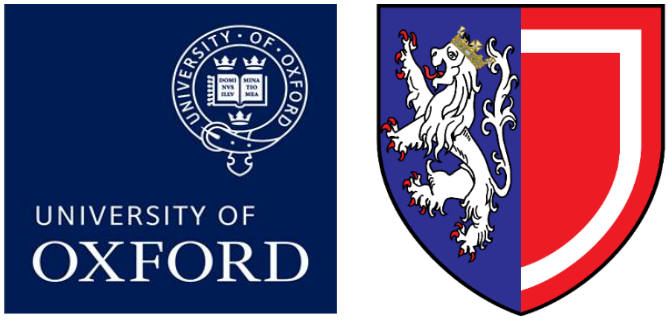


The Development of Novel Artificial Intelligence Tools for the Early Detection of Cardiovascular Disease Using Computerised Tomography

Henry William West

Balliol College
Division of Cardiovascular Medicine
Radcliffe Department of Medicine
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy in Medical Sciences at the University of Oxford, UK

Michaelmas Term 2022

Supervisors:
Professor Charalambos Antoniades & Professor Keith M. Channon

Word Count: <50,000
Excludes references, appendices & figure legends

Contents

Acknowledgements ix
Dedication xi
COVID-19 Statement xii
Abstract xiii
List of Publications and Presentations xiv
Publications since matriculation xiv
Patents since matriculation xv
Awards since matriculation xv
Presentations since matriculation xv
List of Figures xvii
List of Tables xix
List of Abbreviations xx
Chapter 1. Introduction 1
1.1. The burden of cardiovascular diseases including atrial fibrillation and stroke 1
1.2. Cardiovascular risk stratification and prevention 2
1.2.1. Current approaches in cardiovascular disease prevention 2
1.2.2. The residual cardiovascular risk 3
1.3. The atrial myopathy 5
1.3.1. Atrial myopathy biology 7
1.3.2. Detecting the atrial myopathy 10
1.4. The role of epicardial adipose tissue in cardiovascular disease 11
1.4.1. Role of EAT in coronary artery disease 13
1.4.2. Role of EAT in atrial fibrillation 14
1.5. EAT as a biosensor of vascular and cardiovascular dysfunction 15
1.5.1. Vascular wall to EAT signalling 15
1.5.2. Atrial wall to EAT signalling 18
1.6. Imaging the left atrium and surrounding tissues using computed tomography 19
1.6.1. Quantitative assessment of epicardial adipose tissue 20
1.6.2. Quantitative assessment of the left atrium and surrounding tissues 22

1.7. The radiomic approach to computed tomography interpretation 22

1.7.1. Radiomic texture analysis 23

1.7.2. Radiomic features 23

1.7.3. The big data approach to stroke risk with radiomic and clinical data 24

1.8. Hypotheses and aims 25

Chapter 2. General methods 26

2.1. Introduction 26

2.2. Study populations included in this thesis 26

2.2.1. The development of automated segmentation and radiomic feature extraction: The ORFAN Arm 4 Study 27

2.2.2. Epicardial adipose tissue clinical associations: SCOT-HEART Trial & AdipoRedOx Study 29

2.2.3. The Atriomic Risk Algorithm: The ORFAN Arm 4 Study 30

2.3. Coronary computed tomography angiography protocols 31

2.4. The Oxford Academic Cardiovascular Computed Tomography (OXACCT) Core Lab 36

2.5. CT scan transfer 36

2.6. Quality control of image datasets 38

2.7. The automated extraction of radiomic features from CT volumes 39

2.8. Radiomic features 40

2.9. Agatston coronary calcium score 44

2.10. Qualitative coronary analysis 45

2.10.1. Coronary artery stenosis assessment 45

Chapter 3. The Oxford Risk Factors And Non Invasive Imaging Study 47

3.1. Introduction 47

3.2. Study history 47

3.3. Study background and logistics 48

3.3.1. The Oxford University ORFAN Study team and facilities 48

3.3.2. Funding 49

3.3.3. Ethical considerations 49

3.3.4. Site entry into ORFAN 50

3.3.5. Patient and public involvement 51

3.4. Methods 54

3.4.1. Study design 54

3.4.2. Patient populations 54

3.4.3. Study hypothesis 54

3.4.4. Study objectives 55

3.4.5. Participant selection and enrolment 55

4.5.6. Study site patient list & ORFAN identifiers 56

3.4.7. Study data flow outline 56

3.4.8. Study data collection 58

3.4.9. ORFAN sites and patient contribution in the UK 59

3.4.10. ORFAN Sites and Patient contribution from outside the UK 62

3.4.11. Definitions of clinical risk factors and outcomes 64

3.4.12. Safety outcomes 65

3.4.13. Study monitoring 65

3.5. Discussion 65

3.5.1. Context of the ORFAN Study 66

3.5.2. Evidence and technology gap filled by ORFAN 66

3.5.3. Strengths of the ORFAN Study 67

3.5.4. Limitations of the ORFAN Study 67

3.6. Conclusion 67

3.7. Personal contribution 67

Chapter 4. Automated segmentation of CCTA scans 69

4.1. Introduction 69

4.2. Methods 69

4.2.1. Summary of process for technical development 69

4.2.2. Manual segmentation of CCTA 70

4.2.3. Experimentation of peri-LA segmentation 71

4.2.4. Selection of peri-LA segments 72

4.2.5. Whole heart manual segmentation 74

4.2.6. The extraction of EAT volume from the heart segmentation 76

4.2.7. Intra- and inter-observer variability testing for whole heart segments and epicardial adipose tissue volumes 77

4.2.8. Inter-reader variability testing for left atrium and peri-left atrium segments 78

4.2.9. Formalisation of manual segmentation methods 79

4.2.10. The extraction of radiomic features from segments 79

4.2.11. Study populations for technology development and validation 80

4.2.12. Challenging patient populations 83

4.2.13. CCTA acquisition settings and exclusions 84

4.2.14. The automated segmentation development process 84

4.2.15. Manual segmentation of training dataset 86

4.2.16. The deep-learning process 86

4.2.17. Statistical analysis 88

4.3. Results 89

4.3.1. Preliminary validation of whole heart segmentation 89

4.3.2. Internal validation of whole heart segmentation and EAT volume 90

4.3.3. External validation of whole heart segmentation and EAT volume 91

4.3.4. Validation of the automated model for EAT volume quantification in challenging clinical populations 93

4.3.5. Internal validation of the automation of the LA and peri-LA segmentations 95

4.3.6. External validation of the automation of the LA and peri-LA segmentations 97

4.3.7. Validation of the automated LA and peri-LA segmentation in challenging clinical populations 99

4.4. Discussion 101

4.5. Conclusion 102

4.6. Personal contribution 102

Chapter 5. EAT and cardiovascular disease outcomes 104

5.1. Introduction 104

5.2. Methods 105

5.2.1. Overall study design 105

5.2.2. Study populations 105

5.2.3. Coronary computed tomography angiography technical parameters 107

5.2.4. Study ethics 107

5.2.5. Statistical analysis 107

5.3. Results 109

5.3.1. Cross-sectional clinical correlations 111

5.3.2. Longitudinal EAT volume clinical correlations 111

5.3.3. A single predictive value of EAT volume 114

5.3.4. EAT volume and post-operative atrial fibrillation risk 116

5.3. Discussion 119

5.4. Conclusion 120

Chapter 6. ORFAN Arm 4 Study descriptive results & stroke outcomes 121

6.1. Introduction 121

6.2. Methods 122

6.2.1. ORFAN Arm 4 Sites 122

6.2.2. Data sources 122

6.2.3. Data availability 122

6.2.4. Data assembly 123

6.2.5. Statistical analysis 124

6.3. Results 126

6.3.1. CCTA patients at each site 126

6.3.2. Age at time of CCTA 129

6.3.3. Gender 131

6.3.4. Ethnicity 132

6.3.5. Cardiovascular disease risk factors 133

6.3.6. Stroke outcomes 135

6.4. Discussion 138

6.5. Conclusion 140

Chapter 7: Pilot radiomic stroke risk algorithm 141

7.1. Introduction 141

7.2. Methods 142

7.2.1. Study design 142

7.2.2. Outcome data 142

7.2.3. Automated CCTA analysis 142

7.2.4. Radiomic characterisation and machine learning modelling 143

7.2.5. Calculation of the CHA_2DS_2-VASc Score 146

7.2.6. Statistical analysis 146

7.3. Results 147

7.3.1. Internal testing of the pilot stroke risk prediction signature. 151

7.3.2. External testing of the pilot stroke risk prediction signature. 151

7.4. Discussion 152

7.5. Conclusion 155

Chapter 8: Limitations and future directions 157

8.1. Technical considerations 157

8.2. CCTA versus other risk assessment technologies 158

8.3. Artificial intelligence methods 159

8.4. Ongoing study plans and imminent research objectives 159

8.4.1. ORFAN Arm 4 Study priorities 159

8.4.2. CCTA analysis priorities 160

8.4.3. Advanced stroke risk prediction signatures – radiotranscriptomic risk 161

8.4.4. Data methodology research priorities 163

8.5. Long term research plans 163

Chapter 9: Conclusions 164

9.1. The ORFAN Arm 4 Study 164

9.2. Automated EAT volume assessment 165

9.3. Stroke risk prediction 166

References 168

Appendix 192

Appendix 1. Radiomic definitions 192

Appendix 2. ORFAN Study approvals 208

Appendix 3. ORFAN Study patient and public involvement documentation 227

Appendix 4. ORFAN Study protocol - version 10 242

Appendix 5. ORFAN Arm 4 Study variables 293

Appendix 6. ORFAN Study investigators and affiliations 339

Appendix 7. Disease and risk factor ICD-10 groupings 342

Acknowledgements

Oxford, December 16, 2022

It is through the **Rhodes Trust** that my whole Oxford experience became possible. Thank you to the Trust, its Administrators, and its benefactors for this incredible gift and for empowering me so significantly over the last 4 years.

The research achieved in this project would not have been possible without funding from the **British Heart Foundation**. Thank you to the Foundation and all its donors for trusting that our work is important and worthy of support.

To **Professor Charis Antoniades** – thank you for the years of effort put into my work at Oxford, and for the guidance you have provided. We have created a study that will answer important research questions for many years to come and I am so glad that I was able to work in your laboratory and gain such important experience under your tutelage.

To **Professor Keith Channon** - thank you for your assistance and sage council throughout my time as a student. Your professionalism, clinical expertise and deep knowledge is a powerful inspiration.

My life in Oxford commenced with the assistance of **Evangelos Oikonomou** whom taught me many things that I still rely upon today – from practical research skills through to the important soft skills required to navigate life in a DPhil program and in Oxford generally. I am very thankful for your friendship and assistance as you finished your own DPhil.

The technical aspects of this project were expertly supported by **Muhammad 'Musib' Siddique**. I am very thankful for your friendly professionalism, frequent assistance and deep knowledge in AI.

I am also very thankful to the team at **Caristo Diagnostics Ltd** for their generous support of my project. They contributed to critical infrastructure for this project including data transfer, data management and the development of deep-learning systems, all of which made this project possible.

To my fellow **DPhil colleagues** – Christos Kotanidis, Laura Klüner, Murray Polkinghorne, Cheng Xie, Parijat Patel, Michail Mavrogiannis, Ken Chan and Rafail Kotronias. Thank you for sharing this experience with me and for your assistance in my project along the way.

To my **friends and colleagues in the Antoniades Lab** – in no particular order: Lucrezia Volpe, Maria Lyasheva, Sheena Thomas, Alexios Antonopoulos, Ileana Badi, Amy Pritchard, Tom Halborg, Edmarcia De Araujo, Yogi Sohan, Elizabeth Wahome, Nigar Syed, Pete Tomlins, Jagat Chauhan, Lucy

Kingham and John Denton. Thank you for the opportunity to collaborate and have fun along the way!

Thanks go to the **administrative staff in the Division of Cardiovascular Medicine** and the Radcliffe Department of Medicine who assisted me with grant applications, expense claims and all manner of logistics, especially Alina Florea, Lucy Lattimore, Sally Lacey, Harry Clarke, Stephen Woodhouse and Heidi Crook.

My DPhil required collaboration and the creation of professional relationship with academics and clinicians across the world. I am thankful to all who were willing to engage on the ORFAN Study, especially the **local investigators** from the first 7 sites outside of Oxford that joined our study. I also owe thanks to the SCOT-HEART Trial team, led by Professor David Newby.

To my **friends in Oxford** - I have been so enriched by your care and support. I am thankful for all the opportunities that brought us together, be it through Rhodes, Holywell Manor, BCBC, OUARFC, OUANZ, or just through being in the same place at the same time.

To my **parents** – thank you for your love that allowed, and continues to allow, me to flourish. Everything I have achieved was made possible through this.

To **Ashleigh** - thank you for your abundant love throughout our time in Oxford. We have many adventures ahead of us.

Dedication

The work of this thesis is dedicated to my love, Ashleigh, as nothing is more important.

Sincere thanks are extended to The Rhodes Trust for providing me the opportunity of a lifetime.

COVID-19 Statement

My DPhil commenced in October 2018. The global COVID-19 pandemic impacted upon my DPhil project from February 2020 through until its conclusion. It impacted the timeline for the completion of my project and the work that was possible.

The most significant impact of the pandemic was the delay in data collection for the ORFAN Arm 4 Study. This study relied upon the collaboration of clinicians and research staff across over a dozen NHS sites and internationally. Prior to the pandemic we intended to have data collected for the first 8 NHS sites by mid-2021. However, this was only recently realised in November 2022. Major reasons for this were: clinical research staff allocation was reduced to zero at some of our collaborating sites; resources at NHS Digital were re-assigned to COVID-19 surveillance and NHS Test & Trace services leading to months long delays in the processing of data access applications and the processing of the data production following approval; staff at national clinical outcomes registries such as NICOR and SSNAP were re-assigned, furloughed or were working from home without access to data servers which led to months-long delays in data access requests; and IT and PACS staff across our collaborating sites were unable to prioritise the ORFAN Study project in the face of overwhelming demand for COVID-19 related service priorities and research tasks, often compounded by reduced staff numbers.

The result of this impact is that less clinical data (i.e. clinical risk factors, medications, pathology results and disease outcomes) and CT scans are currently available for the ORFAN Arm 4 Study cohort than was intended at the commencement of this thesis project.

Other impacts due to the pandemic included significant delays in the procurement of required research equipment such as the graphics processing units (2 months behind usual service) and servers (4-6 months behind usual service) for automated CT scan segmentation training and high through-put analysis of scans. This significantly delayed the developmental aims of this project.

I was also limited by a delay in funding through the pausing of BHF research funding at the start of the pandemic. We were awarded a BHF Translational Grant in 2019 to fund much of the work within this thesis, however this grant did not commence until September 2020.

Abstract

The Development of Novel Artificial Intelligence Tools for the Early Detection of Cardiovascular Disease Using Computerised Tomography

Henry William West

A thesis submitted for the degree of
Doctor of Philosophy in Medical Sciences at the University of Oxford

The use of big data approaches in medicine promises to fill technology gaps in the translation of basic biological science into clinically actionable tools for the benefit of patient care. This thesis is focused on the development of necessary clinical cohorts and tools to facilitate the translation of our mechanistic understanding of the role of epicardial adipose tissue (EAT) and the atrial myopathy in adverse cardiovascular disease events. This thesis creates novel risk assessment tools that could facilitate timelier implementation of care for those found to benefit.

Visceral obesity is a driver of cardiovascular disease and has been implicated in multiple leading causes of morbidity and mortality including coronary artery disease, atrial fibrillation (AF) and heart failure. EAT is becoming recognised as a critical regulator of myocardial and vascular biology and serves as a surrogate measure of visceral obesity beyond its anatomical relationship with the heart. The atrial myopathy is closely related to EAT but is the specific disease process that drives the risk of AF and stroke through both myocardial and EAT dysfunction.

This thesis first outlines the development of the largest clinical coronary computed tomography angiography (CCTA) cohort study – The ORFAN Arm 4 Study. This study has now enrolled some 90,000 patients from across eight NHS Trusts to provide the necessary baseline clinical data, CCTA scans, and outcomes data to develop novel automated risk technologies. This thesis conceptualised, developed, and externally validated artificial intelligence powered technologies. These include the automated segmentation and radiomic feature extraction of both the whole heart EAT volume, and the tissues surrounding the left atrium of the heart for the detection of the atrial myopathy.

In the largest cohort to date, I show that EAT volume is predictive of key cardiovascular outcomes including myocardial infarction and stroke, but also all-cause and non-cardiac mortality, suggesting that dysfunctional EAT may have far reaching effects beyond the heart. Finally, I demonstrate the successful development and validation of a pilot radiomic risk score for the prediction of ischaemic stroke, laying the foundation for personalised assessment of AF and stroke risk via non-invasive detection of the atrial myopathy.

List of Publications and Presentations

All listed publications, patents, awards, and presentations that occurred following matriculation in Michaelmas Term 2018 (13 October 2018)

Publications since matriculation

2023

West HW, Lyasheva M, Wahome E, Kingham L, Pritchard A, Siddique M, Thomas S, Volpe L, Neubauer S, Channon KM, Antoniades C. The Oxford Risk Factors and Non-Invasive Imaging (ORFAN) Study. *BMJ Open*. 2023. *In preparation.*

West HW, Dangas K, Antoniades C. State-of-the-Art review: Advances in Clinical Imaging of Inflammation. *Journal of the American College of Cardiology: Basic to Translational Science*. 2023. *In preparation*.

2022

West HW, Siddique M, Williams MC, Volpe L, Desai R, Lyasheva M, Thomas S, Kotanidis CP, Dangas K, Tomlins P, Mahon C, Kardos A, Adlam D, Rodrigues JCL, Graby J, Shirodaria C, Deanfield J, Mehta NN, Neubauer S, Channon KM, Desai MY, Nicol ED, Newby DE, Antoniades C. Deep-Learning for Epicardial Adipose Tissue Assessment with Computed Tomography: Implications for Cardiovascular Risk Prediction. *Journal of the American College of Cardiology: Cardiovascular Imaging*, 2022; *Accepted 17 November 2022*

Polkinghorne M, **West HW**, Antoniades C. The role of perivascular fat in the diagnosis and prognosis of atherosclerosis. *CardioMetabolic Syndrome Journal*. 2022;3:e1.

Kotanidis CP, …, Antoniades C; COMBAT Consortium; **ORFAN investigators**. Constructing custom-made radiotranscriptomic signatures of vascular inflammation from routine CT angiograms: a prospective outcomes validation study in COVID-19. *Lancet Digital Health*. 2022 Oct;4(10):e705-e716

Antoniades C & **West HW**. ESC CVD Prevention Guidelines 2021: Improvements, controversies, and opportunities. *Cardiovascular Research*. 2022 Jan 29;118(2):17-19

2021

West HW & Antoniades C. Coronary CT angiography as a 'one-stop shop' to detect the high-risk plaque and the vulnerable patient. *European Heart Journal*. 2021 Oct 1;42(37):3853-3855

2020

West HW & Antoniades C. Imaging and targeting coronary artery inflammation. *Antioxidants & Redox Signaling 2020;10:1089*

West HW & Antoniades C. The Fat Attenuation Index in HIV-infected patients. *AIDS* 2020;34:171–172.

2019

Oikonomou EK, **West HW**, Antoniades C. Computed tomography: Assessment of coronary inflammation and other plaque features. *Arteriosclerosis, Thrombosis, and Vascular Biology 2019;39(11):2207-19*

Patents since matriculation

Antoniades C (40%), Antonopoulos A (40%), **West HW** (20%). 'Radiomic signature of an epicardial region'. PCT/GB2019/053058. Granted in multiple jurisdictions. Managed by Oxford Sciences Enterprises PLC.

Awards since matriculation

September 2022: **BCS Best of the Best Clinical Abstract Prize (1st place),** British Cardiovascular Society (BCS) Annual Conference 2022 (Manchester, UK): West HW et al. '*Automated deep-learning quantification of epicardial adiposity on cardiac CT predicts atrial fibrillation risk immediately following cardiac surgery and long term'.*

September 2021: **BAS Best Early Career Oral Presentation Prize,** British Atherosclerosis Society (BAS) Annual Scientific Meeting (Oxford, UK): West HW et al. *'Deep learning detection of epicardial adiposity on cardiac CT predicts mortality and cardiac events'*.

Presentations since matriculation

2022

West HW, Siddique M, Volpe L, Desai R, Polkinghorne M, Lyasheva M, Dangas K, Tomlins P, Mitchell A, Kardos A, Casadei, B, Channon K, Antoniades C. 'Automatic Epicardial Adiposity Assessment For Atrial Fibrillation Risk In Severe Coronary Atherosclerosis'. Poster Presentation. **British Atherosclerosis Society** Annual Meeting, Oxford, UK, 8th September 2022.

West HW, Siddique M, Volpe L, Desai R, Lyasheva M, Dangas K, Tomlins P, Mitchell A, Kardos A, Casadei B, Channon KM, Antoniades C. 'Automated deep learning quantification of epicardial adiposity on cardiac CT predicts atrial fibrillation risk immediately following cardiac surgery and long term' Oral Presentation. **European Society of Cardiology Congress**, Barcelona, August 26th 2022.

West HW, Siddique M, Volpe, L, Desai R, Lyasheva M, Dangas K, Tomlins P, Mitchell A, Kardos A, Casadei B, Channon K, Antoniades C. 'Automated deep-learning quantification of epicardial adiposity on cardiac CT predicts atrial fibrillation risk immediately following cardiac surgery and long term'. Oral Presentation. **British Cardiovascular Society** Annual Meeting, Manchester, UK, 6th June 2022.

2021

West HW, Siddique M, Volpe L, Desai R, Lyasheva M, Dangas K, Shirodaria C, Neubauer S, Channon K, Desai MY, Newby DE, Rodrigues JCL, Adlam D, Nicol ED, Antoniades C. 'Automated quantification of epicardial adipose tissue on CCTA via deep-learning detection of the pericardium: Clinical implications'. Oral Presentation. **European Society of Cardiology Congress**, Virtual, 27th August 2021.

West HW, Williams MC, Siddique M, Volpe L, Desai R, Lyasheva M, Dangas K, Tomlins P, Oikonomou EK, Shirodaria C, Neubauer S, Channon K, Rodrigues JCL, Nicol ED, Newby D, Antoniades C. 'Deep learning detection of epicardial adiposity on cardiac CT predicts mortality and cardiac events' Oral Presentation. **British Atherosclerosis Society** Annual Meeting, Virtual, 17th September 2021.

