad S. M., Fontana M., Ntusi N. B., Ferreira V. M., Whelan C. J., Myerson S. G., Robson M. D., Hawkins P. N., Neubauer S., and Moon J. C. "Noncontrast T1 mapping for the diagnosis of cardiac amyloidosis". In: *JACC. Cardiovascular imaging* 6.4 (Apr. 2013), pp. 488–497.

[265] Ugander M., Bagi P. S., Oki A. J., Chen B., Hsu L.-Y., Aletras A. H., Shah S., Greiser A., Kellman P., and Arai A. E. "Myocardial edema as detected by pre-contrast T1 and T2 CMR delineates area at risk associated with acute myocardial infarction". In: *JACC. Cardiovascular imaging* 5.6 (June 2012), pp. 596–603.

[266] DallÁrmellina E., Piechnik S. K., Ferreira V. M., Si Q. L., Robson M. D., Francis J. M., Cuculi F., Kharbanda R. K., Banning A. P., Choudhury R. P., Karamitsos T. D., and Neubauer S. "Cardiovascular magnetic resonance by non contrast T1-mapping allows assessment of severity of injury in acute myocardial infarction". In: *Journal of cardiovascular magnetic resonance* 14.15 (Feb. 2012), p. 15.

[267] Dass S., Suttie J. J., Piechnik S. K., Ferreira V. M., Holloway C. J., Banerjee R., Mahmod M., Cochlin L., Karamitsos T. D., Robson M. D., Watkins H., and Neubauer S. "Myocardial tissue characterization using magnetic resonance noncontrast T1 mapping in hypertrophic and dilated cardiomyopathy". In: *Circulation. Cardiovascular imaging* 5.6 (Nov. 2012), pp. 726–733.

[268] Mascherbauer J., Marzluf B. A., Tufaro C., Pfaffenberger S., Graf A., Wexberg P., Panzenböck A., Jakowitsch J., Bangert C., Laimer D., et al. "Cardiac magnetic resonance postcontrast T1 time is associated with outcome in patients with heart failure and preserved ejection fraction". In: *Circulation. Cardiovascular imaging* 6.6 (Nov. 2013), pp. 1056–1065.

[269] Messroghli D. R., Moon J. C., Ferreira V. M., Grosse-Wortmann L., He T., Kellman P., Mascherbauer J., Nezafat R., Salerno M., Schelbert E. B., Taylor A. J., Thompson R., Ugander M., Heeswijk R. B. van, and Friedrich M. G. "Clinical recommendations for cardiovascular magnetic resonance mapping of T1, T2, T2*and extracellular volume: A consensus statement by the Society for Cardiovascular Magnetic Resonance (SCMR) endorsed by the European Association for Cardiovascular

Imaging (EACVI)". In: *Journal of cardiovascular magnetic resonance* 19.1 (Oct. 2017), p. 75.

[270] Moon J. C., Messroghli D. R., Kellman P., Piechnik S. K., Robson M. D., Ugander M., Gatehouse P. D., Arai A. E., Friedrich M. G., Neubauer S., Schulz-Menger J., Schelbert E. B., Society for Cardiovascular Magnetic Resonance Imaging, and Cardiovascular Magnetic Resonance Working Group of the European Society of Cardiology. "Myocardial T1 mapping and extracellular volume quantification: a Society for Cardiovascular Magnetic Resonance (SCMR) and CMR Working Group of the European Society of Cardiology consensus statement". In: *Journal of cardiovascular magnetic resonance* 15.92 (Oct. 2013), p. 92.

[271] Nakamori S., Dohi K., Ishida M., Goto Y., Imanaka-Yoshida K., Omori T., Goto I., Kumagai N., Fujimoto N., Ichikawa Y., Kitagawa K., Yamada N., Sakuma H., and Ito M. "Native T1 Mapping and Extracellular Volume Mapping for the Assessment of Diffuse Myocardial Fibrosis in Dilated Cardiomyopathy". In: *JACC. Cardiovascular imaging* 11.1 (Jan. 2018), pp. 48–59.