West HW, Williams MC, Siddique M, Volpe L, Tomlins P, Shirodaria C, Neubauer S, Channon KM, Newby DE, Antoniades C. 'Epicardial adipose tissue volume measured by an automated computed tomography deep learning network predicts mortality and cardiovascular events'. Oral Presentation. **Cardiology Update 2021**, Davos, Switzerland, 24th October 2021

West HW, Williams, M, Siddique M, Volpe L, Desai R, Lyasheva M, Dangas K, Shirodaria C, Neubauer S, Channon K, Desai M, Rodrigues J, Adlam D, Nicol E, Newby D, Antoniades C. 'Epicardial Adiposity Measured by a Deep Learning Network, Predicts Mortality and Cardiovascular Events in Patients Undergoing Cardiac CT'. Oral Presentation. **American Heart Association** Annual Scientific Sessions, Boston, USA, 14th November 2021

West HW, Siddique M, Volpe L, Desai R, Lyasheva M, Dangas K, Shirodaria C, Neubauer S, Channon K, Desai M, Newby D, Rodrigues J, Adlam D, Nicol E, Antoniades C. 'Applying a Deep-Learning Approach for Automated Quantification of Epicardial Adipose Tissue on Coronary Computed Tomography Angiography in Challenging Clinical Populations'. Oral Presentation. **American Heart Association** Annual Scientific Sessions, Boston, USA, 14th November 2021

List of Figures

Figure 1.1.	The proposed disease spectrum of atrial myopathy	***7***
Figure 1.2.	The pathophysiology of the atrial myopathy	***9***
Figure 1.3.	Role of EAT deposits in coronary artery disease and atrial fibrillation	***12***
Figure 1.4.	Gradient of adipocyte size and FAI around the human coronaries in the presence or absence of coronary atherosclerosis	***17***
Figure 1.5.	Illustration of the major components of the inside to outside interplay between the atrial wall and EAT	***18***
Figure 1.6.	Association of atrial genes with pilot Atriomic Risk Algorithm for ischaemic stroke risk.	***19***
Figure 1.7.	Advantages and disadvantages of imaging modalities to characterise adipose tissue	***21***
Figure 2.1.	John Radcliffe Hospital CT Scanner	***32***
Figure 2.2.	The Manor Hospital CT Scanner	***33***
Figure 2.3.	The flow of CCTA data from the NHS Trust environment to the Oxford ORFAN Core Lab via the Cimar gateway	***38***
Figure 2.4.	Creating a gray level co-occurrence matrix	***40***
Figure 2.5.	Radiomic characterization of textural features	***42***
Figure 2.6.	Wavelet decomposition of an axial slice from computed tomography scan	***44***
Figure 2.7.	Agatston coronary calcium scoring	***45***
Figure 2.8.	Coronary segmentation based on the recommendations of the Society of Cardiovascular Computed Tomography	***46***
Figure 3.1.	Overall logistical flow and data flow within ORFAN Arm 4 Study in England	***57***
Figure 3.2.	UK sites of the ORFAN Study Arm 4	***60***
Figure 3.3.	International sites of the ORFAN Arm 4 Study	***63***
Figure 4.1.	Summary of development steps for the automated segmentation tools	***70***
Figure 4.2.	CCTA segmentations selected for automation	***73***
Figure 4.3.	Screenshot of 3DSlicer work package demonstrating a CCTA scan with manual segmentation on screen	***73***
Figure 4.4.	Manual segmentation of the whole heart	***75***
Figure 4.5.	Image of CaRi-Heart user work view of manual whole heart segmentation	***75***
Figure 4.6.	Automated segmentation of epicardial adipose tissue extracted from a whole heart segmentation	***76***
Figure 4.7.	Inter-reader variability testing of the whole heart volume segmentation	***77***
Figure 4.8.	Inter-reader variability testing of the EAT volume segmentation	***78***
Figure 4.9.	Inter-reader variability testing of the left atrium and peri-left atrium segmentations	***79***
Figure 4.10.	Commonly occurring scan artefacts encountered in clinical CCTA	***84***
Figure 4.11.	Schematic representation of the development pathway for automated segmentations	***85***
Figure 4.12.	Block diagram of the Residual U-Net based CNN architecture employed for all automated DLN	***87***
Figure 4.13.	Preliminary validation of machine vs human whole heart segment volume assessment prior to feedback learning	***89***
Figure 4.14.	Preliminary validation of machine vs human EAT volume assessment prior to feedback learning	***90***

Figure 4.15.	Internal validation of the whole heart segmentation deep-learning model	***91***
Figure 4.16.	Internal validation of the EAT volume model	***91***
Figure 4.17.	External validation of the whole heart segmentation deep-learning model	***92***
Figure 4.18.	External validation of the EAT volume model	***92***
Figure 4.19.	External validation of automated EAT segmentation in challenging clinical CCTA cohorts	***94***
Figure 4.20.	Internal validation of whole left atrium and peri-left atrium automated segmentations	***96***
Figure 4.21.	Internal validation of automated EAT volume segmentation for peri-left atrium segments	***97***
Figure 4.22.	External validation of whole left atrium and peri-left atrium automated segmentations	***98***
Figure 4.23.	External validation of automated EAT volume segmentation for peri-left atrium segments	***99***
Figure 5.1.	Cross-sectional and longitudinal associations between EAT volume and clinical outcomes in the SCOT-HEART trial	***113***
Figure 5.2.	High EAT volume increases risk of major adverse events when assessed with a single cut-point	***115***
Figure 5.3.	Overlayed distribution of SCOT-HEART and the Coronary Artery Bypass Surgery population in the AdipoRedOx study by EAT volume	***117***
Figure 5.4.	Prognostic value of EAT volume for post-operative AF	***118***
Figure 6.1.	Clinical CCTA scans per year in the ORFAN Arm 4 Study	***127***
Figure 6.2.	Age histogram of patients at time of first clinically indicated CCTA	***129***
Figure 6.3.	Median age at time of clinical CCTA within the ORFAN Arm 4 Study	***131***
Figure 6.4.	Diagnosis of cardiovascular disease risk factors each year within the ORFAN Arm 4 Study	***134***
Figure 6.5.	Stroke events prior to and following the clinical CCTA within the ORFAN Arm 4 cohort	***136***
Figure 7.1.	Radiomic feature selection process overview	***143***
Figure 7.2.	Stability analysis of LA and peri-LA volume of interests radiomic features	***144***
Figure 7.3.	Recursive feature elimination	***145***
Figure 7.4.	Radiomic features included in pilot radiomic stroke risk signature	***146***
Figure 7.5.	Study design and flowchart	***148***
Figure 7.6.	Internal testing of the pilot radiomic stroke risk score	***151***
Figure 7.7.	External testing of the pilot radiomic stroke risk score in retraining analysis	***152***
Figure 8.1.	The methodological approach to the development of the radiotranscriptomic Atriomic Risk Algorithm	***162***

List of Tables

Table 1.1.	Radiomic features available for extraction from computed tomography images	***23***
Table 2.1.	Technical review and quality assurance of obtained coronary CT angiograms in the ORFAN Arm 4 Study	***39***
Table 3.1.	Patient and Public Involvement Survey Results for ORFAN Study Protocol Amendment to include Arm 4	***52***
Table 3.2.	Formal Study Objectives of the ORFAN Study	***55***
Table 3.3.	ORFAN Study Arm 4 patient identifier key	***56***
Table 3.4.	Sources of clinical data in ORFAN Study Arm 4	***59***
Table 3.5.	Active ORFAN Arm 4 Study Sites in England	***61***
Table 3.6.	International sites of the ORFAN Arm 4 Study	***64***
Table 4.1.	Manual CCTA segmentations for development of automated segmentation methods	***72***
Table 4.2.	Radiomic feature databases for segmented CCTA regions of interest	***80***
Table 4.3.	ORFAN Arm 4 study populations used for deep-learning model development and validation	***82***
Table 4.4.	Validation of left atrium and peri-LA automated segments in commonly occurring challenging patient groups	***100***
Table 5.1.	Demographics and scan characteristics of external clinical cohorts	***110***
Table 5.2.	Outcomes from the prospective clinical cohorts	***111***
Table 6.1.	ORFAN Arm 4 Sites included in first tranche of data	***122***
Table 6.2.	ORFAN Arm 4 Study CCTA patients by site and year	***128***
Table 6.3.	Age of clinical CCTA patients by NHS Trust	***130***
Table 6.4.	Gender split of ORFAN Arm 4 CCTA patients across NHS Trusts	***132***
Table 6.5.	High level ethnic categories of the ORFAN Arm 4 CCTA patient cohort across NHS Trusts	***133***
Table 6.6.	ORFAN Arm 4 cohort cardiovascular disease risk factors	***133***
Table 6.7.	ORFAN Arm 4 cohort demographics and risk factors at time of CCTA	***137***
Table 7.1.	ORFAN Arm 4 cohort demographics for the pilot radiomic risk score for ischaemic stroke	***150***

List of Abbreviations

ACC	American College of Cardiology
AF	Atrial fibrillation
AHA	American Heart Association
AI	Artificial intelligence
AMI	Acute myocardial infarction
AUC	Area under the curve
BMI	Body mass index
CABG	Coronary artery bypass graft surgery
CAC	Coronary artery calcium
CAD	Coronary artery disease
CCC	Lin's Concordance correlation coefficient
CCS	Coronary calcium score
CCTA	Coronary computerised tomography angiography
CMR	Cardiac magnetic resonance
CRF	Case record file
CT	Computerised tomography
DLN	Deep-learning network
DOB	Date of birth
DPIA	Data protection and impact assessment
EAT	Epicardial adipose tissue
ECG	Electrocardiogram
ESC	European Society of Cardiology
GLCM	Gray Level Co-occurrence Matrix
GLDM	Gray Level Dependence Matrix
GLRLM	Gray Level Run Length Matrix
GLSZM	Gray Level Size Zone Matrix
GPU	Graphical processing unit
GUI	Graphical user interface
HDL	High-density lipoprotein cholesterol
HES	Hospital episode statistics
HR	Hazard ratio
HRA CAG	Health Research Authority Confidentiality Advisory Group
Hs-CRP	High-sensitivity C-reactive Protein
HU	Hounsfield units
ICC	Intraclass correlation coefficient
LA	Left atrium
LAA	Left atrial appendage
LAD	Left anterior descending coronary artery
LCX	Left circumflex coronary artery
LDL	Low density lipoprotein cholesterol
LMS	Left main stem coronary artery
MI	Myocardial infarction
MRN	Medical record number

NDOP	National Data Opt-Out
NGTDM	Neighbouring Gray Tone Difference Matrix
NHS	National Health Service
NICOR	National Institute for Cardiovascular Outcomes Research
OR	Odds ratio
ORFAN	Oxford Risk Factors And Non Invasive Imaging Study
OXACCT	Oxford Academic Cardiovascular CT
PCI	Percutaneous coronary intervention
PCSK9	Proprotein convertase subtilisin/kexin type 9
PET	Positron emission tomography
RCA	Right coronary artery
REC	Research ethics committee
ROC	Receiver operator characteristic
ROI	Region of interest
SCOT-HEART	Scottish COmputed Tomography of the HEART Trial
SD	Standard deviation
SSNAP	Sentinel Stroke National Audit Programme
VOI	Volume of interest

1

Chapter 1. Introduction

This introductory chapter provides an overview of the relevant background to the novel disease risk assessment tools presented within this thesis, namely the automated assessment of epicardial adipose tissue (EAT) on computed tomography (CT), the biology of EAT, and the use of CT radiomics for the assessment of atrial myopathy risk.

The introduction includes a summary of the current understanding of the atrial myopathy and how this relates to atrial fibrillation (AF) and subsequent stroke risk. It also includes relevant background as to how big data approaches can be harnessed for the development of novel cardiovascular risk assessment tools is included. Finally, the scientific hypotheses and aims that are addressed within this thesis are presented.

Note: Parts of this chapter are based on previously published material[1]

1.1. The burden of cardiovascular diseases including atrial fibrillation and stroke

The burden of cardiovascular diseases as a broad group that includes coronary artery disease (CAD) cerebrovascular disease, and AF is far larger than any other grouping of diseases when arranged by bodily system. Morbidity related to CAD and cerebrovascular disease have declined over the last 50 years in the UK. Despite this, CAD continues to be the leading cause of death for males within the UK, and when combined with cerebrovascular death, account for almost half of all deaths[2]. For women, CAD and cerebrovascular disease are respectively second and third to dementia as the leading cause of death[2]. Other Western nations report similar figures. Within the United States heart disease and cerebrovascular events each remain a leading cause of death ahead of cancers, respiratory disease, and injuries[3].

AF is the most common cardiac arrhythmia. It is estimated to have a prevalence in the UK of at least 2.5% of the total adult population[4]. Some 1.4 million people in England alone are estimated to have AF and at least one third of these are undiagnosed. AF contributes to at least 1 in 5 strokes and strokes that are related to AF are typically more severe and carry a higher morbidity and mortality than strokes without AF[4].

In the UK, there are over 100,000 stroke events each year, of which over 35,000 are fatal. In England, one in six people will have a stroke event in their lifetime[5]. This is despite significant advances in prevention, community awareness, earlier treatment accessibility, and more advanced treatment

options. Beyond mortality, stroke is associated with severe morbidity for survivors. 65% of stroke survivors who are hospitalised will leave hospital with a disability[5]. The economic costs of disability due to stroke are enormous. Recent analysis of UK stroke event data suggests the aggregate societal cost of stroke is £26 billion per year, not considering loss of economic output[6].

The continued high prevalence and enormous socioeconomic burden of cardiovascular diseases highlights the urgent need for improved primary and secondary prevention strategies, and specifically new strategies for the identification of high-risk individuals who may go on to suffer from heart disease, be it CAD, AF, or stroke. Any novel risk stratification approaches that can provide more personalised and cost-effective practices in cardiovascular medicine will contribute to reducing the burden of disease and the costs of downstream care.

1.2. Cardiovascular risk stratification and prevention

1.2.1. Current approaches in cardiovascular disease prevention

Population based risk factors remain the main tools utilised for individual assessment of a patient's cardiovascular disease risk profile, for both IHD and stroke. Diagnostic data such as presence of hypertension, hypercholesterolaemia, and diabetes remain the core tenants of event risk assessment. For risk of IHD in those with no prior evidence of cardiovascular disease the latest European Society of Cardiology (ESC) CVD prevention guidelines[7] employ complex yet crude risk scores to estimate the ten-year risk of atherosclerotic CVD. The score is reliant upon gender, age, smoking status, systolic blood pressure, non-HDL cholesterol level and country of residence. Such scores are relied upon to guide subsequent clinical decisions regarding lifestyle recommendations and medical therapy. It is acknowledged that these scores are critical for current clinical practice and have been extensively tested and validated. However, significant concerns remain about the utility and acceptability of such scores. Concerns include the calibration of the scores, as they have been reported to overestimate an individuals' risk of IHD[8] and the continued resistance to the inclusion of novel technologies to enhance individual risk prediction[9], including the possible incorporation of imaging methodologies into the calculation of risk[10]. There is one example of this currently included within clinical guidelines. The coronary artery calcium score (CCS) is included as a risk modifier in patients where medical therapy (i.e., treatment with statins) remains unclear. Guidelines currently recommend a non-contrast cardiac computed tomography (CT) scan for CCS assessment in this situation. The CCS, as calculated by the Agatston method, serves as a surrogate marker of coronary artery atherosclerosis, and provides a quantitative assessment of plaque burden that has been linked to disease outcomes[11]. This is an excellent example from within the guidelines themselves as to how imaging technologies can be incorporated into guidelines for improved patient care.

The field of risk assessment for stroke is far less advanced than that for IHD. Unlike the well-publicised and adhered to guidelines released periodically by major learned societies such as the ESC and American Heart Association (AHA), stroke medicine has not advanced to the point of regular population risk assessment for stroke, except in certain high-risk populations. Risk scores that are currently used are similar in nature to the IHD risk scores – they incorporate population level risk data and are strongly influenced by demographical (i.e., age and sex) and clinical diagnostic data (i.e. presence of disease such as hypertension and diabetes). The most widespread risk assessment tools for stroke are the QRISK3 calculator[12] and the Framingham Stroke Risk Score (FSRS)[13]. Limitations of these scores are myriad. In the UK clinical context, the QRISK3 is the most common stroke risk assessment tool. Limitations of QRISK3, which relies upon 22 individual patient factors for a maximally accurate score, include that it over-predicts risk in older patients and those with multi-morbidity, and as per the IHD risk score, does not incorporate any targeted assessment of the underlying organ through imaging or pathology data beyond LDL cholesterol level. Importantly QRISK3 also fails when consideration is made of competing mortality risk – that is the risk of death of another cause which influences the stroke risk model due to presumed lack of risk of stroke in a censured cohort.

The most widespread clinical risk tool for stroke is the CHA_2DS_2-VASc score for risk of stroke in non-valvular AF[14]. The score incorporates key stroke risk factors including congestive heart failure, hypertension, age >75 years [doubled], type 2 diabetes mellitus, previous stroke, transient ischaemic attack or thromboembolism [doubled], vascular disease, age of 65-75 years, and sex as a predictive scoring model for stroke in patients with AF. The score has become the universal risk score for decisions related to the initiation of anti-coagulation therapy in those with AF and has been incorporated into clinical guidelines, with patients with one or more stroke risk factors (i.e., a CHA_2DS_2-VASc score ≥1 in men or ≥2 in women) treated therapeutically. Moreover, it has been shown that the CHA_2DS_2-VASc score can predict the incidence of stroke in some populations without AF, including those with heart failure and coronary artery disease (CAD)[15-21]. Notwithstanding this, the use of such scores for the initiation of anti-coagulation therapy for stroke risk reduction in those without AF remains highly controversial.

1.2.2. The residual cardiovascular risk

Residual cardiovascular disease risk is the disease risk that remains for an individual following optimisation of lifestyle and medical therapy. Optimal therapy is currently informed by evidence-based guidelines and these strategies have contributed enormously to the decline in cardiovascular disease mortality seen over the preceding half century. Such therapies include lifestyle modification (i.e., avoiding sedentary behaviour, smoking cessation, alcohol reduction, dietary recommendations

and stress reduction), blood pressure lowering medications, cholesterol lowering therapies, blood thinning agents (anti-platelet therapy and anti-coagulants), and coronary artery revascularisation (percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) surgery) where indicated[22,23]. These approaches to disease risk reduction continue to be the mainstay of basic cardiovascular care across the world. As already outlined, despite these therapies there remains an unacceptably high burden of mortality and morbidity due to cardiovascular and cerebrovascular disease within our communities. This has led to the concept of residual cardiovascular disease risk - risk that exists in populations with access to therapies as recommended by the existing guidelines[24]. The major reason that this residual risk remains is that the population wide guidelines do not consider individual patient level risks and specific causal pathways that are likely to be influencing disease risk(s) for a single patient. The 'one-size-fits-all' therapeutic approaches work well as a general approach for all, but there is massive potential for more personalised therapeutics and risk assessment strategies to combat the remaining residual risk.

Efforts to reduce the residual atherosclerotic disease risk for coronary artery disease have been far more advanced than those for the risk of cerebrovascular disease. Specific residual risks have been identified that are related to druggable pathways. These include inflammatory risk[25], hypercholesterolemic risk[26], hyperglycaemic risk[27], and coagulation risk[28]. Should a patient have a CAD event and at follow up be found to have diabetes mellitus, elevated cholesterol, and an elevated level of high-sensitivity C-reactive protein (hsCRP) then current guidelines will recommend strict lifestyle management, the initiation of high intensity statin-therapy, beta-blocker, anti-hypertension therapy with escalating number of agents to control blood pressure, aspirin therapy, a P2Y12 inhibitor and appropriate antihyperglycemic therapy for blood glucose control. Despite this, we know significant risk will likely remain and we now have the evidence to indicate benefit from enhanced therapeutic approaches: hypercholesteraemic risk can be further reduced with ezetimibe or a proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitor. Inflammatory risk can be reduced with the interleukin-1β-targeting monoclonal antibody canakinumab[25]. The residual hyperglycaemic risk can be reduced by a glucagon-like peptide 1 (GLP-1) agonist or a sodium/glucose cotransporter 2 antagonist[27,29]. Coagulation risk can be reduced with low-dose rivaroxaban (or other direct oral anti-coagulant) in conjunction with aspirin after the first twelve months[28]. It is through the targeting of these, and possibly other, druggable pathways that the residual risk for CAD can be reduced.

Significant hurdles must be resolved before this level of risk reduction focused therapy will be achieved outside select advanced clinics and research settings. The most pressing issue is the selection of patients for targeted therapies. Enhanced risk assessment technologies that allow

patients with active residual risk pathways (i.e., active inflammation around the coronary tree, or significant epicardial adipose tissue volume) to be identified for therapy will be the solution to the question of which patients stand to enjoy the most benefit from such therapies. Ideally any such risk assessment technology should be acceptable to the patient, non-invasive, cheap and allow longitudinal follow-up of the patient's ongoing risk through a dose-response relationship with corresponding therapeutic agent(s). Given the cost of such therapeutics (the cost-effectiveness ratio of adding canakinumab to standard of care was estimated at USD$6.4 million per quality-adjusted life-years gained[30]) it is only through select patient initiation and continuation according to effect that will allow these agents to achieve clinical acceptance.

For stroke events, the residual risk must be considered in two separate groups of patients: those with known and treated AF, and those without known AF. Those with AF for which they are anti-coagulated according to current practice guidelines do have a residual risk of ischaemic stroke events. These events are largely attributed to similar disease pathways as CAD and include residual risk through hypertension, diabetes mellitus and dyslipidaemia[31,32]. It is likely that optimisation with strategies as outlined above will benefit such patients. The risk of stroke in those without known AF is fundamentally more challenging as this group is the majority of the adult population and the majority of ischaemic strokes occur in this population group[5]. Despite this, there has been limited research into stroke risk assessment and risk reduction for those without AF. As previously outlined, the majority of stroke risk research in those without AF is focused on risk assessment via risk scores usually used in those with AF such as the CHA_2DS_2-VASc score, with some success but little clinical application as yet[14,33]. There is enormous potential for the discovery and targeting of risk pathways that contribute to ischaemic stroke risk in those without AF, with the major focus area being the identification of patients with atrial myopathy that predisposes them to both AF and stroke[34].

1.3. The atrial myopathy

AF is the most common arrhythmia in the world and is associated with a 5-fold increased risk of stroke. AF-related strokes are typically attributed to the formation of blood clots within or adjacent to the left atrial appendage (LAA) because of local blood stasis. These clots then dislodge and embolize to the brain. This classical description of the risk of embolic stroke in AF has been challenged in recent years through observations made in clinical trials; if stasis in the LAA or LA is the primary cause of stroke then it is maintenance of sinus rhythm that would logically prevent such stasis and hence prevent embolism. Over the preceding twenty years this has repeatedly been proven not to be the case, with two important randomized clinical trials, RACE (Rate Control versus Electrical Cardioversion for Persistent Atrial Fibrillation Study)[35] and AFFIRM (Atrial Fibrillation

Follow-up Investigation of Rhythm Management)[36] both demonstrating that rhythm control does not equate to a lower risk of embolic events.

More recent data has shown that there is little to no temporal relationship between atrial rhythm and clinical AF or stroke events[37,38]. This has been proven through the use of continuous atrial monitoring, such as in the ASSERT (Asymptomatic Atrial Fibrillation and Stroke Evaluation in Pacemaker Patients and the Atrial Fibrillation Reduction Atrial Pacing) trial, where 2580 patients with cardiac implantable devices were followed over time. Only 8% of patients having ischaemic stroke had fibrillation events within the 30-day window prior to the stroke, and 16% had their first atrial fibrillation event after their first stroke[39]. These data are important for understanding the disease spectrum that can lead to AF and stroke: the lack of a temporal relationship between AF and stroke indicates that other factors are likely to be important contributors to the occurrence of stroke and that the presence of AF is not necessary for stroke events to occur.

The clinical evaluation of patients with AF, or at risk of AF, predominantly focuses on upstream risk factors such as hypertension, diabetes, heart failure, and valvular heart disease[40]. There is minimal clinical focus on the organ specifically implicated in the risk of AF and the subsequent risk of ischaemic stroke – the left atria of the heart. Systematic evaluation of the left atrium (LA) is not commonly actioned bar for echocardiographic assessment of the LA size and/or volume when clinician initiated. Notably, the universally used risk score for stroke risk in non-valvular AF, the CHA_2DS_2-VASc score, does not include atrial size or any direct or surrogate measure of atrial pathophysiology. Evidence is growing that atrial myopathy can not only lead to blood stasis within the LA and LAA but that it can also lead to endothelial and/or endocardial dysfunction and the hypercoagulable state. It is through this lens that the atrial myopathy should be considered - that AF may not be the root cause of stroke, but rather, a marker that the atria are diseased. As schematically demonstrated in Figure 1.1, the development of an atrial myopathy predisposes to AF[41], and thromboembolic events including stroke[42,43], but AF is not necessary for stroke. Patients may be at risk of AF due to atrial myopathy, but the first clinical manifestation may be an ischaemic stroke. The clinical ramifications of the early detection of atrial myopathy in an individual patient are obvious – early opportunity for therapeutic intervention to decrease the risk of both AF (through direct intervention in the myopathy pathways) and stroke (through anti-coagulation).