[272] Puntmann V. O., Voigt T., Chen Z., Mayr M., Karim R., Rhode K., Pastor A., Carr-White G., Razavi R., Schaeffter T., and Nagel E. "Native T1 mapping in differentiation of normal myocardium from diffuse disease in hypertrophic and dilated cardiomyopathy". In: *JACC. Cardiovascular imaging* 6.4 (Apr. 2013), pp. 475–484.

[273] Mikami Y., Alfagih R., Khan A., Fine N., and White J. "Value of non-contrast T1 mapping MRI for the differentiation of hypertrophic cardiomyopathy, cardiac amyloid and Fabry cardiomyopathy". In: *Canadian Journal of Cardiology* 33.10 (Oct. 2017), S22.

[274] Pagano J. J., Chow K., Khan A., Michelakis E., Paterson I., Oudit G. Y., and Thompson R. B. "Reduced Right Ventricular Native Myocardial T1 in Anderson-Fabry Disease: Comparison to Pulmonary Hypertension and Healthy Controls". In: *PloS one* 11.6 (June 2016), e0157565.

[275] Sado D. M., White S. K., Piechnik S. K., Banypersad S. M., Treibel T., Captur G., Fontana M., Maestrini V., Flett A. S., Robson M. D., et al. "Identification and assessment of Anderson-Fabry disease by cardiovascular magnetic resonance

noncontrast myocardial T1 mapping". In: *Circulation. Cardiovascular imaging* 6.3 (May 2013), pp. 392–398.

[276] Thompson R. B., Chow K., Khan A., Chan A., Shanks M., Paterson I., and Oudit G. Y. "T1 mapping with cardiovascular MRI is highly sensitive for Fabry disease independent of hypertrophy and sex". In: *Circulation. Cardiovascular imaging* 6.5 (Sept. 2013), pp. 637–645.

[277] Wilson H. C., Ambach S., Madueme P. C., Khoury P. R., Hopkin R. J., and Jefferies J. L. "Comparison of Native T1, Strain, and Traditional Measures of Cardiovascular Structure and Function by Cardiac Magnetic Resonance Imaging in Patients With Anderson-Fabry Disease". In: *The American journal of cardiology* 122.6 (Sept. 2018), pp. 1074–1078.

[278] Krittayaphong R., Zhang S., Saiviroonporn P., Viprakasit V., Tanapibunpon P., Rerkudom B., Yindeengam A., and Wood J. C. "Assessment of Cardiac Iron Overload in Thalassemia With MRI on 3.0-T: High-Field T1, T2, and $T2*$ Quantitative Parametric Mapping in Comparison to T2∗ on 1.5-T". In: *JACC. Cardiovascular imaging* 12.4 (Apr. 2019), pp. 752–754.

[279] Torlasco, Cassinerio, Roghi, Faini, Capecchi, Abdel-Gadir, Giannattasio, Parati, Moon, Cappellini, and Pedrotti. "Role of T1 mapping as a complementary tool to $T2*$ for non-invasive cardiac iron overload assessment". In: *PloS one* 13.2 (Feb. 2018), e0192890.

[280] Cui Y., Cao Y., Song J., Dong N., Kong X., Wang J., Yuan Y., Zhu X., Yan X., Greiser A., Shi H., and Han P. "Association between myocardial extracellular volume and strain analysis through cardiovascular magnetic resonance with histological myocardial fibrosis in patients awaiting heart transplantation". In: *Journal of cardiovascular magnetic resonance* 20.1 (Apr. 2018), p. 25.

[281] Kammerlander A., Tufaro C., Bachmann A. F., Marzluf B. A., Aschauer S., Greiser A., Bonderman D., and Mascherbauer J. "Extracellular matrix expansion by cardiac magnetic resonance T1 mapping- validation with myocardial biopsy". In: *Journal of cardiovascular magnetic resonance* 17.S1 (Feb. 2015), P308.