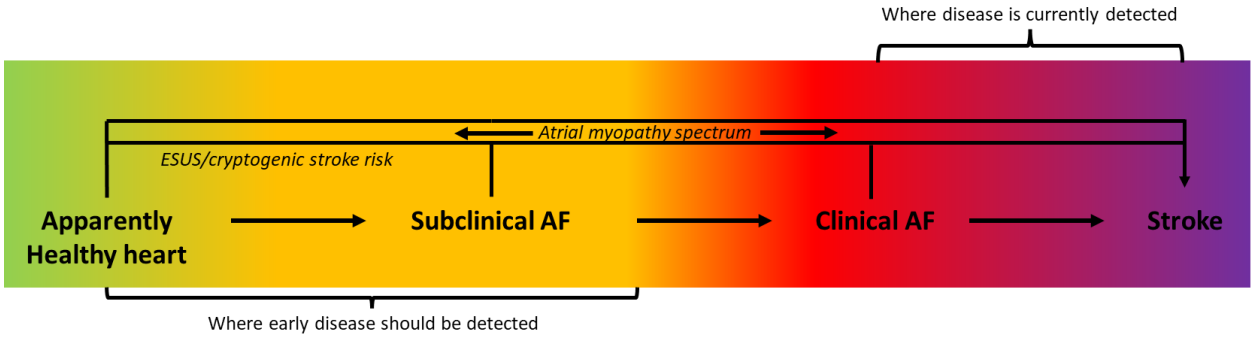


Figure 1.1. The proposed disease spectrum of atrial myopathy. *Atrial myopathy can lead to AF and stroke. AF is not required for stroke to occur, but it is likely that a patient will be predisposed to fibrillation. Currently disease is detected only at the stage of clinically apparent AF, frequently after a long period of non-diagnosis, or at the time of a stroke event. If the atrial myopathy can be detected earlier in the disease spectrum the risk of undiagnosed AF and stroke events could be significantly reduced through earlier monitoring and intervention if indicated.*

1.3.1. Atrial myopathy biology

The concept of an atrial myopathy was first described by Zipes in 1997, when he observed that AF can lead to remodelling of the atria[44]. Since that time, enormous progress has been made in identifying the underlying causal pathways implicated in atrial myopathy. The key processes are aging, inflammation, oxidative stress, and mechanical stretching of the atria. Through these processes the myopathic changes alter the functioning of the myocardial electrophysiology and the cardiac autonomic nervous system and can lead to structural changes within the atria through fibrosis.

The relationship between chronic systemic inflammatory disorders and atrial myopathy is also revealing. Patients with rheumatoid arthritis are at a 40% greater risk of AF, even following adjustment for traditional AF risk factors[45]. Similar elevated risk profiles are seen in patients with psoriasis[46], systemic lupus erythematous[47], ankylosing spondylitis[48] and inflammatory bowel disease[49]. It has been shown that many patients with these disorders have structural or functional evidence of atrial myopathy[50,51]. The role of metabolic dysfunction in the risk of AF is also relevant, with disorders that are characterised by adipose tissue inflammation known to lead to an increased risk of AF. Obesity itself is well known to increase both the incidence and prevalence of AF[52], with bariatric surgery shown to reduce the risk of future AF. Interestingly, type 2 diabetes mellitus carries an increased risk of AF that is proportional to the degree of glycaemic (non)control[53]. It is well describe that the adipose tissue inflammation and insulin resistance associated with conditions such as the metabolic syndrome[54] and non-alcoholic fatty livery disease[55] lead to increased risk of AF. Hormonal diseases that promote adipogenesis such as hypothyroidism[56] and primary hyperaldosteronism[57] also all lead to an increase in AF risk.

Any disease associated with systemic inflammation can cause damage to the atria of the heart through inflammatory injury to the microcirculation and microvascular dysfunction[34]. States of systemic or adipose tissue specific inflammation can lead to transformation of the adipose tissue within the epicardium, the epicardial adipose tissue (EAT), leading to an expansion of its volume and dysregulation of its physiology to become pro-inflammatory[58]. The EAT shares a microcirculation with the cardiac myocardium – there is no fascial barrier. This means that any dysregulation in EAT is readily transmitted via paracrine signalling to the myocardial tissues directly involved in electrical impulses of the atria. The secretion of pro-inflammatory adipocytokines such as tumour necrosis factor-α, interleukin 1-β and interleukin-6 from dysfunctional EAT has been demonstrated to cause microcirculatory injury and fibrosis of myocardial tissue[59-61]. Importantly, the chronic systemic inflammatory and metabolic disorders linked to AF risk are all accompanied by an increase in EAT volume[62-69] and the expansion in the EAT is proportional to the severity of the underlying systemic inflammatory disorder[34]. Another key observation is that in patients with known AF there is a direct association between the volume and inflammatory state of the EAT and the degree of electrical dysfunction in the adjacent myocardium[70,71]. The role of EAT also extends into incident AF in those without previous diagnosis. EAT volume has been shown to predict the incidence AF in those without any other know cardiovascular disease risk factors[72], and it is EAT alone (not other visceral fat depots) that are associated with such risk[73]. Finally, EAT volume has been shown to increase when patients move from paroxysmal to persistent AF[74]. Proinflammatory mediators released by EAT (outside-to-inside signalling) that influence the atrial myocardium are numerous. Leptin, aldosterone and natriuretic peptides have all been implicated in the disruption of myocardial cell activity and associated electrophysiological derangement. [75] Up-regulated activity of aldosterone and down-regulation of endogenous natriuretic peptides can promote EAT deposition and a shift of its biology from a protective or benign state to a proinflammatory state[76]. Further, local rises in aldosterone can cause microvascular injury and fibrosis[77]; these actions are normally opposed by those of endogenous natriuretic peptides[78], but their influence may be minimized by increased clearance of natriuretic peptides by hypertrophied and inflamed adipocytes[79]. Finally, EAT inflammation causes the release of leptin[80], which can itself cause microvascular endothelial dysfunction and promote both inflammation and fibrosis within the atria wall[81].

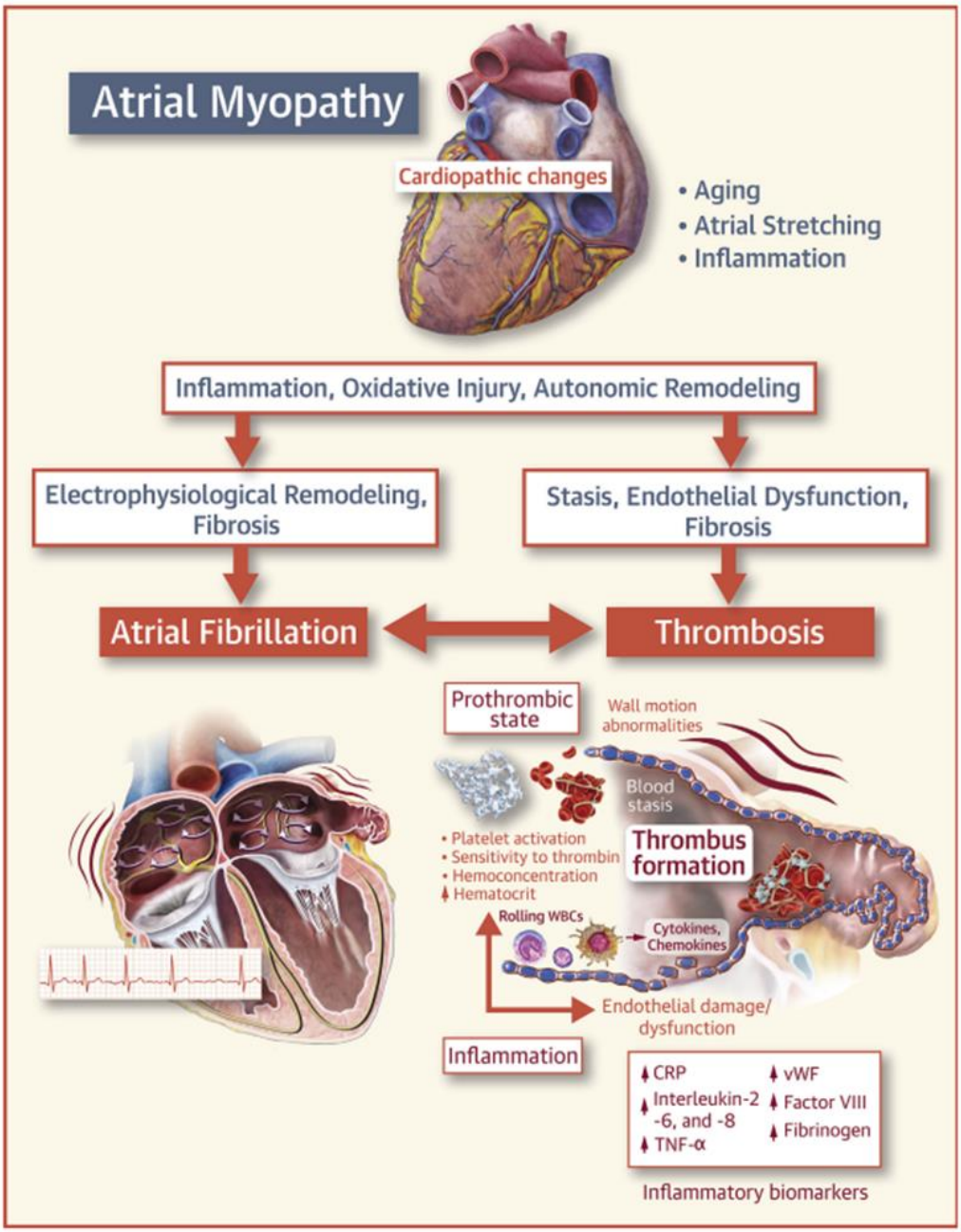


Figure 1.2. The pathophysiology of the atrial myopathy. *Atrial myopathy is typically caused by insults such as aging, inflammation, oxidative stress, and stretching of the atria. These myopathic changes alter the properties of myocardial electrophysiology and the cardiac autonomic nervous system. They can also lead to architectural structural changes characterized by fibrosis. Furthermore, atrial myopathy results in endothelial dysfunction and stasis, thereby a prothrombotic state. Electrophysiological remodelling and fibrosis facilitate the development of atrial fibrillation, which leads to more inflammation, fibrosis and autonomic remodelling, all of which contribute to a worsening prothrombotic environment, mediated by circulating inflammatory cytokines, chemokines and other molecules such as C-reactive protein (CRP), interleukin (IL)- 2, -6 and -8, tumour necrosis factor (TNF)-α, etc. Atrial fibrillation and thrombosis can develop separately and interact closely to further aggravate the underlying atrial myopathic processes. CRP: C-reactive protein; vWF: von Willebrand's Factor; WBC: white blood cell.* ***Reproduced with permission from Shen, Arora, Jalife. JACC Basic Transl Sci, 2019.***

1.3.2. Detecting the atrial myopathy

Currently within clinical practice there is no widespread approach to the detection and management of the atrial myopathy. The typical diagnosis of AF remains through the performance of an electrocardiogram (ECG) following a patient's report of symptoms (such as palpitations, fatigue, dyspnoea, or chest pain), or an incidental finding on an ECG performed for unrelated clinical purposes. The rise of wearable heart rate and rhythm monitors in smart watches has also contributed to the diagnosis of sub-clinical AF[82]. It has been proposed that the atrial myopathy will manifest as either AF itself, non-AF atrial arrhythmias, or ischaemic stroke, and that surrogate markers such as atrial dilation and impaired atrial systole may be useful in the clinical diagnosis of atrial myopathy[83]. Frequent atrial premature beats or paroxysmal atrial tachycardia have been implicated in abnormal atrial substrate and a predisposition to AF and stroke (regardless of AF status), however the detection of such symptoms in apparently healthy populations is very difficult[84,85] and requires targeted screening with expert ECG analysis. Recently there has been research focus on the use of artificial intelligence (AI) to analyse ECG for the detection of AF risk[86]. There are promising results with AI analysis of 12-lead ECG found to have a similar predictive usefulness as a clinical risk factor model (CHARGE-AF score; Cohorts for Heart and Aging Research in Genomic Epidemiology–Atrial Fibrillation[87]) for incident AF[88]. The AI trained ECG scores have not been found to be predictive of incident ischaemic stroke risk as yet, and the relationship between the AI risk assessment models and the atrial myopathy is not fully understood. Analysis of a recent AI powered ECG model for AF risk using saliency mapping and median waveform analysis showed that the model was critically influenced by the period of atrial depolarization and repolarization[88] (i.e., the P-wave and adjacent period of the ECG), a reflection of atrial structure and function that is likely to be strongly influenced by any underlying atrial myopathy.

It is through imaging technologies that hope lies for the most targeted and specific solutions to detecting atrial myopathy. Basic assessments of LA size and volume, through both echocardiography and cardiac CT[89], have been shown to be useful in detecting atrial disease and stroke risk, however these imaging examinations are not interrogating the atrial substrate itself, but the surrogate marker of atrial stretch and dilation. Echocardiography fails to produce the spatial resolution required for thorough interrogation of the LA wall and surrounding tissue, such as EAT. It is through targeted imaging of the atrial wall and surrounding tissues that it is hoped that characterisation of atrial myopathy, and its severity, will be possible. Cardiac magnetic resonance (CMR) has been used for this task, with delayed-enhancement CMR used to assess LA myocardial fibrosis. Such CMR detected fibrosis has been shown to be an independent risk factor for ischaemic stroke following adjustment for clinical risk factors[90]. CMR can also be used for 4-dimensional flow analysis for the development

of stasis maps in the LA and LAA. These maps can reveal physiological changes in LA haemodynamic in AF and reveal an increased risk to atrial thromboembolism that is not detected via the CHA_2DS_2-VASc score[91].

Despite the promising CMR data, the adoption of such technologies for atrial myopathy risk detection is limited by several challenges, often shared by other advanced imaging modalities. These include a requirement for a high level of financial investment for quality image acquisition and significant service and experienced clinician time.

1.4. The role of epicardial adipose tissue in cardiovascular disease

The research focus on the role of EAT in cardiovascular disease risk is not exclusively because of its implication in the atrial myopathy. EAT has been strongly implicated in both CAD and heart failure[92], and may play a role in broader metabolic and inflammatory disease, including COVID-19 cardiac syndrome[93]. EAT is unique due to its anatomical position with unobstructed proximity to the heart muscle, and due to its distinct transcriptome[94] which renders it different to other visceral and subcutaneous adipose tissue deposits.

EAT is largely composed of adipocytes, but also contains nerve cells, inflammatory cells such as macrophages and mast cells, stromal cells, vascular cells and immune cells. EAT is traditionally thought of as white adipose tissue but it also exhibits brown fat-like features[95]. No fascia is present between EAT and the myocardium meaning that the two tissues share the same microcirculation, facilitating crosstalk between EAT and the contiguous myocardium. The role of EAT in specific disease context for CAD and AF is demonstrated in Figure 1.3 and explored in detail in the subsequent two sections.

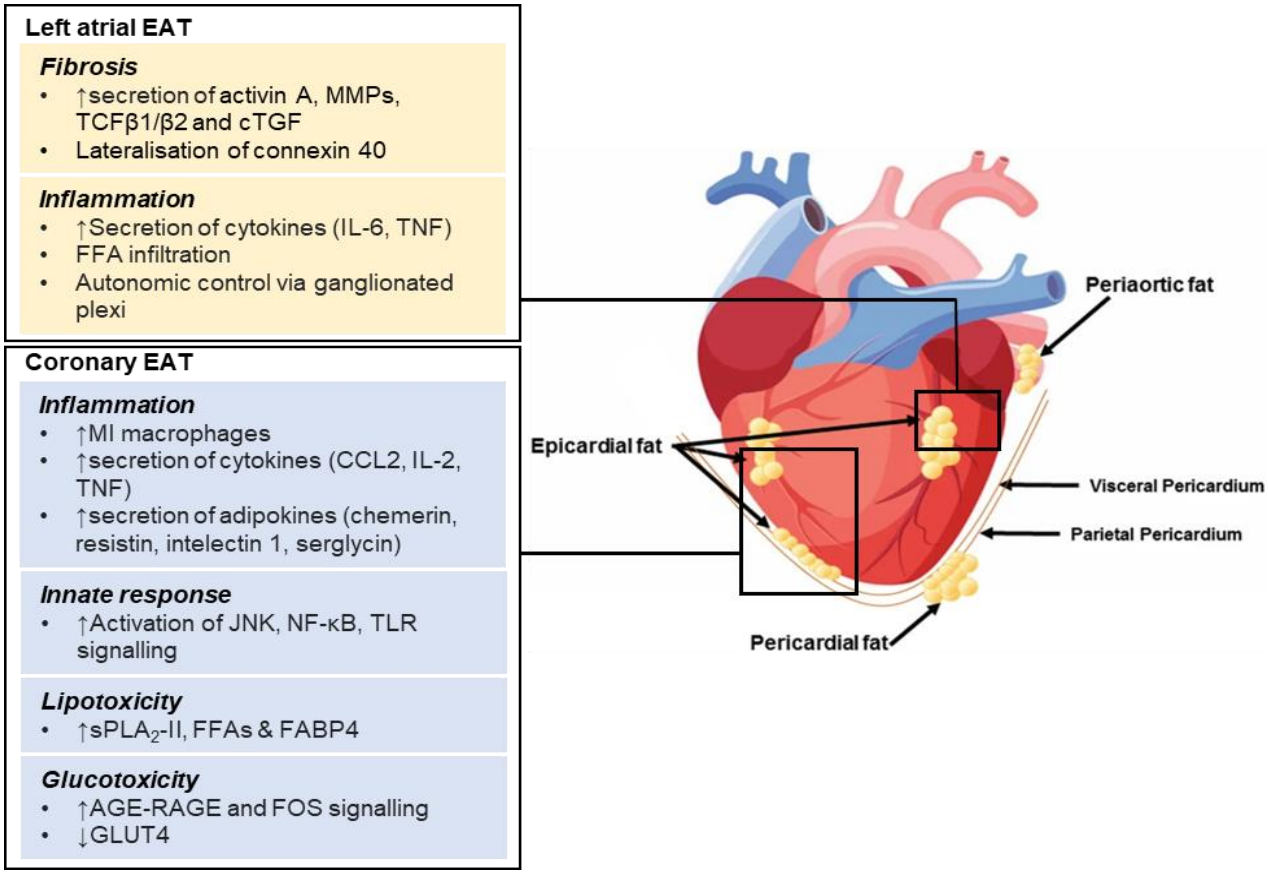


Figure 1.3. Role of EAT deposits in coronary artery disease and atrial fibrillation. The epicardial adipose tissue (EAT) exists as local depots between the myocardium and the visceral layer of the pericardium. EAT can therefore infiltrate the left atrium (left atrial EAT) and surround the coronary arteries (coronary EAT, or peri-coronary EAT). By contrast, pericardial adipose tissue is located between the visceral and parietal layers of the pericardium. Left atrial EAT has a high expression of genes encoding pro-arrhythmogenic factors. Left atrial EAT can contribute to atrial fibrillation through the local secretion of profibrotic factors (matrix metalloproteinases (MMPs), transforming growth factor-β1 (TGFβ1) and TGFβ2, connective tissue growth factor (cTGF) and activin A and inflammatory factors (IL-6 and tumour necrosis factor (TNF)) as well as free fatty acid (FFA) infiltration and increased autonomic control via ganglionated plexi. The coronary EAT has a high expression of genes encoding pro-inflammatory adipokines and factors regulating glucose and lipid metabolism. Coronary EAT can influence the development and progression of coronary artery disease through increased infiltration of pro-inflammatory M1 macrophages from EAT into the adjacent myocardium, the paracrine or vasocrine release of several pro-inflammatory cytokines (CCL2, IL-6 and TNF) and adipokines (chemerin, intelectin 1 (also known as omentin 1), resistin and serglycin), and the activation of innate immune response factors such as JUN N-terminal kinase (JNK), nuclear factor-κB (NF-κB) and Toll-like receptors (TLRs). Upregulation of signalling via advanced glycation end products (AGE) binding to their receptor RAGE in EAT can contribute to the oxidative stress and endothelial damage associated with coronary atherosclerosis in patients with diabetes mellitus. The excessive influx of FFAs from EAT into the coronary arteries is mediated by enzymes such as group II secretory phospholipase A2 (sPLA2-II) and adipocyte fatty acid-binding protein (also known as FABP4). GLUT4, glucose transporter type 4. **Adapted with permission from Berg et al, 2019, Epicardial Adipose Tissue in Cardiovascular Disease, from 'Bioactive Lipids in Health and Disease' part of the 'Advances in Experimental Medicine and Biology' book series (AEMB, volume 1127;131-143) and Iacobellis, 2022, Epicardial adipose tissue in contemporary cardiology, Nat. Rev. Cardiol. 19(9):593-606.**

1.4.1. Role of EAT in coronary artery disease

Inflammation is the major feature in the CAD risk conveyed by EAT. In patients with known CAD, EAT possesses thick infiltrates of pro-inflammatory MI macrophages, mast cells and T cells[96,97]. Microarray analysis has demonstrated that EAT has a transcriptomic profile that is pro-atherogenic in patients with CAD[94]. Genes strongly implicated in inflammatory pathways are upregulated in the EAT of patients with CAD and in patients who have suffered AMI. For example, pro-inflammatory genes that encode for cytokines such as IL-6, CCL2 and TNF have been implicated, along with chemokine ligands and receptors, and pro-inflammatory adipokines including chemerin[98] and resistin[99]. Importantly this upregulation of pro-inflammatory genes in CAD has been shown to be present in EAT but not in other visceral fat depots in the same patients, such as omental fat[100]. The effects of this rich pro-inflammatory proteasome within the tissues immediately abutting the coronary artery adventitia, and intersperses with the adventitia, is the primary source of inflammatory signalling via paracrine means. There is a disequilibrium between pro-inflammatory and anti-inflammatory adipokines secreted by EAT in patients with CAD, exhibited most clearly through the gene and protein expression of adiponectin, a cytokine with anti-inflammatory properties, which has been found to be supressed in those with CAD[94]. The role of adiponectin as an antioxidant on cardiomyocytes through the downregulation of myocardial NADPH oxidase activity has been described, offering insight into the protective mechanisms that EAT can play in myocardial and vascular oxidative stress[101,102].

Patients with CAD are more likely to have an upregulation in inflammatory cytokine expression through the activation of mediators of the EAT innate immune response. These include Toll-like receptors, JUN N-terminal kinase and nuclear factor-κB[100]. The relationship between endothelial function and EAT is well researched. The Antoniades laboratory has proven that there is significant and important paracrine cross-talk between EAT and the vascular wall[103]; in the presence of coronary inflammation and atherosclerotic disease, the release of pro-inflammatory mediators from the vascular wall into the surrounding EAT results in significantly impaired adipocyte differentiation and intracellular lipid formation, with important implications in cardiovascular diagnostics[104]. This crosstalk has been demonstrated to be influenced by a number of molecules including endothelium-derived microparticles containing miR-19b, which have been shown to induce EAT inflammation and promote atherogenesis in Apoe$^{-/-}$ mice[105]. The role of EAT in endothelial disease is clearly of fundamental interest for CAD risk, and those with known CAD have been shown to have EAT that is secreting greater levels of molecules that can have deleterious effects on endothelial cells, such as resistin, which causes endothelial cell permeability. Other mechanisms that EAT can influence CAD

and AMI include through the activation of coagulation implicated mediators such as tissue plasminogen activator[94].

The EAT is also implicated in ectopic lipids. EAT is a reservoir of fatty acids that if releases have been shown to influence lipid accumulation within the coronary artery tree. It has been shown that group II secretory phospholipase A2, the enzyme necessary for the synthesis of pro-inflammatory lipids, is upregulated in those with symptomatic CAD compared to healthy individuals[106]. The potential mechanisms that implicate EAT in vascular lipid rich lesions are myriad and include the expression of FABP4[107], the high degree of conjugated fatty acids in EAT compared to other visceral fat deposits[108], and the elevated impact of insulin-stimulated lipogenesis in EAT compared to other fat deposits. The upregulation of advanced glycation end products and their receptors (AGE-RAGE) can also drive EAT towards a state of oxidative stress and subsequent endothelial damage[109].