[282] Wong T. C., Piehler K., Meier C. G., Testa S. M., Klock A. M., Aneizi A. A., Shakesprere J., Kellman P., Shroff S. G., Schwartzman D. S., Mulukutla S. R., Simon M. A., and Schelbert E. B. "Association between extracellular matrix expansion quantified by cardiovascular magnetic resonance and short-term mortality". In: *Circulation* 126.10 (Sept. 2012), pp. 1206–1216.

[283] Puntmann V. O., Carr-White G., Jabbour A., Yu C.-Y., Gebker R., Kelle S., Rolf A., Zitzmann S., Peker E., DÁngelo T., et al. "Native T1 and ECV of Noninfarcted Myocardium and Outcome in Patients With Coronary Artery Disease". In: *Journal of the American College of Cardiology* 71.7 (Feb. 2018), pp. 766–778.

[284] Puntmann V. O., Carr-White G., Jabbour A., Yu C.-Y., Gebker R., Kelle S., Hinojar R., Doltra A., Varma N., Child N., et al. "T1-Mapping and Outcome in Nonischemic Cardiomyopathy: All-Cause Mortality and Heart Failure". In: *JACC. Cardiovascular imaging* 9.1 (Jan. 2016), pp. 40–50.

[285] Andersen S., Nielsen-Kudsk J. E., Vonk Noordegraaf A., and Man F. S. de. "Right Ventricular Fibrosis". In: *Circulation* 139.2 (Jan. 2019), pp. 269–285.

[286] Gottbrecht M., Kramer C. M., and Salerno M. "Native T1 and Extracellular Volume Measurements by Cardiac MRI in Healthy Adults: A Meta-Analysis". In: *Radiology* 290.2 (Feb. 2019), pp. 317–326.

[287] Zhuang B., Sirajuddin A., Wang S., Arai A., Zhao S., and Lu M. "Prognostic value of T1 mapping and extracellular volume fraction in cardiovascular disease: a systematic review and meta-analysis". In: *Heart failure reviews* 23.5 (Sept. 2018), pp. 723–731.

[288] Minegishi S., Kato S., Takase-Minegishi K., Horita N., Azushima K., Wakui H., Ishigami T., Kosuge M., Kimura K., and Tamura K. "Native T1 time and extracellular volume fraction in differentiation of normal myocardium from non-ischemic dilated and hypertrophic cardiomyopathy myocardium: A systematic review and meta-analysis". In: *International journal of cardiology. Heart & vasculature* 25 (Dec. 2019), p. 100422.

[289] Boomen M. van den, Slart R. H., Hulleman E. V., Dierckx R. A., Velthuis B. K., Harst P. van der, Sosnovik D. E., Borra R. J., and Prakken N. H. "Native T1 reference values for nonischemic cardiomyopathies and populations with increased cardiovascular risk: A systematic review and meta-analysis". In: *Journal of magnetic resonance imaging: JMRI* 47.4 (Apr. 2018), pp. 891–912.

[290] Sterne J. A. C., Egger M., and Moher D. "Addressing Reporting Biases". In: *Cochrane Handbook for Systematic Reviews of Interventions*. 2008, pp. 297–333.

[291] Spruijt O. A., Vissers L., Bogaard H.-J., Hofman M. B. M., Vonk-Noordegraaf A., and Marcus J. T. "Increased native T1-values at the interventricular insertion regions in precapillary pulmonary hypertension". In: *The international journal of cardiovascular imaging* 32.3 (Mar. 2016), pp. 451–459.

[292] Tello K., Dalmer A., Axmann J., Vanderpool R., Ghofrani H. A., Naeije R., Roller F., Seeger W., Sommer N., Wilhelm J., Gall H., and Richter M. J. "Reserve of Right Ventricular-Arterial Coupling in the Setting of Chronic Overload". In: *Circulation. Heart failure* 12.1 (Jan. 2019), e005512.

[293] Asano Ryotaro, Ogo Takeshi, Morita Yoshiaki, Tsuji Akihiro, Fukui Shigefumi, Ueda Jin, Kotoku Akiyoshi, Fukuda Tetsuya, Ohta-Ogo Keiko, Ishibashi-Ueda Hatsue, Noguchi Teruo, and Yasuda Satoshi. "Abstract 15190: Native T1 Mapping: A Novel Prognostic Marker in Pulmonary Hypertension". In: *Circulation* 138.Suppl_1 (Nov. 2018), A15190–a15190.