1.4.2. Role of EAT in atrial fibrillation

EAT influences the risk of atrial fibrillation in multiple ways[110], as has already been explored above in the mechanisms of the atrial myopathy. Key mechanisms of how EAT influences AF risk include genetic and neural factors, the direct impact of inflammation and subsequent fibrosis on cardiomyocytes, and the effect of fatty infiltration on atrial remodelling which in turn impact on atrial conduction (Figure 1.3). The embryonic origins of EAT are important in its possible roles in AF and atrial myopathy. The embryonic epicardium progenitor cells can differentiate into cardiac fibroblasts, smooth muscle cells or adipocytes[111]. It has been shown that the atrial natriuretic factor secreted by atrial myocytes when stressed has adipogenic potential that leads to greater differentiation of epicardial progenitors into adipocytes[112]. Patients with AF have been found to have a stronger adipogenic potential within their atrial myocardial secretome than matched healthy controls[112]. Given this, it is postulated that under exposure to stressors such as hypertension or diabetes mellitus the atrial myocytes could contribute to the expansion of EAT and therefore to the development of the atrial myopathy substrate itself[92].

The EAT located around the LA and infiltrating the atrial myocardium has been found to exhibit increased gene expression for proteins implicated in oxidative stress, muscle contraction and calcium signalling. These findings are not reported in EAT extracted from peri-coronary or periventricular EAT[113], indicating the unique role that EAT adjacent to the atria may play in AF risk.

The absence of any fascia between the EAT and left atrial myocardium, and the shared blood supply via the coronary system provides an environment for both inwards and outwards communication between EAT and the atrial myocytes and electrical conduction cells. There is excellent evidence that profibrotic factors including matrix metalloproteinases (MMPs) and activin A diffuse locally from EAT

amongst cardiac myocytes, and that pro-fibrotic cytokines such as the interleukins and TNF also do the same[113]. In vitro studies have shown that conditioned EAT medium upregulated transforming growth factor-β1 (TGFβ1) and TGFβ2 and led to fibrosis of EAT tissues, which is mediated by EAT secretion of activin A.

The role of FFAs in AF risk is a fascinating area of research. FFAs can diffuse locally into the atrial myocardium and contribute to electromechanical changes in myocardial tissue through the physical disruption and separation of cardiomyocytes[114]. This results in conduction slowing, the loss of side-to-side cell connections, and cardiomyocyte disorganisation that may result in conduction delay and re-entry conduction as is reported in clinical AF[115,116].

The role of the nervous system in AF risk is influenced by EAT. Both sympathetic and parasympathetic nerve fibres are found in EAT and contribute to physiological autonomic control of the heart. The ganglionated plexi that are responsible for the initiation and maintenance of AF is typically located surrounded by EAT[117]. The activation of these plexi can lead to an increase in the calcium transient amplitude in the atrial myocardium[118]. That botulinum injection into EAT during cardiac surgery can suppress ganglionated plexi, reduce autonomic nervous activity and have longer-term beneficial effects on atrial fibrillation is likely related to this mechanism[119]. As with the risk of CAD due to EAT, some mechanisms have bene found that are protective of AG risk. There is excellent evidence that in a similar fashion to coronary vessels, myocardial derived oxidation products can lead to PPARγ-mediated upregulation of adiponectin biosynthesis in the adjacent EAT, which then exerts antioxidant effects on cardiomyocytes[101].

1.5. EAT as a biosensor of vascular and cardiovascular dysfunction

1.5.1. Vascular wall to EAT signalling

The inside to outside signalling from the vascular wall to proximal EAT via paracrine means has been well described[120]. Adipose tissue deposits such as EAT contain adipocytes with highly variable sizes and differentiation status[121]. This feature of adipose tissue lends itself to the investigation of signalling effects upon adipocyte structure and has revealed significant findings related to the role of 'inside to outside' signalling influencing EAT structure and function. It has previously been shown that PPAR-γ and CEBPA are synergistically responsible for the triggering and early control of differentiation within preadipocytes and their subsequent development into mature adipocytes[121]. Maturation of adipocytes is characterised by the expression of adipocyte specific genes and the subsequent upregulation of lipogenic enzymes. This is associated with increased accumulation of intracellular lipid which determines the volume of each adipocyte. The Antoniades laboratory has demonstrated[122] the differences in adipocyte volume and the differing tissue biology between

various adipose tissue depots. We have found that adipocyte size is on average smaller in human EAT and thoracic adipose tissue compared to subcutaneous adipose tissue. This is partly due to the presence of larger populations of immature preadipocytes identified by the lower expression of PPAR-γ, CEBPA, and FABP4.

Our laboratory group has also demonstrated the consistent bi-directional communication between the EAT and the vascular wall. We previously postulated that vascular inflammation may release inflammatory mediators that induce PPAR-γ–mediated structural changes in proximal EAT depots, the peri-vascular adipose tissue. We have shown that when the vascular wall expresses proinflammatory cytokines the inflamed vascular tissue prevents lipid droplet formation in EAT adipocytes in an ex vivo coculture system by interfering with the ability of preadipocytes to differentiate, as well as by triggering their proliferation[104]. The effect within the EAT is significant: there is marked increase in lipolysis, a decline in adipogenesis and an increase in edema throughout the EAT. Importantly, we have demonstrated a clear gradient in adipocyte size and the expression of adipogenic genes including FABP4, CEBPA, and PPAR-γ, when moving from EAT adjacent to the human coronary arteries to adipose tissue further away. This is demonstrated in Figure 1.4A-C which shows the gene expression level gradient between proximal and more remote adipose tissue, and Figure 1.4D for the gradient in adipocyte size, both utilising EAT from around the coronary arteries. These findings led to the successful development of the CT powered tool for the non-invasive assessment of coronary artery inflammation – the Fat Attenuation Index (FAI). This new imaging tool enabled the analysis of coronary artery FAI in cylindrical layers around the human coronary arteries, demonstrating a parallel shift of FAI to more negative values moving away from the vascular wall, reflective of the increase in adipocyte size and the adipose tissue lipid content (see Figure 1.4E-I). The pathobiological process of inside to outside signalling leading to a shift in EAT structure and function, and the subsequent detection of this phenomenon via CT, was demonstrated in a video developed and published[123] from the Antoniades group which is available online as supporting information to the original article.

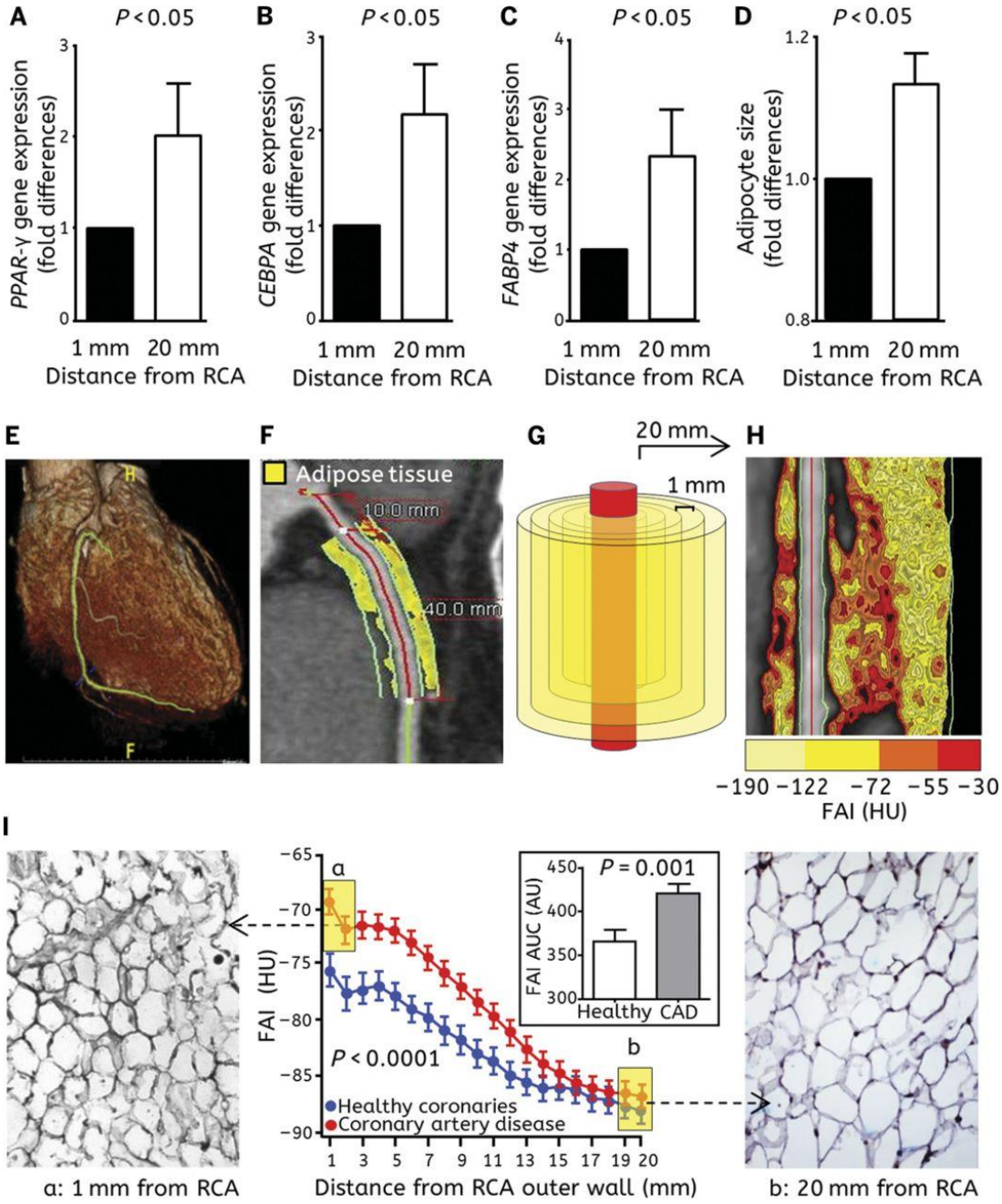


Figure 1.4. Gradient of adipocyte size and FAI around the human coronaries in the presence or absence of coronary atherosclerosis. *(A)PPAR-γ, (B)CEBPA, and (C) CEBPA4 gene expression and (D) adipocyte size in pericoronary adipose tissue samples attached to the RCA and paired samples ~20 mm away from it [n= 6 to 12 pairs for (A) to (D)]. (E to G) In study arm 3, FAI around the RCA of patients undergoing CTA was calculated for each cylindrical 1-mm-thick layer of pericoronary tissue, whereas for a radial distance, FAI was calculated from RCA wall 1 to 20 mm. (H) FAI mapping of PVAT around the RCA. (I) FAI and radial distance from vascular wall in patients with CAD (n = 149) versus healthy individuals (n= 117) [comparison of the area under the curve (AUC) using unpaired t test and comparison of the curves using two-way ANOVA for repeated measures with "FAI × distance" interaction].P values by Wilcoxon signed-rank test (A to D).*

Reproduced with permission from Antonopoulos A et al. Sci Trans Med, 2017. 9(398)

1.5.2. Atrial wall to EAT signalling

The inside to outside signalling discussed so far focuses on the signalling from the coronary vascular wall to the proximal EAT surrounding the coronary tree. We have also undertaken work related to the role of inside to outside signalling from the myocardial wall of the heart influencing the EAT proximal to the left atrium through similar mechanisms, as summarised in Figure 1.5.

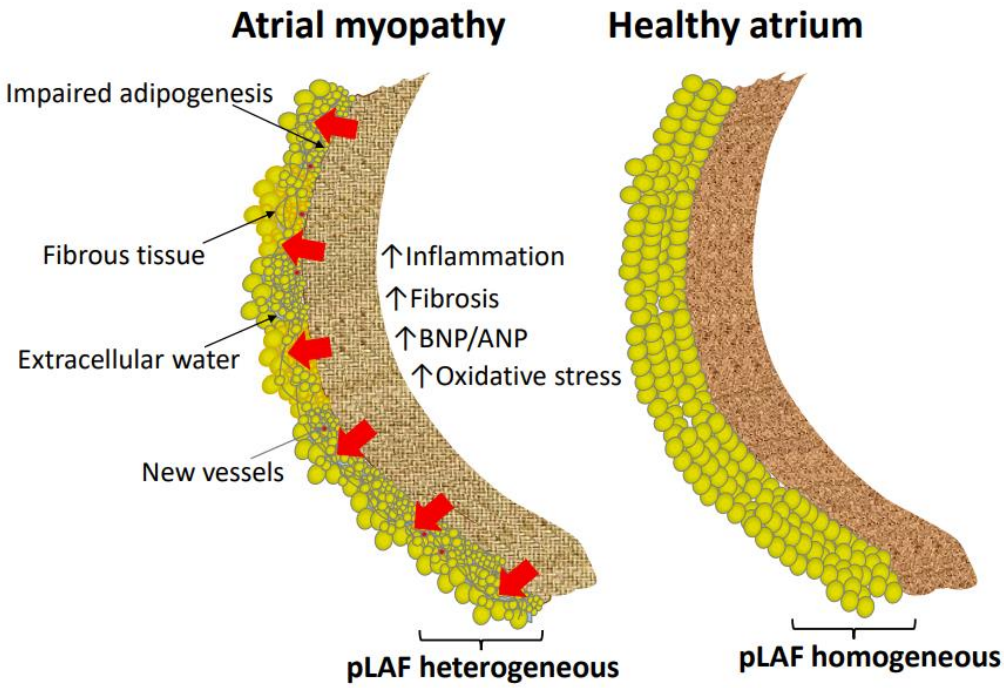


Figure 1.5. Illustration of the major components of the inside to outside interplay between the atrial wall and EAT. *Similar to the relationship between the vascular wall and proximal EAT, the atrial wall also exhibits inside to outside signalling with strong influence of the atrial endothelium and myocardium on the adjacent EAT. This is mediated through fibrotic, inflammatory changes within the atrial myocardium that drives tissue changes within the EAT such as impaired angiogenesis, fibrotic change, and neoangiogenesis. ANP: atrial natriuretic peptide; BNP: brain natriuretic peptide; pLAF: peri left atrium fat.*

The inside to outside signalling effect of the atrial wall on EAT has been demonstrated in experiments performed within the Antoniades lab that discovered the influence of atrial myocardial oxidative stress on EAT[101]. We demonstrated a positive association between circulating adiponectin and the expression of the ADIPOQ gene in EAT with NADPH oxidase activity as a measure of atrial myocardial redox state. This finding was independent of plasma BNP levels. Beyond this association it was also discovered that 4HNE, a product of oxidation in the human heart that can cross cellular membranes as a signalling molecule, leads to upregulation of ADIPOQ gene expression in human EAT via a PPAR-γ–dependent mechanism. This finding was then confirmed using a cardiac myocyte–

specific NOX2 transgenic mouse and a pig model of rapid atrial pacing (as means to increase NADPH oxidases activity and O_2^- generation in the atrial myocardium).

The effects of the inside to outside signalling have been further elucidated through work investigating the EAT structural and functional changes that can occur in states of atrial endocardial and myocardial dysfunction which can be detected through CT imaging in a similar manner to the peri-coronary FAI[124]. The tissue changes in the peri-atrial space, including in the EAT, have been successfully identified on CT imaging[124] using radiomic analysis. When some or all of the hallmarks of the atrial myopathy, including fibrosis, angiogenesis, oedema, lipolysis, and decreased adipogenesis, are present the density and texture of the EAT and the adjacent connective tissues surrounding the LA changes in ways that are identifiable on CT. It is the detection of these processes that allows the use of radiomics for the assessment the clinical events that can result from atrial myopathy such as AF and ischaemic stroke. In pilot data analysis of an early version of the Atriomic Risk Algorithm, a radiomic risk signature for the prediction of ischaemic stroke, we have found that the upregulated gene expression of atrial COL1A1, a gene associated with fibrosis, and atrial TNFa, a gene associated with inflammation, was strongly associate with a high risk for stroke as assessed via the pilot Atriomic Risk Algorithm (Figure 1.6).

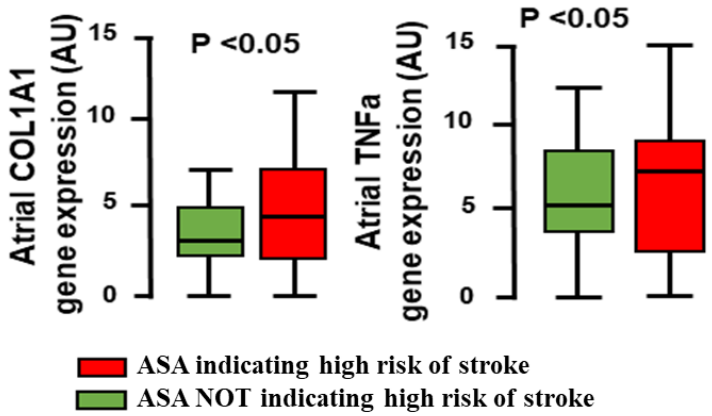


Figure 1.6. Association of atrial genes with pilot Atriomic Risk Algorithm for ischaemic stroke risk. *When tested in pilot analysis, the initial pilot version of the ASA risk prediction tool for ischaemic stroke was found to correlate high risk with genes known to be associated with atrial fibrosis and atrial inflammation. ASA: Atriomic stroke algorithm.*

1.6. Imaging the left atrium and surrounding tissues using computed tomography

Cardiac CT has traditionally fulfilled two important clinical roles; the assessment and quantification of coronary artery calcium load via non-contrast CT assessment, and the exclusion of anatomically

significant CAD via contrast enhanced coronary CT angiography (CCTA). As opposed to invasive coronary angiography or functional testing that detects luminal narrowing and myocardial ischaemia, respectively, CCTA provides a platform to assess the coronary lumen, plaque if present[125], and the structure of the cardiac chambers and myocardium and the surrounding extra-cardiac tissues and organs[126]. The use of CT in this thesis is to enable assessment of the LA, the LA wall and the surrounding tissues. This use of CT is unusual in that it is focused on structures not typically assessed, either qualitatively or quantitatively, in clinical interpretation of non-contrast cardiac CT or CCTA.

1.6.1. Quantitative assessment of epicardial adipose tissue

Numerous imaging modalities enable the characterisation of adipose tissue within the human body. Each modality has its own strength and weaknesses which has seen different modalities gain acceptance in different clinical settings, as outlined in Figure 1.7. For example, dual energy X-ray absorptiometry (DEXA) is commonly used for whole body fat mass as it is easy to use and is widely available. Despite this, DEXA does not allow segmentation and quantification of specific adipose tissue depots (e.g., EAT), so its use is limited to whole body assessment only. Ultrasound assessment of adipose tissue is commonly carried out, including of EAT. Ultrasound allows two-dimensional thickness assessments of adipose tissue depots in an easy to use, yet operator-dependent manner[127]. The major limitation of ultrasound assessment of adipose tissue depots is the poor association between three-dimensional volumes. The use of magnetic resonance for assessment of adipose tissue overcomes these issues, but is limited by the low levels of availability and prolonged acquisition times required for accurate assessment[128]. When MR has been applied on ex vivo human myocardial samples, saturation-recovery T1 mapping and DIXON images were found to be excellent at identifying overall fat, interstitial fat and fatty fibrosis as confirmed by histology examination of the same samples[129], however such results have not been able to be replicated in vivo yet. MR analysis of adipose tissue has advantages in hybrid structural and functional settings. An excellent example of this is LA strain assessed through feature tracking. Peak longitudinal atrial strain has been found to correlate with the degree of fibrofatty replacement within the left atrial myocardial wall on histological analysis[130].

CT imaging has become the most widespread clinical imaging modality for the characterisation of any structure within the human body and has become the gold-standard for adipose tissue quantification. This is owing to its submillimetre spatial resolution and ease of access and use across the world. CT offers 3D views of the heart and its epicardial surface. Based upon the reconstructed

cross-sectional images, volumetric measurements can be performed with a high degree of accuracy[128]. The assessment of adipose tissue on CT relies upon pre-defined and validated attenuation thresholds (typically defined as -190 to -30 Hounsfield Units [HU])[120]. The quantitative nature of CT (i.e. each voxel is characterised by an attenuation value – a Hounsfield Unit that describes its relative radiodensity compared to water) enables an indirect assessment of relative tissue density, therefore adding a qualitative dimension to the volumetric characterisation of adipose tissue depots[131]. CT scanning protocols rely on the administration of intra-venous contrast for the opacification of the vascular lumen for CCTA. This allows the simultaneous assessment of coronary artery anatomy and the segmentation of tissue depots such as EAT or perivascular adipose tissue.

The structural information gleaned by CT can be further supplemented by 18F-FDG-PET imaging. PET-CT can reflect the metabolic and inflammatory activity of adipose tissue[132]. However, PET-CT imaging has the same limitations such as radiation exposure, clinical availability, and cost.

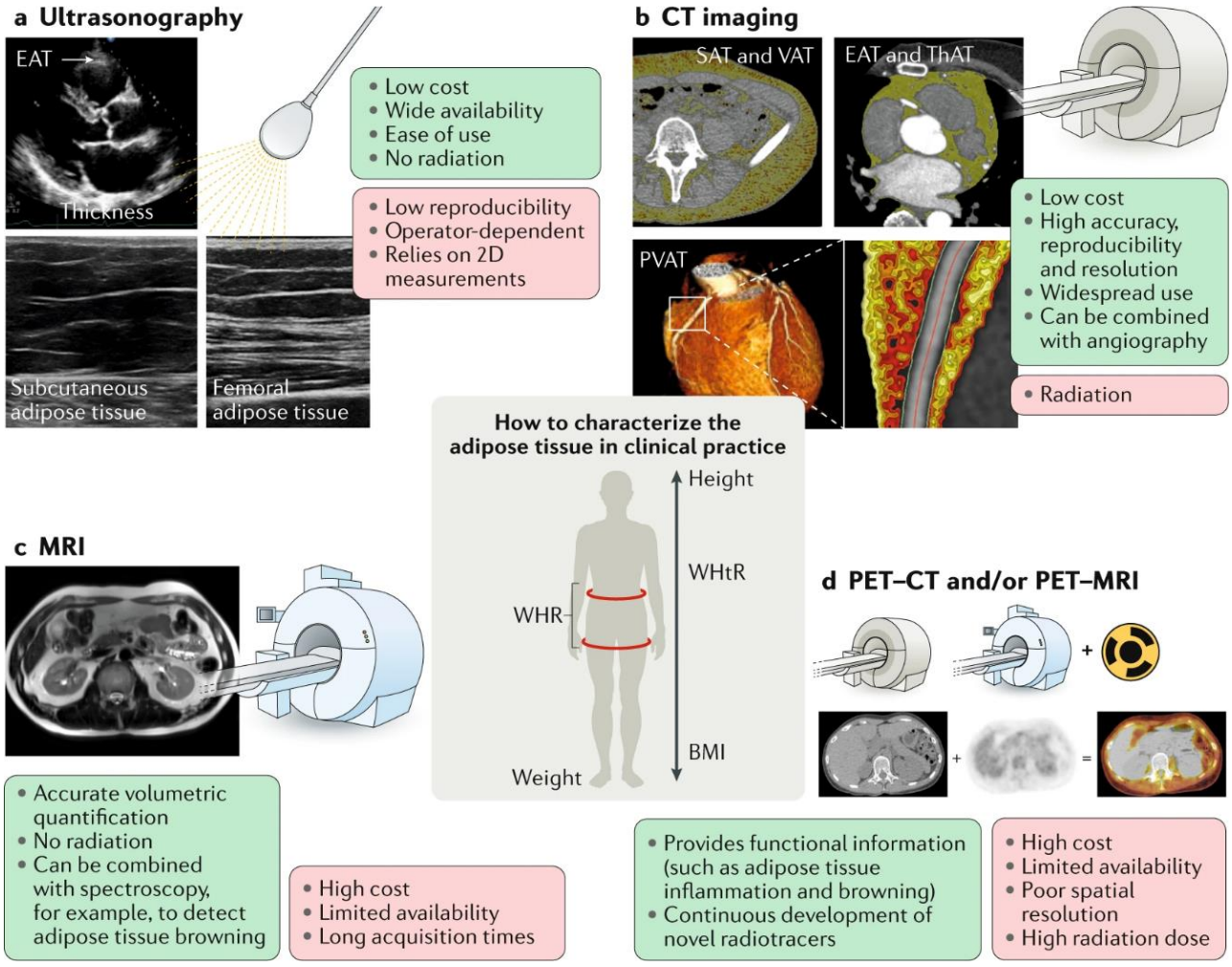


Figure 1.7. Advantages and disadvantages of imaging modalities to characterise adipose tissue. *Commonly used anthropometric indices such as the body mass index (BMI) do not capture critical*

adipose tissue features, such as distribution, biology and overall body fat composition. Alternative anthropometric measurements, such as waist-to-hip ratio (WHR) and waist-to-height ratio (WHtR), are more appropriate obesity indices of cardiometabolic risk than BMI but still do not provide important information on adipose tissue biology and visceral distribution. Several widely used clinical imaging modalities can be employed to provide this information, with each modality having its relative advantages (green boxes) and limitations (red boxes) compared with the others. 2D: two-dimensional; CT: computed tomography; EAT: epicardial adipose tissue; ThAT: thoracic adipose tissue; MRI: magnetic resonance imaging; PET: positron-emission tomography; PVAT: perivascular adipose tissue. SAT: subcutaneous adipose tissue; VAT: visceral adipose tissue. ***Reproduced with permission from Oikonomou EK & Antoniades C. Nat Rev Cardiol, 2019. 16(2):83-99***

1.6.2. Quantitative assessment of the left atrium and surrounding tissues

The assessment of the LA and peri-LA tissues on CT has not been a major focus of research. Most work to date relates to the quantification of LA volume alone including the establishment of normal ranges for LA volume, and the creation of specialised indices of LA size[133].