[294] Habert P., Capron T., Hubert S., Bentatou Z., Bartoli A., Tradi F., Renard S., Rapacchi S., Guye M., Bernard M., Habib G., and Jacquier A. "Quantification of right ventricular extracellular volume in pulmonary hypertension using cardiac magnetic resonance imaging". In: *Diagnostic and interventional imaging* 101.5 (May 2020), pp. 311–320.

[295] Broberg C. S., Prasad S. K., Carr C., Babu-Narayan S. V., Dimopoulos K., and Gatzoulis M. A. "Myocardial fibrosis in Eisenmenger syndrome: a descriptive cohort study exploring associations of late gadolinium enhancement with clinical status and survival". In: *Journal of Cardiovascular Magnetic Resonance* 16.1 (2014).

[296] Blyth K. G., Groenning B. A., Martin T. N., Foster J. E., Mark P. B., Dargie H. J., and Peacock A. J. "Contrast enhanced-cardiovascular magnetic resonance imaging in patients with pulmonary hypertension". In: *European heart journal* 26.19 (Oct. 2005), pp. 1993–1999.

[297] Bratis K., Lindholm A., Hesselstrand R., Arheden H., Karabela G., Stavropoulos E., Katsifis G., Kolovou G., Kitas G. D., Sfikakis P. P., Koutsogeorgopoulou L., Mavrogeni S., and Ostenfeld E. "CMR feature tracking in cardiac asymptomatic systemic sclerosis: Clinical implications". In: *PloS one* 14.8 (Aug. 2019), e0221021.

[298] Ntusi N. A. B., Piechnik S. K., Francis J. M., Ferreira V. M., Rai A. B. S., Matthews P. M., Robson M. D., Moon J., Wordsworth P. B., Neubauer S., and Karamitsos T. D. "Subclinical myocardial inflammation and diffuse fibrosis are common in systemic sclerosis – a clinical study using myocardial T1-mapping and extracellular volume quantification". In: *Journal of Cardiovascular Magnetic Resonance* 16.1 (2014), p. 21.

[299] Brown J., Norrington K., Kotecha T., Martinez-Naharro A., Fayed H., Teresi L., Denton C., Schreiber B., Fontana M., Kellman P., Coghlan J., and Knight D. S. "Subclinical myocardial abnormalities in systemic sclerosis-associated versus non-connective tissue disease pulmonary hypertension by CMR multiparametric mapping". In: 40.Supplement_1 (2019).

[300] Jellis C. L., Yingchoncharoen T., Gai N., Kusunose K., Popović Z. B., Flamm S., and Kwon D. "Correlation between right ventricular T1 mapping and right ventricular dysfunction in non-ischemic cardiomyopathy". In: *The international journal of cardiovascular imaging* 34.1 (Jan. 2018), pp. 55–65.

[301] McCann G. P., Beek A. M., Vonk-Noordegraaf A., and Rossum A. C. van. "Delayed contrast-enhanced magnetic resonance imaging in pulmonary arterial hypertension". In: *Circulation* 112.16 (Oct. 2005), e268.

[302] Sanz J., Dellegrottaglie S., Kariisa M., Sulica R., Poon M., ODéonnell T. P., Mehta D., Fuster V., and Rajagopalan S. "Prevalence and correlates of septal delayed contrast enhancement in patients with pulmonary hypertension". In: *The American journal of cardiology* 100.4 (Aug. 2007), pp. 731–735.

[303] De Lazzari M., Cipriani A., Rizzo S., Famoso G., Giorgi B., Tarantini G., Thiene G., Tona F., Iliceto S., Basso C., and Perazzolo Marra M. “Right Ventricular Junctional Late Gadolinium Enhancement Correlates With Outcomes in Pulmonary Hypertension”. In: *JACC. Cardiovascular imaging* 12.5 (May 2019), pp. 936–938.

[304] Garcıéa-Álvarez A., Garcıéa-Lunar I., Pereda D., Fernández-Jimenez R., Sánchez-González J., Mirelis J. G., Nuño-Ayala M., Sánchez-Quintana D., Fernández-Friera L., Garcıéa-Ruiz J. M., et al. “Association of myocardial T1-mapping CMR with hemodynamics and RV performance in pulmonary hypertension”. In: *JACC. Cardiovascular imaging* 8.1 (Jan. 2015), pp. 76–82.

[305] Kellman P. and Hansen M. S. “T1-mapping in the heart: accuracy and precision”. In: *Journal of cardiovascular magnetic resonance* 16.2 (Jan. 2014), p. 2.

[306] Jang J., Ngo L. H., Captur G., Moon J. C., and Nezafat R. “Measurement reproducibility of slice-interleaved T1 and T2 mapping sequences over 20 months: A single center study”. In: *PloS one* 14.7 (July 2019), e0220190.

[307] Rogers T., Dabir D., Mahmoud I., Voigt T., Schaeffter T., Nagel E., and Puntmann V. O. “Standardization of T1 measurements with MOLLI in differentiation between health and disease–the ConSept study”. In: *Journal of cardiovascular magnetic resonance* 15.78 (Sept. 2013), p. 78.

[308] Zhu Y., Fahmy A. S., Duan C., Nakamori S., and Nezafat R. “Automated Myocardial T2 and Extracellular Volume Quantification in Cardiac MRI Using Transfer Learning-based Myocardium Segmentation”. In: *Radiology. Artificial intelligence* 2.1 (Jan. 2020), e190034.

[309] Raman F. S., Kawel-Boehm N., Gai N., Freed M., Han J., Liu C.-Y., Lima J. A. C., Bluemke D. A., and Liu S. “Modified look-locker inversion recovery T1 mapping indices: assessment of accuracy and reproducibility between magnetic resonance scanners”. In: *Journal of cardiovascular magnetic resonance* 15.64 (July 2013), p. 64.

[310] Fortin J.-P., Sweeney E. M., Muschelli J., Crainiceanu C. M., and Shinohara R. T. "Removing inter-subject technical variability in magnetic resonance imaging studies". In: *NeuroImage* 132 (May 2016), pp. 198–212.

[311] Fortin J.-P., Parker D., Tunç B., Watanabe T., Elliott M. A., Ruparel K., Roalf D. R., Satterthwaite T. D., Gur R. C., Gur R. E., and al. et. "Harmonization of multi-site diffusion tensor imaging data". In: *NeuroImage* 161 (Nov. 2017), pp. 149–170.

[312] Humbert M., Kovacs G., Hoeper M. M., Badagliacca R., Berger R. M. F., Brida M., Carlsen J., Coats A. J. S., Escribano-Subias P., Ferrari P., et al. "2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension". In: *European Respiratory Journal* 61.1 (Jan. 2023).

[313] Ennals E. "AI spots damage on heart scans in seconds". In: *The Daily Mail* (Dec. 2022).

[314] Harrison H. "NHS heart patients to recieve quicker diagnosis due to Sheffield University and Teaching Hospitals innovation". In: *The Star* (Dec. 2022).

[315] Ridley E. L. *AI-based measurements increase utility of cardiac MRI*. `https://www.auntminnie.com/index.aspx?sec=road&sub=aic_2021&pag=dis&ItemID=133997`. Accessed: 2023-1-6. Dec. 2021.

[316] Chu L. *Validation of AI Cardiac MRI Measurements*. Oct. 2022.

[317] Backhaus S. J., Rosel S. F., Stiermaier T., Schmidt-Rimpler J., Evertz R., Schulz A., Lange T., Kowallick J. T., Kutty S., Bigalke B., Gutberlet M., Hasenfuß G., Thiele H., Eitel I., and Schuster A. "Left-atrial long-axis shortening allows effective quantification of atrial function and optimized risk prediction following acute myocardial infarction". In: *European Heart Journal Open* 2.5 (Sept. 2022), oeac053.