There has been previous work related to the semi-automated[134,135] and automated assessment of EAT volume on CT. The focus has been on whole heart EAT (not limited to peri-LA EAT) and automated algorithms via deep-learning approaches have been created for this purpose[136,137]. The majority of segmentation models published to date lack robust training cohorts (i.e. trained in convenience samples or selected samples of high quality scans) and/or lack true external validation in clinically relevant population samples. No automated segmentation algorithms have been applied to real-world clinical cohorts of all-comers for clinically indicated CCTA. Currently, there are no examples of automated EAT segmentation algorithms being incorporated into clinical workflow for patient risk assessment purposes. Moreover, the tissues surrounding the left atrium chamber such as the LA wall and adjacent tissues within the visceral pericardium (EAT, extra-cellular matrix and fibrous connective tissues) have not been reported as being investigated specifically on CT, either manually or in an automated manner.

1.7. The radiomic approach to computed tomography interpretation

The use of radiomics in CT has rapidly expanded in recent years. CT scans are more than plain images, they are data. Radiomics is the field of imaging processing that extracts quantitative information from imaging data to generate datasets of data mining through traditional statistical approaches and novel AI approaches[138]. Radiomic features range from simple volumetric, shape-related or first order statistics, to second and higher-order statistics that describe the texture of a segmented volume and the spatial relationship of voxels with similar or different attenuation values[139]. Such features can identify imaging patterns of significant clinical value that may not be at all recognisable by the naked eye[140].

1.7.1. Radiomic texture analysis

Following the segmentation of an independent volume(s) of interest on a CT scan, the radiomic features can be extracted for analysis purposes. These segmentations can be 2D or 3D in shape. 2D segmentations are typically upon a single slice of a CT scan, while 3D span multiple slices. Within this thesis radiomic features were extracted by our own technique using a custom tool – the RadExtractor, which is discussed in Chapter 2. To undertake radiomic analysis the attenuation values present in the volume of interest are discretised into bins of a given size/HU range. This is the mathematical transformation stage of radiomic analysis and forms a critical aspect of the methodology to successfully extract the features for further analysis as is undertaken within this thesis.

1.7.2. Radiomic features

For any unique CT segmentation, a total of 872 radiomic features can be calculated (Table 1.1 and Appendix 1). Radiomic features are defined through consensus, and have been formalised through the Pyradiomics library[141] which allows the features to be incorporated into image segmentation software such as 3DSlicer[142]. Shape-related and first-order radiomic features can be calculated using the raw HU values of the segmented region. For calculation of texture features (Gray Level Co-occurrence Matrix [GLCM], Gray Level Dependence Matrix [GLDM], Gray Level Run-Length Matrix [GLRLM], Gray Level Size Zone Matrix [GLSZM], and Neighbouring Gray Tone Difference Matrix [NGTDM]), adipose tissue voxels can be discretised into 16 bins of equal width (width of ten HU), to reduce noise while allowing a sufficient resolution to detect biologically significant spatial changes in segmentation attenuation. Discretisation into 16 bins has been previously recommended as the optimal approach to increase the signal-to-noise ratio of CT images for radiomic analysis[139]. To enforce symmetrical, rotationally-invariant results, texture statistics (GLCM etc.) can be calculated in all directions and then averaged, as previously described[139]. First order and texture-based statistics can also be calculated with three-dimensional wavelet transformations of the original image resulting in eight additional sets of radiomic features[143].

Table 1.1. Radiomic features available for extraction from computed tomography images

	Original	**Wavelet transformations (n=8)**	**All**
First order	19	152	171
Shape related (both 2D & 3D)	26	-	26
GLCM	24	192	216
GLDM	14	112	126
GLRLM	16	128	144
GLSZM	16	128	144
NGTDM	5	40	45
Total	**120**	**753**	**872**

1.7.3. The big data approach to stroke risk with radiomic and clinical data

Utilising radiomics features extracted from clinical CCTA scans it is possible to create radiomic risk signatures for disease outcomes such as AF and stroke. This approach to radiomics has been utilised previously within the Antoniades laboratory, with successful development of radiomics based risk scores for CAD[144] and for vascular inflammation in COVID-19[145].

Both risk scores relied upon a similar methodology in their development. The extraction of radiomic features from a large and relevant population sample is necessary for the training and internal testing of the risk signatures which can be developed via machine learning. It is possible to train the radiomic signature to detect relevant biology, such as gene expression for certain inflammation or fibrosis related pathways. This is achieved through the collection of myocardial tissue and/or EAT from patients undergoing cardiac surgery, along with the acquisition of CCTA for the same patients. A radiomic signature trained within the cohort with biological ground truth available can then be tested and fine-tuned within a broader cohort of clinical CCTA patients and used as a non-invasive means to detect the radiomic changes associated with the harmful processes that the algorithm was originally trained to detect.

To train such algorithms, large cohorts of clinical CCTA patients are required, with detailed clinical risk factors from prior to the time of the CCTA, and detailed outcomes following the scan. There are no clinical CCTA cohorts of the required size for such workflows reported within the literature to date. Others have previously created clinical CCTA registries, such as the CONFIRM (COronary CT Angiography EvaluatioN For Clinical Outcomes: An InteRnational Multicenter) Registry[146] and large cohorts for trials such as PROspective Multicenter Imaging Study for Evaluation of chest pain (PROMISE) trial and the Scottish computed tomography of the heart (SCOT-HEART) Trial[147,148], but none of these have been on the required scale for the development of radiomic risk signatures for conditions such as AF and stroke – conditions which although common, require very large numbers and prolonged periods of follow-up for enough outcomes to accrue for statistical power to be acceptable. Such cohorts must be powered with enough patients who have sufficient follow-up following the date of the clinical CCTA so that algorithms can be trained by splitting the cohort by the relevant disease outcomes (such as a confirmed ischaemic stroke). Such cohorts require outcomes collected from national registries or nation-wide electronic healthcare record systems as opposed to traditional follow-up of prospective cohorts through patient clinics and telephone consults.

A major issue restricting the creation of such large-scale clinical cohorts is ethical considerations. How tens of thousands to hundreds of thousands of clinical patients can be consented for such

cohorts is challenging from both a logistical and ethical view. Such research can only be carried out in a medical system with highly connected data flows and centrally stored registries of clinical events across a large number of hospital sites. The ethical considerations for such studies are also challenging, as no study could consent enough patients directly for inclusion in such large scale cohorts. High level ethical approvals (i.e. governmental) are needed to facilitate acceptable and safe access to clinical data at this scale. Such studies can only be performed in jurisdictions with mechanisms for the administration and oversight of such research.

1.8. Hypotheses and aims

The principal hypothesis of my doctoral project is that CCTA-based assessment of the EAT and of the tissues surrounding the LA of the heart is a promising means to extract information about the disease state of the heart. Therefore, it could be used to improve early cardiovascular disease diagnosis and risk stratification in patients undergoing CCTA as a part of clinical care.

To investigate these hypotheses, this thesis aims to:

a) Develop a major clinical CCTA cohort study for the development and testing of novel CT technologies such as radiomics based stroke risk assessment (Chapter 3 & 6);
b) Automate the quantification of whole heart EAT volume on clinical CCTA (Chapter 4);
c) Explore the ability of automated EAT volume quantification to detect cardiovascular disease, and predict risk of future cardiovascular disease (Chapter 5);
d) Automate the segmentation of the LA and peri-LA tissues, including EAT, on clinical CCTA (Chapter 4); and,
e) Explore the role of radiotranscriptomics for risk stratification for ischaemic stroke utilising LA and peri-LA segmentations from CCTA (Chapter 7).

2

Chapter 2. General methods

2.1. Introduction

This chapter examines the study design, patient populations and general methods employed to explore the diagnostic and prognostic value of the automated CCTA technologies developed within this thesis.

First, I present an overview of the individual studies that were used in the thesis. This is complemented by the following Chapter 3 which is dedicated to the creation and delivery of the overarching study, The Oxford Risk Factors And Non Invasive Imaging (ORFAN) Arm 4 Study.

Second, I discuss the CCTA protocols that were used in the studies and the retrieval and post-processing of the images. This section is complemented by Chapter 4, which deals with the specific development and testing of the novel CCTA automated analysis techniques used in this thesis. This section includes a description of the workflow in the Oxford Academic Cardiovascular Computed Tomography (OXACCT) core laboratory and the standard operating procedures that have been designed to govern this process.

2.2. Study populations included in this thesis

The ORFAN Arm 4 Study population was assembled to address the scientific aims of this thesis, and other scientific questions. The origins and development of the ORFAN Arm 4 study is explained in detail in Chapter 3. Here I include the clinically relevant details concerning the populations themselves and the clinical CCTA scans of these populations.

All clinical patient populations, bar one, are under the umbrella of the ORFAN Arm 4 Study. The exception is the SCOT-HEART Trial population, which is utilised as an external test cohort in Chapter 5.

Different groupings of patients from the ORFAN Arm 4 study are used for different developmental, validation and clinical association studies. This includes:

1) the manual training, internal and external validation of the automated deep-learning networks for whole heart segmentation, left-atrium (LA), and peri-LA segmentation;

2) the clinical associations of automated EAT volume and cardiovascular disease; and,

3) the development and validation of the pilot Atriomic Risk Algorithm for ischaemic stroke.

2.2.1. The development of automated segmentation and radiomic feature extraction: The ORFAN Arm 4 Study

The ORFAN Arm 4 Study is a multi-centre, international, prospective observational cohort study of clinical patients who undergo CCTA, with retrospective inclusion of eligible patients (CCTA for any clinical indication from 01/01/2010 to date) and collection of linked clinical data (01/01/2005 to date). This study was created for the purpose of this project and other similar projects. The planned follow-up of this cohort for clinical outcomes is until 2030 (maximum data collection of 25 years; 2005-2030). The study is described in Chapter 3. In this Chapter I outline the patient groups used for the purposes of the development of the automated CCTA segmentation networks.

A total of 3100 UK based patients in ORFAN Arm 4 were utilised for the developmental aim of this thesis – namely to:

1) Automate the quantification of EAT volume on clinical CCTA; and,

2) Automate the quantification of LA and peri-LA tissue segmentation on clinical CCTA.

2200 CCTA were used for the manual training of both deep-learning networks (DLN) for the quantification of the whole heart within the pericardium for the quantification of the EAT volume within that segmentation, and the LA and peri-LA tissues. This included 100 sequestered scans for preliminary internal validation following initial network training for the whole heart segmentation model, three groups of 200 scans unseen by the algorithm were used for feedback learning to improve the performance of the DLN (600 patients in total), and 200 unseen UK ORFAN Arm 4 scans were used for internal validation of the models.

UK based ORFAN Arm 4 patients utilised for the DLN development were sourced from the Oxford University Hospital NHS Foundation Trust, the Royal United Hospitals Bath NHS Foundation Trust, University Hospitals of Leicester NHS Trust, and the Milton Keynes University Hospital NHS Foundation Trust.

International patients from the Cleveland Clinic, Ohio, USA, were utilised for the external validation cohort. Utilising an entirely geographically and clinician/scientist distinct cohort allowed this validation to be a broad external validation cohort, as opposed to employing UK based or a UK/USA split cohort as a narrow external validation group[149]. A total of 720 patients were included in the external validation cohort. The Cleveland Clinic cohort consisted of a total of 2246 CCTA. Of the 2246 CCTA scans, 720 was randomly selected from the group for external validation of the DLN. The

number of 720 was selected based upon researcher availability for manual segmentation – the entire cohort was not manually segmented.

To be sure of the validity and applicability of the developed DLN for the automated segmentation of all volumes (whole heart, the LA and the peri-LA volumes) we tested the network in challenging yet commonly encountered CCTA patient groups. Two disparate cohorts were employed for this purpose, the AdipoRedOx Study, which included patients with CCTA within six weeks of major cardiac surgery, and the SCOT-HEART Trial, which included a variety of patients with relevant characteristic that rendered them challenging CCTA candidates.

AdipoRedOx Study

The AdipoRedOx Study is a University of Oxford sponsored study which includes patients undergoing coronary artery bypass grafting (CABG) surgery at the John Radcliffe Hospital (Oxford Research Ethics Committee C reference 11/SC/0140) who agreed to undergo CCTA soon after their surgery. The AdipoRedOx Study is primarily interested in the collection and analysis of tissue samples collected from the time of surgery, however for the purposes of this aim no tissue samples were utilised. The CCTA scans were employed as challenging validation patients with CCTA within six weeks of the open chest cardiac surgery. The DLN for segmentation of the whole heart, the LA and the peri-LA regions of interest were tested upon these patients. These scans were collected locally with a total of 253 available CCTA scans transferred to the OXACCT.

SCOT-HEART Trial

In the same manner as the AdipoRedOx Study patients, to ensure validity of the DLN in a wide diversity of patients with challenging CT characteristics our team collaborated with the SCOT-HEART Trial investigators who granted me access to the CCTA arm of the SCOT-HEART study (ethics NCT01149590)[148,150]. Briefly, SCOT-HEART was a prospective open-label, parallel group, multicentre trial, which recruited a total of 4146 patients aged 18–75 years referred for the assessment of suspected angina due to coronary heart disease from 12 cardiology chest pain clinics across Scotland. The study participants were randomly assigned (1:1) to standard care plus CCTA or standard care alone. The full protocol and results of this randomised controlled trial have been previously published[148,150,151]. SCOT-HEART was a randomised trial in which CCTA scanning was the intervention of interest. As such, patients were closely followed and management was optimised based on the results of the CCTA scan (antiplatelet and/or statin therapy, revascularisation as appropriate). Overall, a total of 1786 available CCTA scans were transferred to the OXACCT. For this aim, only 189 specific CCTAs were selected due to the characteristics of the patients. These characteristics included those with BMI ≥ 40, those with reported coronary artery calcium (CAC)

score of ≥ 400, and those with significant metallic artefact within the pericardium (metallic valves and/or pacing wires in-situ). The remainder of the SCOT-HEART Trial patients were utilised in subsequent aims of this thesis.

2.2.2. Epicardial adipose tissue clinical associations: SCOT-HEART Trial & AdipoRedOx Study

To understand the clinical associations of EAT volume with long term cardiovascular disease risk, the DLN for the automated quantification of EAT was applied into the full AdipoRedOx Study and SCOT-HEART Trial populations. These studies are described in general above, so only relevant information for this aim is included here.

AdipoRedOx Study

There were 253 AdipoRedOx Study patients included in the longitudinal clinical association study as part of this aim. Two relevant outcomes were utilised for the analysis: new onset AF occurring in-hospital following the cardiac surgery procedure with a median follow up period of 6 days, and new onset AF occurring long term out of hospital with a median follow up period of 90 months. The in-hospital outcomes were collected locally by research staff at the John Radcliffe Hospital, Oxford, utilising review of written and electronic medical records, and automated telemetry monitoring including continuous ECG monitoring during intensive care unit admission and daily ECGs plus additional ECGs on report of symptoms. Any episode of AF lasting >30 s was classified as post-operative AF[152].

The long term AF outcomes were sourced from NHS Digital Hospital Episode Statistics (HES). These data were collected through data sharing agreement NIC-392669 (University of Oxford and NHS Digital) and included all diagnostic codes for AF from inpatient admissions episodes, outpatient clinic episodes, and emergency care episodes. No primary care data were available for inclusion in this analysis.

SCOT-HEART Trial

All patients in the SCOT-HEART Trial who were randomised to receive CCTA were included in the analysis of outcomes relevant to EAT volume. Data obtained at both baseline and at the most recent five-year follow up, as reported previously[151], were utilised for analysis. Data for 1558 unique patients were included. The baseline variables utilised for cross-sectional association were CAD and AF (prevalent CAD and AF at the time of randomisation).

The five-year follow up data utilised in the analysis were all-cause mortality, non-cardiac mortality, non-fatal myocardial infarction (MI), and non-fatal stroke.

The cross-sectional outcome of AF was defined as any documented diagnosis of paroxysmal, persistent, or chronic AF within a patient's medical history, including self-report by the patient at the time of enrolment in the study. Outcomes within the SCOT-HEART Trial are all defined within the Study Protocol[148,153]. Importantly, MI was defined as per the Universal Definition of the Joint ESC/ACCF/AHA/WHF Task Force for the Redefinition of Myocardial Infarction[154]. For the purposes of this study, the investigators followed the original SCOT-HEART investigators adjudication of events. This adjudication occurred according to the guidelines of the ACC/AHA[155] and the Academic Research Consortium for definition of the cause of death[156]. Cardiac mortality within the study was defined as 'any death due to proximate cardiac causes (e.g., myocardial infarction, low-output heart failure, fatal arrhythmia)'[148]. Sudden cardiac death was also included within this group[155,156]. Any other death that was not captured by the previous definition was classified as of non-cardiac cause. Non-cardiac deaths included vascular causes such as stroke and embolism within the pulmonary circulation. Deaths were classified as 'deaths of unknown cause' only when information on the exact cause could not be obtained with certainty, which was at the discretion of the local SCOT-HEART investigators. The protocol of the SCOT-HEART study required that all adjudication of study events and cause of death be performed at each local site by the local study investigators. These investigators were not involved in the CT scan analysis or the statistical analysis of the study data. Events were ascertained via patient chart review, death certificate review and/or telephone follow-up or in-person verification with family members. No patient events were re-adjudicated for this study. All information utilised in this study was sent for statistical analysis to an independent team based at the University of Oxford, United Kingdom, which performed all analyses blind to scan data.

2.2.3. The Atriomic Risk Algorithm: The ORFAN Arm 4 Study

The preliminary development of the radiomic version of the Atriomic Risk Algorithm was developed utilising cohorts from individual sites of the ORFAN Arm 4 Study.

The Royal Brompton and Harefield and Oxford sites were utilised for training and internal testing of the algorithm, while the Papworth NHS Foundation Trust cohort was utilised for external testing.

The ORFAN Arm 4 Study is described in detail in Chapter 3, including the development of the study, the ethical approvals, the local investigators, the number of patients from each site, and the methodology of data and CCTA transfer. For this analysis, outcomes were sourced through UK wide registry data which the entire ORFAN Arm 4 cohort was linked to. Critical for the training of the radiomic stroke risk algorithm were the NHS Digital HES data which included all hospital events for ischaemic stroke, and the national stroke event registry data (SSNAP).

2.3. Coronary computed tomography angiography protocols

This section describes the methodology and standardised operating procedures for the acquisition of CCTA scans in the aforementioned studies. Since some studies were performed in Oxford (i.e., local ORFAN Arm 4 patients & AdipoRedox) by myself and other members of the research team, whereas others were performed in external centres (external ORFAN Arm 4 sites and SCOT-HEART sites), a slightly modified protocol was used in each imaging centre in accordance with local expertise, hardware, and guidelines.

John Radcliffe Hospital, Oxford, United Kingdom: Participants in the ORFAN (clinically-indicated CCTA at baseline) and AdipoRedox (research CCTA patients undergoing clinically indicated cardiac surgery) studies underwent CCTA using a 64-slice scanner (General Electric LightSpeed Ultra, General Electric Healthcare, Chicago, USA) (Figure 2.1). Heart rate was optimised using intravenous injection of beta blockers to achieve a heart rate of less than 65 beats per minute, unless medically contraindicated. Sublingual glyceryl-trinitrate (800ug) was also administered to achieve maximum coronary vasodilatation, unless contraindicated or declined by the patient. In the AdipoRedox Study (but not ORFAN) a non-contrast prospectively ECG triggered axial acquisition CT scan was obtained first (0.35 sec rotation time, 2.5 mm axial slice thickness, 20 mm detector coverage, tube energy of 120 kVp and 200 mA) with the carina and the diaphragm used as cranial and caudal landmarks respectively.

Lung field of view was extended to cover the entire thoracic soft tissue (for subcutaneous adipose tissue analysis). In all studies, CCTA was performed following intravenous injection of 60-95ml of iodine-based contrast medium (Niopam 370, Bracco Imagin, Milan, Italy) at a flow rate of 5-7 mL/sec (tube energy of 100 or 120 kVp, axial slice thickness of 0.625 mm, rotation time of 0.35 sec). The selection of these technical parameters was based on body type to ensure optimal image quality. Prospective image acquisition was used in most cases by ECG-gating at 75% of cardiac cycle (with additional padding if needed). In the presence of irregular rhythm or high heart rate not responsive to beta blockers, retrospective gating was used as a last resort.



Figure 2.1. John Radcliffe Hospital CT Scanner. *The 64-slice General Electric LightSpeed Ultra (General Electric Healthcare) system in the John Radcliffe Hospital, Oxford, UK.*

Manor Hospital, Oxford, United Kingdom: Clinical participants in the ORFAN study underwent CCTA using a 320-slice scanner (Aquilion ONE, Toshiba Medical Systems (now Canon Medical Systems), Otawara, Japan) (Figure 2.2). Heart rate was optimised using intravenous injection of beta blockers to achieve a heart rate of less than 65 beats per minute, unless medically contraindicated. Sublingual glyceryl-trinitrate (800ug) was also administered to achieve maximum coronary vasodilatation, unless contraindicated or declined by the patient. Between 60 and 80 ml of iodine-based contrast medium (Niopam 370, Bracco Imaging, Milan, Italy) were given at a flow rate of 5-7 mL/sec (tube energy of 100 or 120 kVp, axial slice thickness of 0.50 mm, tube rotation 0.35 sec). A standard CCTA reconstruction algorithm was used to generate a set of images at the 75% phase of the heart cycle as well as the best phase +/- 10% temporal window. Prospective image acquisition was used in most cases by ECG-gating at 75% of cardiac cycle. In the presence of irregular rhythm or high heart rate not responsive to beta blockers, retrospective gating was used as a last resort.



Figure 2.2. The Manor Hospital CT Scanner. *The Toshiba Aquilion ONE CT scanner in The Manor Hospital, Oxford, United Kingdom.*

Bath (ORFAN Arm 4 UK): Clinical patients undergoing CCTA at Royal United Hospitals Bath NHS Foundation Trust were scanned on either a 64-detector row scanner (Edge, Siemens, Munich, Germany) or if scanned after 2016 a 256-detector row scanner (Drive; Siemens, Munich, Germany). All scans were performed from the carina to the base of the heart and were prospectively gated unless the patient was scanned on the Edge and had a very irregular heart rate, in which case a retrospectively gated scan would be performed as a last resort. All Drive CCTA were prospectively gated. The Edge had 64 physical detector rows but 128 channels and used z-sharp technology to effectively allow 128-slice scanning. For both scanners the heart rate was optimised to be 60bpm or below using the intravenous injection of beta-blockers or ivabradine if beta blockers were contraindicated. Sublingual glyceryl-trinitrate was given to all patients immediately prior to the scan. For the Edge, a 20ml test bolus of iodine contrast medium, Niopam 340 (Bracco Imaging S.p.A, Milan, Italy) or Visipaque 350 (GE, Milwaukee, USA), was given initially, followed by 80ml of the same contrast for the main acquisition. Further 80mls would be given if a repeat was necessary. The Edge used an injector pump (Medrad Inc., Warrendale, USA). For the Drive, a 12ml test bolus of iodine contrast medium (Omnipaque 340, GE, Milwaukee, USA) was given followed by 60ml of the same contrast medium for the main acquisition, with a possible further 60mls for repeat dosage if necessary. An injector pump (Bracco Imaging S.p.A, Milan, Italy) was used for the Drive. All contrast

was given at 5.0ml/s with a sodium chloride chase. The Edge has a rotation time of 0.28sec with an axial slice thickness of 0.5mm and a 500mm detector coverage. A tube voltage of 120kVp was used for most patients with 140kVp sometimes used for patients with a larger habitus at the discretion of the radiologist. The Drive has a 0.28s rotation time with a 0.6mm axial slice thickness. The detector coverage was between 300-500mm. The tube voltage was at the radiologist discretion between 70 to 140kVp, with most scans performed at 100kVp or 120kVp.

Leicester (ORFAN Arm 4 UK): Clinical patients undergoing CCTA at University Hospitals of Leicester NHS Trust were scanned on a 128-detector row scanner (Definition Flash, Siemens, Munich, Germany). All scans were performed from 6 slices above the coronary artery ostia to 6 slices below the base of the heart. IV beta blockers were used for heart rate optimisation at 60bpm. Iodine contrast medium (Iomeron 350, Bracco, Milan, Italy) was given intravenously, with an 80ml dosage given at 6mls/s for the main acquisition. The rotation time of the scanner was 280ms, with an axial slice thickness of 0.75mm. The tube voltage was set at 100kVp and could be altered at the radiologist discretion to be 80, 120 or 140kVp based on patient factors. All scans were prospectively gated at 60-75% of cardiac cycle.

Milton Keynes (ORFAN Arm 4 UK): Patients undergoing clinical CCTA at Milton Keynes University Hospital NHS Foundation Trust were scanned on either a 64-slice scanner if prior to 2022 (Prime, Canon Medical Systems, Otawara, Japan), or a 160-slice scanner if after January 2022 (Prime SP, Canon Medical Systems, Otawara, Japan). All scans were performed from carina to the cardiac apex and were prospectively gated with reconstruction at 75% of cardiac cycle and a 'best phase + window'. Local protocol for heart rate optimisation was 2 days of ivabradine prior to the scan if HR >65bpm (7.5mg oral BD), or if HR <65bpm 5mg oral BD on the day prior to the scan and on the morning of the scan. Intravenous beta blockers were utilised for those in AF and if HR was >60bpm. Iodine contrast medium (Omnipaque 350, GE, Milwaukee, USA) was utilised with 63ml the standard dose for a patient scanned at a tube voltage of 100kVp (flow rate of 4.5ml/s). If scanning at 120kVp due to a larger patient, then 77ml of contrast was given (flow rate of 5.5ml/s) and if a smaller patient, then a tube voltage of 80kVp was utilised with 49ml of contrast given (flow rate 3.5ml/s). The rotation time was 275ms, with a slice thickness of 0.5mm.

Royal Brompton & Harefield (ORFAN Arm 4 UK): Clinical patients undergoing CCTA at the Royal Brompton and Harefield Hospitals were scanned on a 256-slice count scanner (Somatom Definition Flash, Siemens, Munich, Germany). All scans were performed from the carina to the posterior heart border with no field of view limitations. Heart rate was optimised with intravenous beta blocker in 5mg increments with a target heart rate of 60bpm. Glyceryl-trinitrate was given sublingually for

maximal vasodilation if not contraindicated. Iodine based contrast medium (Iomeron 400, Bracco, Milan, Italy) was utilised unless the patient's renal function was compromised, in which case Visipaque 320 (GE, Milwaukee, USA) was used. For small (BMI <20) and average sized (BMI 20-28) patients a 15ml test bolus followed by a 60ml main acquisition bolus was used. For large patients (BMI >28) a 15ml test bolus followed by an 80ml main acquisition bolus was used. If an 18 gauge cannula was sited then the contrast was given at 6ml/s, or if a 20 gauge cannula was sited then it was given at 5ml/s. For small patients a tube voltage of 80kVp was utilised with quality reference of 400mAs, for average size patients 100kVs was used with a quality reference of 300mAs, and for large patients 120kVp was utilised with a quality reference of 200mAs. The rotation time of the scanner was 280ms, with a pitch of 3.4 and an axial slice thickness of 0.75mm. For adaptive sequential acquisition all parameters remained the same with a feed of 34.5mm. Prospective gating was used for all acquisitions, with high pitch helical (flash) acquisition at 70% R-R interval, adaptive sequential end diastolic acquisition at 60-80% R-R interval, and adaptive sequential end systolic timing all available.

Royal Papworth Hospital (ORFAN Arm 4 UK): Clinical patients undergoing CCTA at the Royal Papworth Hospitals were scanned on a 2 x 192 slice count scanner (Somatom Force, Siemens, Munich, Germany). All scans were performed from the carina to the posterior heart border with no field of view limitations. Heart rate was optimised to be 60bpm with intravenous beta blocker. Glyceryl-trinitrate was given sublingually for maximal vasodilation if not contraindicated. Iodine based contrast medium (Iomeron 400, Bracco, Milan, Italy; or Omnipaque 350, GE, Milwaukee, USA). For small and average sized patients 85ml of contrast was utilised for the main acquisition. For large patients a 90ml main acquisition bolus was used. Contrast was given at 5ml/s for standard patients, or if an 18-gauge cannula was sited then the contrast was given at 6ml/s. Tube energy was set as standard at 100kVp with clinician discretion. The rotation time of the scanner was 250ms, with an axial slice thickness of 0.75mm.

Cleveland (ORFAN Arm 4 USA): CCTA scans were performed using either a 256-slice Brilliance iCT scanner (Philips Medical Systems, Best, The Netherlands), a 2 x 128-slice Definition Flash scanner (Siemens Healthcare, Erlangen, Germany) or a 2 x 192-slice Somatom Force CT scanner (Siemens Healthcare, Forchheim, Germany). In patients with heart rate > 60bpm, 5 mg of intravenous metoprolol (with incremental 5 mg doses up to a maximum dose of 30 mg) or intravenous diltiazem (5 mg increments up to 20 mg maximum), if the heart rate remained above 60 beats per minute once the patient was positioned on the CT table. Patients also received 0.3 mg of nitroglycerin

sublingually immediately before CCTA and iodine contrast (Omnipaque 350, General Electric, Milwaukee, USA) was administered at flow rate of 5-6 ml/s.

SCOT-HEART Trial: CCTA scans were performed using either two 64 detector row scanners (Brilliance 64, Philips Medical Systems, Netherlands, and Biograph mCT, Siemens, Germany) and one 320 detector row scanner (Aquilion ONE, Toshiba Medical Systems, Japan) at three imaging sites according to each site's local protocol[148,150,153].

2.4. The Oxford Academic Cardiovascular Computed Tomography (OXACCT) Core Lab

For all research studies included in this thesis, image analysis was performed in the OXACCT Core Lab (Acute Multi-disciplinary Imaging and Intervention Centre, John Radcliffe Hospital, Oxford, U.K.,). I designed a set of systems and standardised operating procedures to govern the processes relating to image transfer, storage, analysis and data control in conjunction with other OXACCT team members. In all cases, scans were collected from their source as dicom files and transferred to the Core Lab as described below. A server dedicated to ORFAN Study CT scans was used to create an accessible library of all research CCTA scans. Access was restricted to authorised users affiliated with the OXACCT Core Lab and at all stages only pseudo- or fully anonymised data were used for analysis.

Analysis for the work in this thesis were performed on two systems. 3DSlicer[142] (version 4.10.0 through 4.11.0) was used initially for manual segmentation of whole heart, LA and peri-LA segments, before shifting to the bespoke research image analysis software CaRi-Research (the research version of the CaRi-Heart® system) version 1.0 through 2.4.2 (Caristo Diagnostics Ltd., Oxford, UK).

2.5. CT scan transfer

The transfer of the CCTA scan to the ORFAN Core Lab at the University of Oxford is an integral part of this thesis, and in particular within the ORFAN Arm 4 Study. Transfer of data from within the ORFAN Arm 4 study has been coordinated in conjunction with the Oxford University spin-out company Caristo Diagnostics Ltd. (Oxford, UK). The research team would bring each site's clinical and technical team into contact with the project lead at Caristo Diagnostics so that they could support the site to install a scan gateway locally at each site. Caristo work with Cimar UK Ltd to provide the CT transfer platform for the Study. Cimar UK Ltd (an acronym meaning Chartered Image Management And Reporting) is a separate company specialising in the large transfer of medical data via secure cloud-native means. The transfer of CT data was Caristo's contribution in-kind to this study.

All CT scans from each NHS site and most international sites are sent to the Oxford Core Lab via the locally managed Cimar Gateway (virtual server provided to each Trust by Caristo Diagnostics Ltd and residing within the Trust's own network). The process is demonstrated graphically in Figure 2.3.

Access to Cimar Gateway is only available to each Trust's authorised personnel with occasional monitored maintenance access provided as required. The Cimar gateway uploads the data to Cimar Cloud via a 256-bit encryption. Once the data is uploaded to Cimar Cloud, it is automatically purged from the Cimar Gateway. In theory, individuals with administrator access to Cimar Cloud can see data containing personal identifiable information, however, this access is limited to only certain individuals who need to support the data transfer. All access to Cimar Cloud has access restricted by username and password with audit logging enabled for access to data. Cimar Cloud also requires multi-factor authentication which is necessary for all administrator users. The data in Cimar Cloud is pseudo-anonymised and then saved to the University of Oxford servers (off-line). The data is only stored for a short duration on Cimar Cloud (purged ~7-14 days after it is saved onto the University of Oxford servers) with no ability to recover the data. Both Caristo Diagnostics Ltd. and Cimar are GDPR complaint with ISO 27001 certification.

At the time of upload at the local site to the Cimar Gateway the CT scan is stripped of all patient identifiers and the assigned unique ORFAN Study participant ID is inserted in place of the patient's name. This fully automates the pseudonymisation of the clinical scans. The pseudonymisation rules automatically remove secondary captures and all screengrabs that may contain burnt-in patient information. All ECG data that commonly contains patient information are automatically removed. In total some 226 possible DICOM tags that could provide patient identifiable data are removed at the time of upload to the portal. Only technical tags such as scan slice thickness, phase, kilovoltage, are retained for analysis purposes.

The gateway system allows each NHS Trust site to easily send their ORFAN Arm 4 CCTA scans to the Oxford Core Lab for analysis within the study. Certain longstanding international collaborating sites such as the Cleveland Clinic, USA, site of ORFAN Arm 4 had transferred their data to the Oxford Core Lab using encrypted external hard-drives prior to the formal establishment of the ORFAN Arm 4 Study and the adoption of the Cimar gateway system.

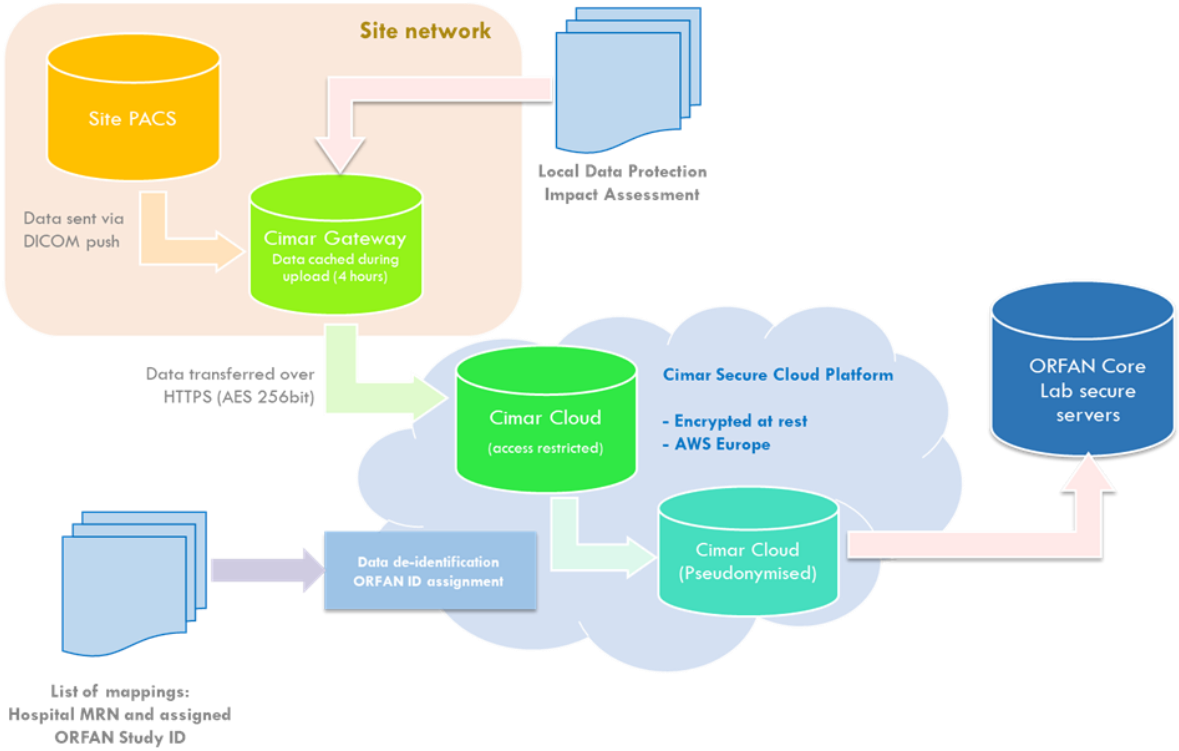


Figure 2.3. The flow of CCTA data from the NHS Trust environment to the Oxford ORFAN Core Lab via the Cimar gateway. *Data is pushed or retrieved from the local site network PACS. Data is sent via a dicom push to the Cimar gateway and uploaded to the Cimar Cloud, which are encrypted Europe based servers. Within the cloud the pseudonymisation occurs, with all patient identifiers including dicom tags stripped from the scan, and the ORFAN Study ID inserted as the only scan identifier. Once pseudonymisation is complete the scan is pushed to the ORFAN Core Lab server at The University of Oxford. All data in the cloud is purged regularly and no cached data is held.*

2.6. Quality control of image datasets

The first step in the processing and image analysis of the obtained clinical CCTA scans was the quality control check, based on the guidelines of the Society of Cardiovascular Computed Tomography for the interpretation and reporting of CCTA. Based on these recommendations, CCTA should preferably be acquired as isotropic submillimeter 3-dimensional electrocardiography-gated data sets to enable reliable reconstruction of the obtained images and display in multiple formats[157]. The checklist used for the quality control of the CCTA scans is presented in Table 2.1. If a CCTA scan was deemed to be of low quality unsuitable for analysis then it was excluded from any inclusion in the research analysis. For instance, scans that were performed at tube voltage settings other than 120 or 100 kVp (most commonly used and recommended settings in most centres and scanners) would be excluded from inclusion in some analysis that relies upon attenuation-based metrics as these are not adequately validated at these settings. Similarly, CCTA scans with suboptimal field of view of the entire heart from the bifurcation of the pulmonary trunk superiorly to the apex of the heart at the diaphragm were excluded. Images with minor quality issues such as <5 missing slices, minor motion blur, minor stair-step or blooming artefacts were retained in analysis.

Table 2.1. Technical review and quality assurance of obtained coronary CT angiograms in the ORFAN Arm 4 Study

Field	Checklist
File format	Are the images stored as DICOM files (or related DICOM-files, e.g. .IMA files)?
Anonymisation	Do the DICOM files contain non-anonymised or pseudo-anonymised data?
Date and time	Does the Study ID, date and time of the scan match what is expected?
CT scanner	Scans should be performed using 64-slice CT scanners or higher
Contrast	Is there adequate opacification of the aortic root? (mean HU of aortic root lumen >200)?
Completeness	Are all slices present?
Image window	Does the window extend from the bifurcation of the pulmonary artery to the apex of the heart?
Slice thickness	Was the scan obtained as an isotropic submillimeter 3-dimensional ECG-gated data set?
Tube voltage	Ideally, tube voltage should be set at 120 or 100 kVp (analysis dependent)
Motion artefact	Are there any motion or step artefacts affecting the interpretation of the scan?
Blooming artefact	Are there artefacts caused by stents/pacemaker leads etc. that affect the interpretation and visualisation of adjacent coronary vessels?
Other artefacts	Are there other issues with the scan?
Patient factors	Are there patient factors rendering this scan unusable? (i.e. extreme body habitus or extremely unusual anatomy)

CT: Computed Tomography; ECG: electrocardiogram; HU: Hounsfield Units.

2.7. The automated extraction of radiomic features from CT volumes

The methods concerning the training of automated networks for the segmentation of CCTA regions of interest are detailed in Chapter 3 of this thesis. Here, I present the general methods for the extraction of radiomic features from CCTA segments. The extraction of radiomic features required an automated approach to avoid the manual work of extracting and saving the features scan by scan and segment by segment. A graphical user interface (GUI) was created to facilitate the automated extraction of radiomic features for bulk segments. This tool was named the RadExtractor by our research group. Created by Dr Musib Siddique and me, the RadExtractor enables the selection and merging of discreet segmentations that have been undertaken on a CT scan and the selection of custom Hounsfield unit (HU) ranges for voxel selection. This is very useful for adipose tissue, as adipose is detected within the window of −30 to −190 Hounsfield Units (HU) on CT.

The tool allows greater utility than traditional manual extraction of radiomic variables as segments can be merged to create permeations of discreet segments for analysis. For example, segment A and segment B may have been segmented independently but are adjacent segments. The RadExtractor allows the user to extract the radiomic features of segment A and B independently of each other,

but also to merge the regions and extract the overall features for the region combined. The extracted radiomic features are saved as .csv files for each segment.

2.8. Radiomic features

As explained in Chapter 1, the radiomic features that can be extracted from CCTA segmentations are already defined within the literature. For each segment 872 radiomic features can be calculated, ranging from shape-related to first- and higher-order (texture) statistics using the SlicerRadiomics extension which incorporates the Pyradiomics library into 3D Slicer[141]. All definitions of radiomic features are included in Appendix 1.

A key aspect of radiomic methodology is the selection of the range and number of HU bins for discretisation. This depends upon two factors: a) the range of HU present in the original segmentation; and b) signal/noise ratio. As an example, for a tissue segmentation with HU values in the range of 30 to 190 HU (160 HU-wide range), voxels can be discretised into two bins (each with a range of 90 HU), four bins (each with a range of 40 HU), or even 160 bins with a range of a single HU value. The latter approach introduces a significant degree of mathematic complexity in the computation of the radiomic features. Treating HU values of 50 and 51 HU as significantly different decreases the signal-to-noise ratio of radiomic measurements by introducing potential noise. On the other hand, using just two bins leads to loss of potentially useful information since HU values of 30 and 80 fall into the same bin. Once the volume has been segmented and discretised into bins, the next step involves the generation of the mathematical matrices[139-141,158] (Figure 2.4).

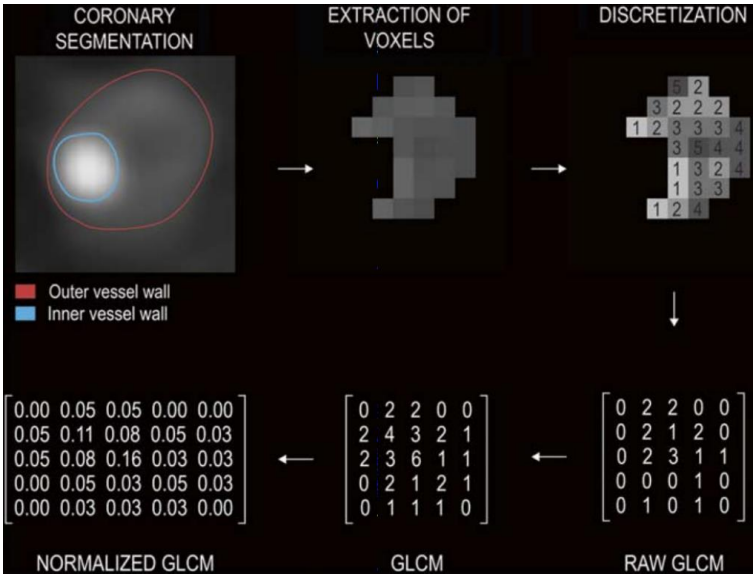


Figure 2.4. Creating a gray level co-occurrence matrix. *This example demonstrates the generation of a gray level co-occurrence matrix (GLCM) for a segmented region of a CCTA, in this case a coronary artery plaque. Firstly, the voxels of interest are extracted (based on manual or attenuation-based segmentation which can be performed either manually or automatically). The voxels are then split*

into n different value groups, known as "bins". For a given direction and distance (e.g. distance 1 voxel and angle 0 degrees) we calculate the number of times a value j is found to the right of value i. This number is then inserted into the ith row and jth column of the raw matrix. To ensure symmetry, the transpose is added to the raw matrix. Each number is then substituted by its frequency to generate a normalised GLCM. To ensure rotationally invariant results, the same process is repeated in all directions and then the averaged GLCM is computed and used to generate GLCM-derived radiomic features. ***Reproduced with permission from Kolossvary, M., et al. J Thorac Imaging, 2018. 33(1): p. 26-34.***

Gray level co-occurrence matrix (GLCM): First described by Haralick et al in 1973[159], GLCMs are second-order statistics, which means that they examine the relationship between two voxels. For a given angle (e.g., 0, 90, 180 or 270 degrees) a GLCM (i, j) matrix is calculated by computing the number of times a value j occurs to the right (or bottom, left, top) of value i. To compute rotationally invariant results, the transpose is added to the raw GLCM and the absolute values are then substituted by the frequency values. Using this normalised GLCM, multiple metrics can be calculated as described in Appendix 1. For instance, GLCM entropy will increase in the presence of multiple types of gray-level transitions in a volume. On the other hand, correlation assesses the linear dependence between adjacent voxels throughout a given volume, which is inversely proportional to an image's contrast (i.e. it is helpful to think of checkerboard with alternating black and white squares/voxels resulting in low correlation and high contrast)[139-141,158].

Gray level run-length matrix (GLRLM): Contrary to second-order statistics, higher-order statistics examine the spatial relationship between three or more voxels. The easiest concept within this group is that of GLRLM matrices, which assesses how many voxels are next to each other with same value. In the GLRLM matrix, the rows refer to the attenuation values (bins) and the columns the run lengths. For instance an element in the ith row and jth column describes the number of times in the volume that i value voxels are next to each other j times. Similarly to GLCM, GLRLM are calculated in all directions and then normalised. Based on the definitions used to calculate the individual features in this group (Appendix 1 & Figure 2.5), different weightings can be applied to short versus long runs and high versus low gray level values, thus describing different attenuation/texture patterns in the volume of interest[139-141,158].