[318] Messika-Zeitoun D., Bellamy M., Avierinos J.-F., Breen J., Eusemann C., Rossi A., Behrenbeck T., Scott C., Tajik J. A., and Enriquez-Sarano M. "Left atrial remodelling in mitral regurgitation–methodologic approach, physiological determinants, and outcome implications: a prospective quantitative Doppler-echocardiographic

and electron beam-computed tomographic study". In: *European Heart Journal* 28.14 (July 2007), pp. 1773–1781.

[319] Garg P., Gosling R., Swoboda P., Jones R., Rothman A., Wild J. M., Kiely D. G., Condliffe R., **Alabed** S., and Swift A. J. *Cardiac magnetic resonance identifies raised left ventricular filling pressure: prognostic implications.* European Heart Journal, May 2022.

[320] Kawel-Boehm N., Hetzel S. J., Ambale-Venkatesh B., Captur G., Francois C. J., Jerosch-Herold M., Salerno M., Teague S. D., Valsangiacomo-Buechel E., Geest R. J. van der, and Bluemke D. A. "Reference ranges ("normal values") for cardiovascular magnetic resonance (CMR) in adults and children: 2020 update". In: *Journal Cardiovascular Magnetic Resonance* 22.1 (Dec. 2020), p. 87.

[321] Macdonald A., Salehi M., **Alabed** S., Maiter A., Goh Z. M., Dwivedi K., Johns C., Cogliano M., Alandejani F., Condliffe R., Wild J. M., Kiely D. G., Garg P., and Swift A. J. "Semi-automatic thresholding of RV trabeculation improves repeatability and diagnostic value in suspected pulmonary hypertension". In: *Frontiers in cardiovascular medicine* 9 (Jan. 2023).

[322] **Alabed** S., Uthoff J., Zhou S., Garg P., Dwivedi K., Alandejani F., Gosling R., Schobs L., Brook M., Capener D., et al. *Machine Learning cardiac-MRI features predict mortality in newly diagnosed pulmonary arterial hypertension.* European Heart Journal - Digital Health, May 2022.

[323] Yang J. and Huang S. *Current and Future Application of Artificial Intelligence in Clinical Medicine.* Bentham Science Publishers, June 2021.

[324] Alandejani F., **Alabed** S., Garg P., Goh Z. M., Karunasaagarar K., Sharkey M., Salehi M., Aldabbagh Z., Dwivedi K., Mamalakis M., et al. *Training and clinical testing of artificial intelligence derived right atrial cardiovascular magnetic resonance measurements.* Vol. 24. 1. Journal of Cardiovascular Magnetic Resonance, 2022, p. 25.

[325] Sande D. van de, Van Genderen M. E., Smit J. M., Huiskens J., Visser J. J., Veen R. E. R., Unen E. van, Ba O. H., Gommers D., and Bommel J. van. "Develop-

ing, implementing and governing artificial intelligence in medicine: a step-by-step approach to prevent an artificial intelligence winter". In: *BMJ* 29.1 (Feb. 2022).

[326] Vos J. L., Leiner T., Dijk A. P. J. van, Pedrizzetti G., Alenezi F., Rodwell L., Wegen C. T. P. M. van der, Post M. C., Driessen M. M. P., and Nijveldt R. "Cardiovascular magnetic resonance-derived left ventricular intraventricular pressure gradients among patients with precapillary pulmonary hypertension". In: *European Heart Journal Cardiovascular Imaging* 24.1 (Dec. 2022), pp. 78–87.

[327] Park S. H. and Han K. "Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction". In: *Radiology* 286.3 (Mar. 2018), pp. 800–809.

[328] Griffiths P. D., Bradburn M., Campbell M. J., Cooper C. L., Graham R., Jarvis D., Kilby M. D., Mason G., Mooney C., Robson S. C., Wailoo A., and MERIDIAN collaborative group. "Use of MRI in the diagnosis of fetal brain abnormalities in utero (MERIDIAN): a multicentre, prospective cohort study". In: *Lancet* 389.10068 (Feb. 2017), pp. 538–546.

Index