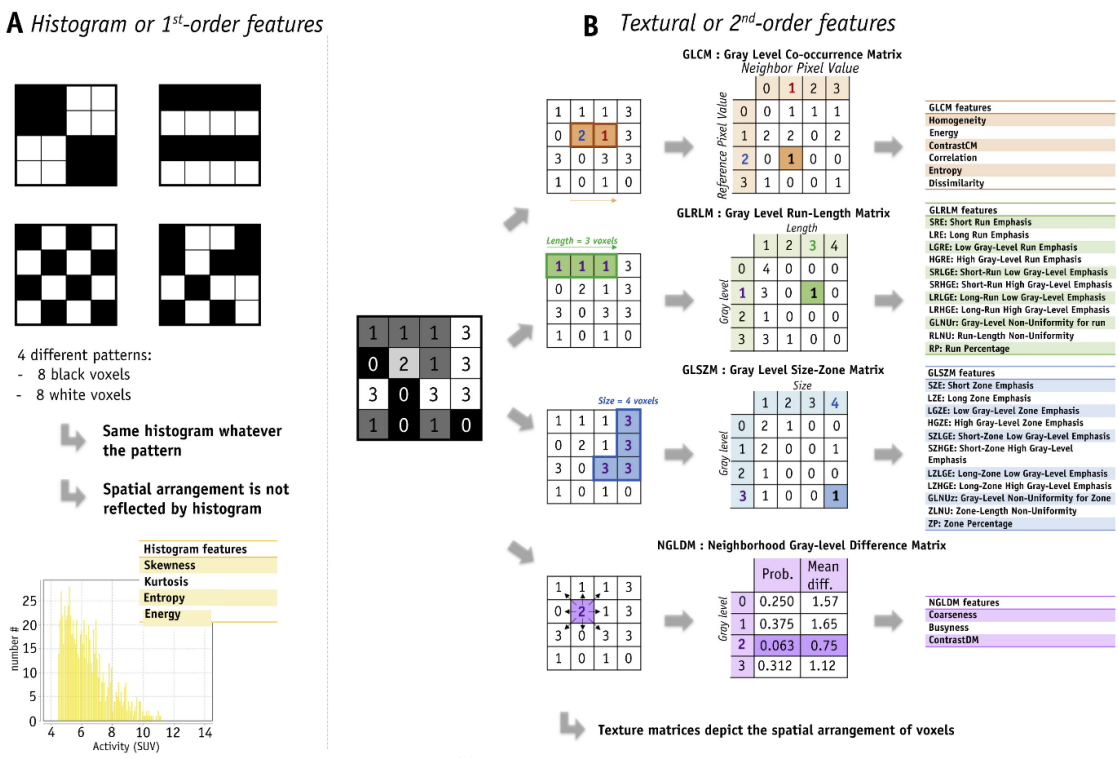


Figure 2.5. Radiomic characterization of textural features. *For a given volume of interest, histogram-based first-order features only reflect the voxel attenuation distribution. Different texture patterns (same number of voxels with similar attenuation values but different location) may still have identical histogram and therefore similar first-order statistics, such as mean attenuation, skewness, kurtosis. These differences can instead be described using higher-order features derived from texture matrices that reflect the unique spatial arrangement of voxels.* ***Reproduced with permission from Reuze, S., et al. Int J Radiat Oncol Biol Phys, 2018. 102(4): p.1117-1142.***

<u>*Neighbouring gray tone/level different matrix (NGTDM):*</u> The neighbouring gray tone/level different matrix is a one-dimensional matrix, where each gray level entry in the matrix (defined as gt) is the sum of the differences between all the voxels with gray-level value gt and the mean gray-level value of its neighbourhood (as described in the Appendix 1 & Figure 1.5). This enables the calculation of radiomic features that described the visual properties of an image, such as coarseness, contrast, busyness, complexity and strength. These metrics are close to the visual interpretation of texture by humans. For instance, coarseness is a quantitative metric that reflects the difference in the radiodensity levels of neighbouring regions, whereas strength has even been associated with the visual aesthetics of an image, with higher texture strength linked to more attractive images to the human observer [139-141,158].

<u>*Gray level size zone matrix (GLSZM):*</u> Using different approaches and matrices more features can be defined. For instance, gray level size zone matrices (GLSZM) are similar to GLRLMs but instead of focusing on one-dimensional lengths, they focus on two-/three-dimensional zones of connected voxels (Appendix 1 & Figure 2.5) [139-141,158].

Gray level dependence matrix (GLDM): Finally, a gray level dependency is defined as the number of connected voxels within distance d that are dependent on the centre voxel. A neighbouring voxel with gray level j is considered dependent on centre voxel with gray level i if $|i-j| \leq \alpha$ (usually d=1 and a=0). Thus, in a GLDM, the (i,j)th element refers to the number of times a voxel with gray level i with j dependent voxels in its neighbourhood appears in image (Appendix 1 & Figure 2.5).

Shape-related features: Contrary to attenuation or texture-based radiomic features, shape features describe the dimensions, shape, surface, volume and other geometric features of a segment volume of interest. As such, they are independent of the attenuation values and derived from simple euclidean geometric analysis of a segmentation's shape (Appendix 1 & Figure 2.5)

Wavelet transformations: Radiomic features can be calculated using both the original images as well as mathematical transformations of the original data. An image can be transformed from the original spatial domain into the frequency domain without any information loss. Rather than assigning radiodensity values to spatial (x, y, z) coordinates, the frequency domain describes the pattern and rate at which attenuation changes along spatial directions. Wavelet transforms can transform an image into the frequency domain by retaining spatial information. They rely on a window function, where the window width reflects both the spatial resolution (localisation precision) and frequency resolution (global image features). Wavelet transformation decomposes the data into high and low-frequency components (narrow and wide window respectively) (Figure 2.6). At high frequency (shorter time intervals), the wavelets can capture discontinuities, ruptures and singularities in the original data. At low frequency (longer time intervals), the wavelet characterises the coarse structure of the data to identify the long term trends. Thus, the wavelet analysis allows extraction of hidden and significant temporal features of the original data, while improving the signal-to-noise ratio of imaging studies[143,160]. High- and low-frequency wavelet transformations can be applied in all directions (x, y, z) resulting in eight additional datasets: HHH, HLL, HHL, HLH, LLL, LHH, LHL and LLH. Attenuation and texture-based metrics can be calculated for all of these wavelet transformations. However, since they are independent of gray level intensities, shape-related statistics are consistent across all wavelet transformations and the original CT image and are therefore calculated only once.

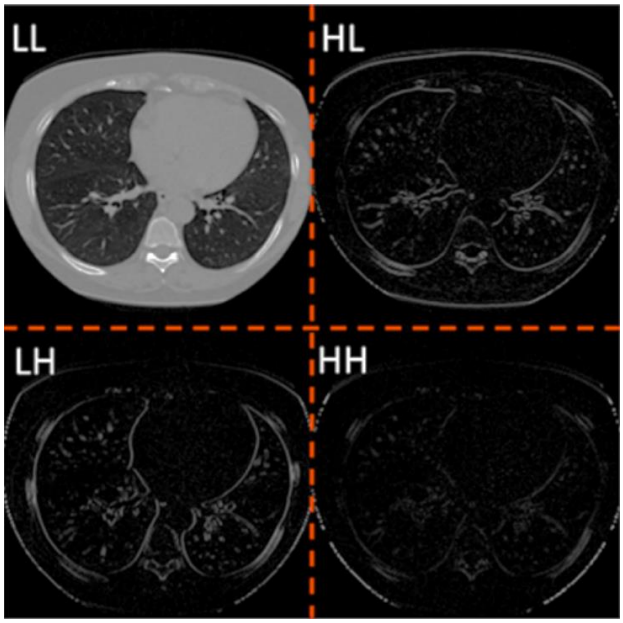


Figure 2.6. Wavelet decomposition of an axial slice from computed tomography scan. *Four wavelet transformation of the original image using high- (H) and low- (L) frequency wavelet filters along the x and y axis. In this example of a patient with lung nodules, wavelet transformations may help to distinguish true nodules from normal parenchymatous injuries or other tissues.* ***Reproduced with permission from Orozco M et al. BioMedical Engineering OnLine (2015) 14:9.***

2.9. Agatston coronary calcium score

When obtained for use as co-variates in risk models (e.g. SCOT-HEART Trial studies), non-contrast ECG-gated cardiac CT scans were used to: a) achieve a preliminary assessment of cardiac anatomy and presence of coronary or extra-coronary calcification; and b) get a quantitative assessment of coronary calcium by means of the coronary calcium score (CCS). CCS was computed using commercially available software (Aquarius Workstation® V.4.4.11-13, TeraRecon Inc., Foster City, CA, USA) using the Agatston method that takes into account the volume as well as density of coronary calcifications (Figure 2.7)[161,162]. Voxels with attenuation values that exceeded 130 HU on non-contrast CT scans corresponding to calcium were highlighted by the semi-automated software. The reader then scrolled through the axial slices of the cardiac CT scan and reviewed each highlighted area, before circling it and assigning it to a coronary territory (left main, LAD, RCA, LCX). The values for each vessel distribution were then summed to generate the Agatston CCS. Extra-coronary calcifications in the aorta, valves, pericardium or myocardium were excluded from the CCS[157,161].

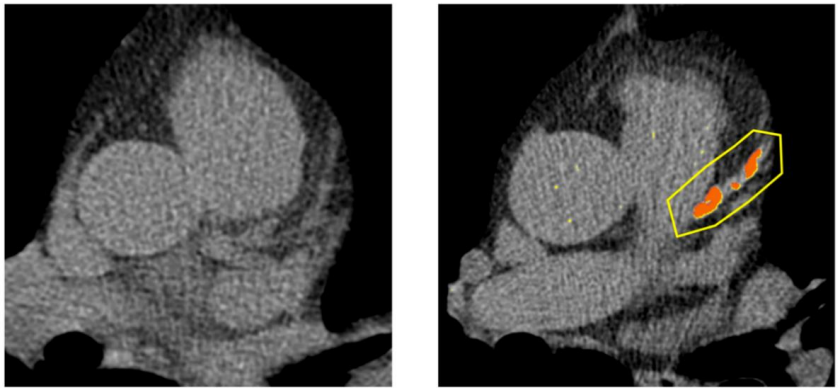


Figure 2.7. Agatston coronary calcium scoring. *Example image of automated calcium quantification on CCTA. On thick slice (2.5-3mm axial thickness) non-contrast cardiac CT scans voxels within the walls of the coronary vessels with a Hounsfield (HU) value of >130 are selected and a score is assigned to each segment based on its volume and attenuation (higher attenuation = higher score). The sum of the scores is then computed to produce the overall Agatston CCS. In the left panel the CCS = 0, in the right panel calcium is detected (orange overlay) within the left anterior descending artery. The calcium has been auto segmented as part of the Agatston CCS. The yellow box is highlighting the vessel with calcium and is added for readers ease only. The segmented calcium can extend in 3 dimensions across slices.*

2.10. Qualitative coronary analysis

This thesis does not focus on traditional coronary analysis of CCTA images, however, stenosis analysis is included as a co-variate in some models of risk for relevant outcomes.

The majority of work to ascertain traditional coronary artery assessment on CCTA was not performed by myself, but by other members of the OXACCT Core Lab.

Because coronary anatomy can be complex, motion or calcium-related (e.g. blooming) artefacts are frequent, and coronary lesions are often morphologically subtle, CCTA scans were reviewed on a dedicated cardiac-specific workstation (Aquarius Workstation® V.4.4.11-13, TeraRecon Inc., Foster City, CA, USA) in various, 2- and 3- dimensional reconstruction formats. This included a review of the 2-dimensional image stacks (i.e. axial, sagittal and coronal) as well as multiplanar reformations (MPR) around the tracked centreline of a given coronary vessel. In the presence of heart-rate related artefacts where padding or retrospective gating were performed, alternate times in the cardiac cycle were reviewed to identify the optimal phase for analysis. High risk plaque features, quantitative coronary plaque analysis and perivascular adipose tissue characterisation with the Fat Attenuation Index[104,163] or the Fat Radiomic Profile[144] were not utilised in the work relevant to this thesis.

2.10.1. Coronary artery stenosis assessment

When reviewing the coronary tree on CCTA for obstructing lesions, a systematic approach was followed by segmenting the coronary vasculature into 18-segments as recommended by the Society of Cardiovascular Computed Tomography (Figure 2.8)[164]. An initial examination was performed for

the course and branching of the main coronary vessels. Then, the lumen and wall of the coronary vessels were examined for the presence of luminal plaque and the maximal percent diameter stenosis was reported. The assessment of stenosis severity was performed using manual measurement tools on the reconstructed images and was used to classify the lesion in one of the pre-defined ranges/groups of stenosis severity. Briefly, mild, moderate and severe coronary stenoses were defined as lesions causing luminal stenosis 25-49%, 50-69% and ≥70%, respectively[164,165]. Where mentioned, obstructive CAD was defined as the presence of at least one coronary stenosis ≥50% on coronary computed tomography angiography (CCTA).

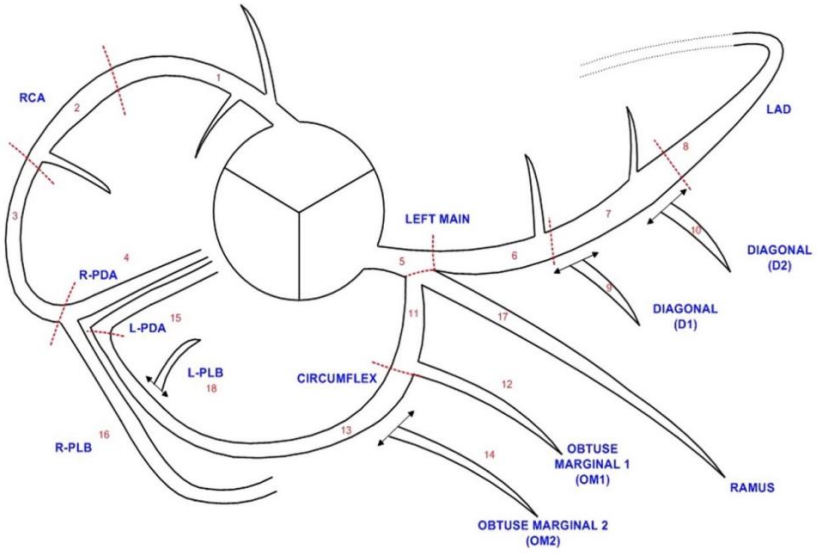


Figure 2.8. Coronary segmentation based on the recommendations of the Society of Cardiovascular Computed Tomography. *1: proximal right coronary artery (RCA); 2: mid RCA; 3: distal RCA; 4: right posterior descending artery (R-PDA); 5: left main coronary artery; 6: proximal left anterior descending (LAD) artery; 7: mid LAD; 8: distal LAD; 9: first diagonal (D1); 10: second diagonal (D2); 11: proximal left circumflex (LCX); 12: obtuse marginal 1 (OM1); 13: distal LCX; 14: obtuse marginal 2 (OM2); 15: left posterior descending artery (L-PDA); 16: right posterolateral branch (R-PLB); 17: ramus intermedius; 18: left posterolateral branch (L-PLB).* ***Reproduced with permission from Leipsic, J., et al. J Cardiovasc Comput Tomogr, 2014. 8(5): p. 342-58***

3

Chapter 3. The Oxford Risk Factors And Non Invasive Imaging Study

3.1. Introduction

The majority of data analysed for the realisation of the aims of this thesis have come from the fourth arm of the Oxford Risk Factors And Non Invasive Imaging Study (ORFAN) study ('the Study'). Firstly, this Chapter outlines the creation of the ORFAN Arm 4 Study within the larger ORFAN Study protocol, including the provisions of the protocol itself, the ethical approvals of the study, the pathway for sites to join the ORFAN Arm 4 Study and patient & public involvement considerations for the Study.

The second portion of this chapter outlines the status of the ORFAN Arm 4 Study regarding active sites, principal investigators, participant numbers per site, and the prospective outlook for the Study moving forward.

The scientific aim of this thesis that is achieved in this chapter is the development of a major clinical CCTA cohort study for the development and testing of novel CT technologies such as radiomics based stroke risk assessment.

3.2. Study history

The ORFAN Study first came into being as a prospective cohort study in February 2016. The study was, and remains, sponsored by the University of Oxford, UK. The study initially consisted of 3 separate but related research programs named 'Arms' and was operating exclusively at the John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK.

In 2018, it became clear that a larger retrospective study arm would be necessary to fully address the scientific aims of the study – explicitly to develop and validate novel imaging biomarkers of cardiovascular disease risk. The rise of big data analysis techniques for both traditional clinical datasets and in the field of radiomics was what motivated the development of the fourth arm of the Study. A cohort of hundreds of thousands of patients with relevant CT scans, risk factors from the time of the scan, and outcomes that have occurred following the CT scan would be a perfect cohort for the development of novel CT biomarkers.

The ORFAN protocol was formally amended in early 2020 to establish ORFAN Arm 4, with submission to the Research Ethics Committee (REC) and the Health Research Authority (HRA) Confidentiality Advisory Group (CAG) for study approval. Initially this amendment allowed the creation of a cohort of 75,000 patients from across England who could be included in the Arm 4 of the ORFAN Study.

In February 2022 the ORFAN Protocol was further amended to expand the ORFAN Arm 4 cohort to include up to 200,000 patients from across the UK, with an allowance for an additional 50,000 international participants within the Study also. This was approved by the REC and the HRA CAG.

3.3. Study background and logistics

3.3.1. The Oxford University ORFAN Study team and facilities

The Chief Investigator (CI) and Study Founder is Professor Charalambos Antoniades. The ORFAN Study is managed by Professor Antoniades' academic research laboratory, The Oxford Translational Cardiovascular Research Group at the Division of Cardiovascular Medicine in the Radcliffe Department of Medicine, University of Oxford.

The ORFAN project team consists of clinical researchers, post-doctoral scientists and DPhil Students. The current project manager is Ms Lucy Kingham, who joined the project that had previously been managed by me and Clinical Researcher Ms Maria Lyasheva. Ms Sheena Thomas, the Lead Radiographer for the Study and the Oxford Cardiovascular Computed Tomography (OXACCT) Core Lab provides technical support to the Study. Other active team members include Ms Parijat Patel, DPhil Student, who is working on data integration and inter-operability within the ORFAN Study. Dr Musib Siddique, a senior post-doctoral scientist in deep-learning, Dr Jagat Chauhan, bioinformatician working with the ORFAN Study data, and Dr Elizabeth Wahome, an epidemiologist who has recently joined the ORFAN Study team to contribute to clinical data analysis.

The ORFAN Study is based at the University of Oxford Acute Multi-Disciplinary Imaging and Intervention Centre (AMIIC) at the John Radcliffe Hospital, Headington, Oxford, UK. Researcher facilities and workstations are located there alongside plant installed specifically for ORFAN Arm 4 data including a major computational resource featuring a custom 12-Core Graphical Processing Unit (GPU) capable of handling very large volumes of DICOM data for deep-learning analysis, and data servers for secure storage of ORFAN study databases including segmented CT scans. These facilities are complemented by the clinical facilities at AMIIC for prospective recruitment into the ORFAN Study (Arms 1-3) including a Naeotom Alpha Photon-Counting CT Scanner (Siemens, Munich, Germany), cardiac intervention suites, and freezer facilities for recruited patient pathology samples.

3.3.2. Funding

The funding for the ORFAN Study has come from several sources, with certain funds specific to certain scientific programs within the ORFAN Study, and other funds provided for facilities and/or plant for wider research purposes. Funders include:

- The NIHR Oxford Biomedical Research Centre
- The British Heart Foundation
- Innovate UK
- The National Consortium of Intelligent Medical Imaging
- The European Commission (MAESTRIA project)
- NHS[X]

Specific to this project, the British Heart Foundation provided significant funds for scientific work to be undertaken within the ORFAN Arm 4 Study. This funding was in the form of a Translational Award (TG/19/2/34831, PI: Professor Charalambos Antoniades) titled 'Using Radiomics and Artificial Intelligence to Predict Cardio-embolic Stroke' which provided the funding for the GPUs within AMIIC and the servers required to handle the CT scan data. This grant has served to provide the staff and plant necessary for this DPhil project to be realised.

3.3.3. Ethical considerations

The overall responsibility of the ORFAN Study ethics approvals lies with the Chief Investigator. Delegation of responsibilities for the study to ORFAN project team members has resulted in Ms Maria Lyasheva being responsible for the consented arms of the study (Arms 1-3) and myself being responsible for the non-consented arm (Arm 4).

The ORFAN Study as a whole has UK Research Ethics Committee (REC) Approval (Oxford Research Ethics Committee C reference 15/SC/0545). The ORFAN IRAS Project ID is 164361 and the project is listed on ClinicalTrials.gov with identifier NCT05169333.

The current version of the ORFAN Study protocol is V10.0 which was last amended and approved by the REC in July 2022 (see Appendix 2 for REC approval). Arm 4 of the ORFAN study has Health Research Authority Confidentiality Advisory Group (HRA CAG) approval (20/CAG/0157) for specific dataflows of non-consented data between ORFAN Study sites in England and Wales and third-party data controllers who perform data linkage as part of the study design to collect both baseline and outcome data for participants (see Appendix 2 for HRA CAG approval). The ORFAN study also undergoes local review at each local ORFAN participating site, with most sites requiring a Data Protection & Impact Assessment (DPIA) to be submitted to the local Research & Development

Department and/or the local Data Protection Officer/Caldicott Guardian for review. This is primarily for Arm 4 of the study.

International sites of the ORFAN Arm 4 study all undergo their own ethical approval within their own institutional or national system. Each site is supported through their application by the ORFAN Arm 4 team, namely Dr Henry West and Ms Lucy Kingham. The sites use the same ORFAN Protocol as is approved in the UK but seek approval only for the Arm 4 of the study, as they do not participate in the prospective consented arms of ORFAN. Many international sites also require a further ethical approval for the transfer of patient data from their home country to the UK for research purposes. These applications are managed locally with the support of the ORFAN team in the UK.

3.3.4. Site entry into ORFAN

The entry of prospective study sites into the ORFAN Study follows a simple pathway. This approach has been used for all eight of the original sites included in the first data collection, and all subsequent sites.

1. Informal discussions with the prospective local principal investigator (PI) and the ORFAN chief Investigator (CI) about possible joining of the study.
2. Teleconference between the Oxford based ORFAN team and the prospective ORFAN site team to formally discuss the feasibility of the ORFAN Study at their site, and for presentations on the ORFAN Study methods by the study leads (Dr Henry West and Ms Lucy Kingham).
3. Sharing of the ORFAN Study protocol, the local site database and data dictionary for local data collection input, technical requirements for the scan portal and the draft Research Contract for data sharing between the prospective site and the University of Oxford.
4. Formal discussions between the University of Oxford Research contracts team and the contract team at the local site.
5. Signing of a Research Contract between the University of Oxford and the local site data controller.
6. Minor amendment to the ORFAN Study Protocol to include the local site as a participating site within the Study Protocol, including listing the local PI as an ORFAN investigator.
7. Minor amendment to the HRA CAG approval to include the participating site within the CAG approval.
8. On-boarding teleconference with ORFAN investigators at Oxford with the local research team at the site, followed by the initiation of research activity at the site.

3.3.5. Patient and public involvement

The ORFAN Study has undergone two major rounds of patient and public involvement (PPI) activity. The first was at the founding of the study in 2015, when lead radiologist of the study Ms Sheena Thomas led patient focus groups to ascertain the acceptability of the prospective arms (arms 1-3) amongst patients referred to CT coronary angiography. Arm 4 was not within the study protocol at that point. This initial PPI included two focus groups with patients referred to CT coronary angiography and members of the public. Overall, there was positive response from these focus groups, and this PPI evidence of acceptability was included in initial applications for ethical approval for the Study.

The second, and more relevant, round of PPI occurred in late-2020 when the ORFAN study protocol was being amended to include Arm 4, the non-consented arm of the study. This round of PPI included a patient survey to evaluate patients' views on the specific dataflows proposed in the Study Protocol to enable the successful realisation of the Arm 4.

The survey was prepared by myself and approved by the Radcliffe Department of Medicine's Head of Clinical Research Operations and Governance, Ms Hanan Lamlum.

The survey was held in an ad-hoc fashion on weekdays when clinical research staff were available in November and December 2020. The patients approached to participate were a convenience sample of those referred to outpatient radiology services at both the John Radcliffe Hospital (OUH NHS Trust) and The Manor Hospital in Oxford. The patients who were asked to complete the survey were all NHS patients aged over 18-years old who were attending radiology for any CT scan. Patients were approached in person and asked if they were interested in completing the survey about a proposed University of Oxford research study. Those who were not interested were thanked for their time and not approached in any further way.

The survey asked three questions that each gave the respondent a scale of 0-10 to respond as to their agreement or disagreement with a certain statement about the proposed data flow within ORFAN. Question 4 was a free comment box with the respondent asked to provide any comments on the type of research described in this survey, including any concerns that they may have had.

The full text of the PPI survey documentation is included in the Appendix 3. The results of the quantitative questions were very positive (Table 3.1) with the majority of those surveyed having no issue with the proposed study methodology related to their data use.

The instruction for questions 1, 2 and 3 was as follows:

''For all questions, please use the scale of 0 to 10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement: …

Please indicate your response to questions 1, 2 and 3 with a mark at the appropriate level of agreement/disagreement''

Table 3.1. Patient and Public Involvement Survey Results for ORFAN Study Protocol Amendment to include Arm 4

Question	***Statement***	***Number of responses (of 25)***	***Range of responses***	***Mean response***	***Mode response***
1	*'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.*	25	7-10	9.12	10
2	*'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties organisations, such as NHS Digital, for the additional collection of relevant research information'.*	25	5-10	8.8	10
3	*'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual patient or contact the patient in any way.'*	25	5-10	8.92	10

The free-text response question elicited 13 comments from the 25 respondents. All comments are included below in transcribed forma and some scanned examples are attached in Appendix 3.

Most comments were very supportive of the ORFAN Arm 4 proposed data flow, with the only concerns raised being in relation to the security of the data, the data flow, and the concern of patients being contacted in an unsolicited manner. The average agreement with each question was very high, and one participant was in the 'neutral'. One respondent also expressed a wish for access to all their own data.

Example responses to the free text comments and concerns question are included here, with a focus on recurring themes:

'*Everything is fine, felt very comfortable'* (respondent 24)

'*I think that volunteers should be given access to their own biometric data, i.e. blood results and images'* (respondent 25)

'*It must be secure for patients' sake'* (respondent 23)

'*Happy to help!*' (respondent 15)

'*It makes sense. Thanks for your work'* (respondent 13)

'*As long as I/patients are not contacted*' (respondent 11)

'*No concerns'* (Respondent 10)

''*Happy to help'* (Respondent 9)

'*As long as it is secure'* (respondent 7)

''*We should support research* initiatives' (Respondent 6)

'*No concerns'* (Respondent 4)

'*Thank you. It must be safe to the patients'* (Respondent 2)

These results, and the very supportive comments provided reassurance that the population referred to public radiology services were supportive of the proposed data flow. Of the 13 who commented, the concerns raised were mostly general data protection concerns of relevance to all research. Possible concerns included unsolicited contact made by the researchers (which is not possible in Arm 4 as contact details are not collected for this arm), and the security of the data (security procedures are detailed at length in the ORFAN Study Protocol and supporting documents.). All such concerns were considered during the planning of the ORFAN Arm 4 methodology and implementation.

3.4. Methods

The design and methodology of the ORFAN Arm 4 Study is outlined in detail below.

3.4.1. Study design

The ORFAN Arm 4 Study is a multi-centre, international, prospective observational cohort study of clinical patients who undergo CCTA, with retrospective inclusion of eligible patients (CCTA from 2010 to date) and collection of data (2005 to date). The study will follow the cohort for outcomes until 2030 (maximum data collection of 25 years; 2005-2030).

The full ORFAN Study protocol is included in Appendix 4.

3.4.2. Patient populations

The patients eligible for the ORFAN Arm 4 Study at large are clinical patients aged between 18-99 years who have attended one of the participating ORFAN sites' CT departments for a CT angiogram including CCTA, or CT chest, or CT abdomen and pelvis scan, or post-mortem CT coronary angiography and are not participants of ORFAN study arms 1, 2 or 3.

This study population is very broad, and the initial focus of the ORFAN Arm 4 Study has been to include only patients who have undergone a clinical CCTA. This CCTA can be for any clinical indication.

Exclusion criteria for ORFAN Arm 4 are few. The only exclusion criteria are participation in ORFAN Study Arms 1, 2 or 3, and an active opt-out from use of data for research purposes. This is assessed via the NHS Digital National Data Opt-Out service.

3.4.3. Study hypothesis

The general ORFAN Study hypothesis is that biomarkers of metabolic risk could predict clinical outcomes in patients undergoing computed tomography angiography. Specifically, we hypothesise that:

1) Imaging biomarkers of body adiposity are associated with plasma biomarkers of metabolic risk.

2) Imaging/plasma/genetic biomarkers predict vascular disease severity (such as extent of coronary atherosclerosis, vascular calcification, aortic aneurysm size).

3) Imaging/plasma/genetic biomarkers predict future cardiovascular events and vascular disease progression in subjects undergoing computed tomography angiography.

3.4.4. Study objectives

The objectives of the ORFAN Study are outlined in Table 3.2. The overall objective of ORFAN for all arms is to investigate the development and validation of biomarkers for the prediction of cardiovascular disease progression or major events.

Most relevant to the ORFAN Arm 4 study is Secondary Objective 1, the objective concerning the identification of novel biomarkers able to predict cardiovascular disease pathogenesis and extent of vascular disease.

In practice this objective empowers the search and discovery of novel CT imaging biomarkers from patient scans.

Table 3.2. Formal study objectives of the ORFAN Study

	Objective	Outcome measures
Primary	To investigate whether biomarkers of metabolic risk can predict atherosclerosis disease progression or major adverse cardiovascular events	1. Measurements of plasma/imaging markers of cardiometabolic risk 2. Atherosclerosis progression by Computerised Tomography 3. Major adverse cardiovascular events over 10 years.
Secondary	Secondary objective 1: To identify novel biomarkers able to predict cardiovascular disease pathogenesis and extent of pre-existing vascular disease	1. Measurement of coronary plaque burden by Computerised Tomography. 2. Measurement of coronary calcification by Computerised Tomography 3. Measurement of aortic dimensions by Computerised Tomography 4. Measurement of changes in cardiac structure by Computerised Tomography 5. Measurement of novel indices of adipose tissue expansion and distribution, by using Computerised Tomography and/or ultrasound
	Secondary objective 2: Develop and validate imaging biomarkers to predict clinical outcomes in patients with COVID-19 infection.	1. Measurement of modification to cardiovascular disease pathogenesis and modulation of pre-existing vascular disease that is attributable to infection with coronavirus disease 2019 (COVID-19).

3.4.5. Participant selection and enrolment

Following the onboarding of ORFAN Study sites, each sites' research team collate their own patient list for inclusion in the ORFAN Arm 4 study. This list is managed by the local site PI, who is a senior clinician within either the cardiology or radiology department of the local site. This list will include all patients aged 18-99 years who have had a CCTA scan from 01/01/2010 through to the date the list is assembled.

The initial cohort will be 75,000 patients from across 8 NHS Trust sites in England.

Following expansion of the cohort in 2022, the cohort will be 200,000 patients from across England, and 50,000 internationally once fully realised.

Patients are not informed of their participation, as per the Study protocol and NHS Act 2006 Section 251 approval (HRA CAG approval).

4.5.6. Study site patient list & ORFAN identifiers

The Study sites are required to collate basic patient identifiers to enable subsequent data collection. These include patient NHS number, date of birth (DOB), postcode, local hospital medical record number (MRN) and the date of the CCTA scan.

Each study site is then required to assign each participant an ORFAN Study identifier. Each site is identified by a specific ID coding. Patients who have had multiple CCTA scans are only assigned a single ORFAN Study ID.

The format of the ORFAN Study patient identifiers is [Study code][Study Arm code][Site Code][Patient number]. The ORFAN Study code is F. The Arm code for ORFAN Arm 4 is 4. The site codes are outlined in Table 3.3. The patient code is a number between 00001 and 99999.

Each patient is assigned their unique ORFAN Study identifier in chronological order of CCTA scan at each site – the earliest CCTA scan will receive the first identifier.

Table 3.3. ORFAN Study Arm 4 patient identifier key

Site	Site code	Example patient identifier
Oxford University Hospital NHS Foundation Trust	OX	F4OX00012
University Hospitals of Leicester NHS Trust	LI	F4LI00012
Milton Keynes University Hospital NHS Foundation Trust	MK	F4MK00012
Royal Brompton and Harefield NHS Foundation Trust	RH	F4RH00012
Royal United Hospitals Bath NHS Foundation Trust	BA	F4BA00012
Royal Papworth Hospital NHS Trust	CA	F4CA00012
Guy's and St Thomas' NHS Foundation Trust	GT	F4GT00012
Barts Health NHS Trust	QM	F4QM00012
Leeds Teaching Hospitals NHS Foundation Trust	LE	F4LE00012

3.4.7. Study data flow outline

To best understand the pathway of patient data within the ORFAN Arm 4 Study, a big picture summary is informative (Figure 3.1).

Once an NHS site has formally joined the ORFAN Study, the site collates their own list of patient identifiers. They assign a unique ORFAN Study patient identifier to their own patients and then they dispatch a file containing the patient identifiers (NHS number and DOB) plus the assigned study ID to the third-party data holders of NHS Digital, The National Institute for Cardiovascular Outcomes Research (NICOR), and The Sentinel Stroke National Audit Programme (SSNAP). The third parties

then perform a National Data Opt-Out check and remove any patients who have opted out of non-consented use of their data in research. This check is also performed at the authorised third parties who have received the linkage file. Having received the linkage file from the NHS sites, the third parties then action the linkage with the patient identifiers within the databases that they control and dispatch the resulting data variables to the ORFAN team at Oxford. Meanwhile, the sites must complete a local clinical database (case record file – CRF) for each participant and dispatch this to the ORFAN team in Oxford. The sites also push each patient's CCTA scan(s) to the Oxford Core Lab via a specially installed portal that removed patient identifiers and inserts the assigned ORFAN ID.

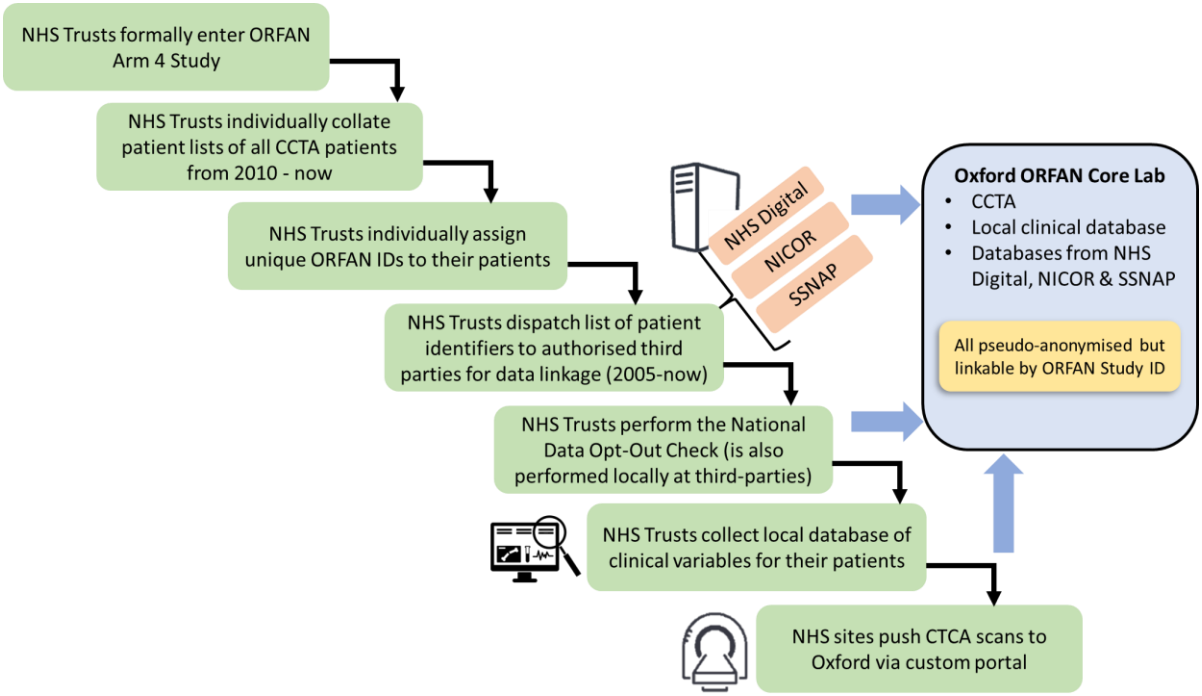


Figure 3.1. Overall logistical flow and data flow within ORFAN Arm 4 Study in England. *Once a Trust enters the study they collate their own list of patients eligible for inclusion (must have had a CCTA after 2010, inclusive) and assign each individual patient a unique ORFAN Study identifier. The sites then dispatch their list of patient identifiers (ORFAN identifier, NHS number and date of birth) to the authorised third parties of NHS Digital, NICOR and SSNAP for data linkage. Once linked the data from these third parties is sent to the ORFAN Core Lab at the University of Oxford. Simultaneously, the sites ae collecting their own local databases of clinical variables and pushing the CCTA through the custom Cimar portal to the Core Lab (see Figure 2.3 for specific of scan transfer). Once complete the ORFAN Core Lab will possess the CCTA scan, the local clinical database and the linked databases from NHS Digital, NICOR and SSNAP with patients all linked through their unique ORFAN Study ID. NHS: national health service; NICOR: National Institute for Cardiovascular Outcomes Research; SSNAP: The Sentinel Stroke National Audit Programme; CCTA: computed tomography coronary angiogram*

The result of this methodology is that the ORFAN team at the University of Oxford receives the CT scan and the relevant CRF for each patient from the sites themselves, and linked patient data from

the third parties, all linked by the unique patient identifier as originally assigned by the clinical teams at each ORFAN Study NHS site.

3.4.8. Study data collection

ORFAN Study Arm 4 data for England sites is collected from four sources:

1. The Local NHS Trust where the patient underwent the CCTA that renders them eligible for participation

    Including:

    a. The CCTA scan(s) and metadata
    b. A local CRF for each patient included at the site
2. NHS Digital
3. The National Institute for Cardiovascular Outcomes Research (NICOR)
4. The Sentinel Stroke National Audit Programme (SSNAP)

The data available from each source is summarised in Table 3.4.

The full variables received from each clinical site in the local CRF, and the variables received from NHS Digital, NICOR and SSNAP are included in Appendix 5.

Table 3.4. Sources of clinical data in ORFAN Study Arm 4

Source	Nature of data	Specific sub-databases
Local NHS Site CRF	Basic demographics, CCTA full request and report, clinical variables from time of scan, pathology results from time of scan and medications from time of scan	NA
NHS Digital	Full NHS records from 2005-now, including complete Hospital Episode Statistics data.	Hospital Episode Statistics: A&E, admitted patients, outpatients, critical care – all diagnoses, treatments & procedures Emergency Care Data Set Medicines Dispensed in Primary Care Civil registration (deaths) COVID-19 Hospitalisations in England Surveillance System
NICOR	Linkage with all cardiac events from the national audits 2005-present including all myocardial events, interventions such as percutaneous coronary intervention and coronary artery bypass graft surgery.	The National Adult Cardiac Surgery Audit (NACSA) dataset The Myocardial Ischaemia National Audit Project (MINAP) dataset The National Audit of PCI (NAPCI) dataset National Audit of Heart Failure (NAHF) dataset National Cardiac Rhythm Management (CRM) – electrophysiological studies (EPS) dataset National Cardiac Rhythm Management (CMR) – pacemaker and implantable cardiac device (PMICD) dataset The National Congenital Heart Disease Audit (NCHDA) dataset The Transcatheter Aortic Valve Implantation (TAVI) dataset
SSNAP	Linkage with all stroke events from the national stroke audit 2005-present	The SSNAP dataset

NICOR: The National Institute for Cardiovascular Outcomes Research; SSNAP: The Sentinel Stroke National Audit Programme

3.4.9. ORFAN sites and patient contribution in the UK

The initial ORFAN protocol allowed for 75,000 participants from the UK. These 75,000 participants were drawn from eight NHS Trusts (green and red in Figure 3.2). The remainder of sites will be included in the expanded cohort of 200,000 participants as approved in the ORFAN protocol amendment of 2022.



Figure 3.2. UK sites of the ORFAN Study Arm 4. *Study sites are represented in differing colours depending on their status. Green is Oxford, the sponsor of the ORFAN Study. Red represents the sites who have participated in the first data collection of 75,000 patients. Yellow represents the sites joining the ORFAN Study who will contribute to the forthcoming second collection of patient data to bring the total UK patients to 200,000.*

Table 3.5 provides the site profile for the UK ORFAN Arm 4 sites who have already contributed patient data into the Study. This includes the first 8 sites plus Leeds, who were able to upload their patient list for NHS Digital linkage, but have not actioned other aspects of the study as yet.

Table 3.5. Active ORFAN Arm 4 Study sites in England

Site	PI	Number of patients originally contributed	Patients remaining following NDOP	Outstanding patients to contribute to second wave	Local data collection	Scan transfer
Oxford University Hospital NHS Foundation Trust	Professor Charalambos Antoniades	5,766 (plus 10,000 non-contrast chest scans †)	5,199 (plus 10,000 non-contrast chest scans)	0*	Complete	Complete
University Hospitals of Leicester NHS Trust	Professor David Adlam	10,051	9,064	0*	Ongoing	Complete
Milton Keynes University Hospital NHS Foundation Trust	Professor Atilla Kardos	1,000	901	0*	Complete	Complete
Royal Brompton and Harefield NHS Foundation Trust	Dr Edward Nicol	16,096	14,514	3,963	Ongoing	Complete
Royal United Hospitals Bath NHS Foundation Trust	Dr Jonathan Rodrigues	3,906	3,523	0*	Complete	Complete
Royal Papworth Hospital NHS Trust	Dr Nicholas Screaton	9,609	8,665	0*	Complete	Complete
Guy's and St Thomas' NHS Foundation Trust	Dr Ronak Rajani	11,609	10,469	0*	Ongoing	Ongoing
Barts Health NHS Trust	Professor Steffen Petersen	13,000	11,722	22,520	Ongoing	Ongoing
Leeds Teaching Hospitals NHS Foundation Trust	Professor John Greenwood	3,729	To be performed	0*	Ongoing	Ongoing
Totals:		84,766	74,057	26,483		

*Zero indicates all patients from 2010 until the list was assembled in late 2021 were included and no patients were left off the list. All sites continue to accrue clinical patients who have undergone scans between the date the list was created and now. These patients will likely be included in the next collection of patient data.
†Already cleared NDOP
NDOP: National Data Opt-Out; PI: principal investigator

Sites yet to contribute patients into the ORFAN Arm 4 Study but who have signed contracts to join the study, or are in the process of, include:

- NHS Greater Glasgow and Clyde
- Royal Wolverhampton NHS Trust
- Sandwell & West Birmingham Hospitals NHS Trust
- University Hospitals Birmingham NHS Foundation Trust
- Manchester University NHS Trust
- University of Edinburgh
- University College London
- Wirral University Teaching Hospital NHS Foundation Trust

3.4.10. ORFAN Sites and Patient contribution from outside the UK

There are 6 international sites (outside UK) who have contributed data and scans into ORFAN Arm 4, with 8 sites actively collecting data to contribute. A further 13 sites are in the process of joining the study. Figure 3.3 displays all international sites. Table 3.6 lists all international sites by institution, with local principal investigators, and the current status and contribution of the site.

The full list of ORFAN Study investigators and affiliations is found in Appendix 6.

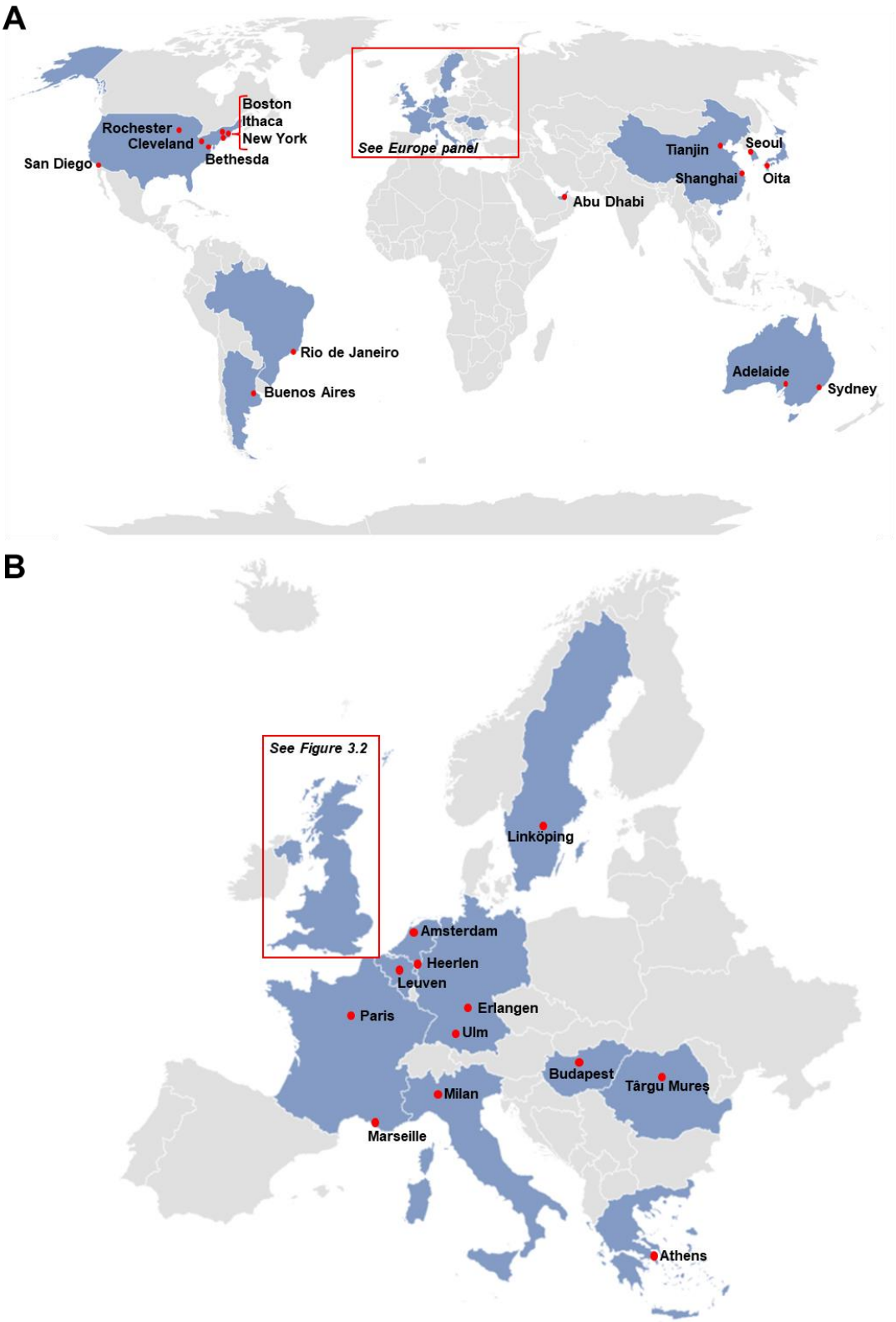


Figure 3.3. International sites of the ORFAN Arm 4 Study. *There are 28 sites that are either already contributing data into the ORFAN Arm 4 Study or are in the process of joining. Panel A shows international sites from outside Europe while panel B shows European sites.*

Table 3.6. International sites of the ORFAN Arm 4 Study

Institute and city*	Country	Local PI	Status†	Patient number‡
Asia				
Fudan University, Fudan	China	Neng Dai	Contractual	NA
TEDA International Cardiovascular Hospital, Tianjin	China	Guo-Wei He	Contractual	NA
Oita University, Oita	Japan	Naohiko Takahashi	Contributing	NA
Seoul National University, Seoul	South Korea	Bon-Kwon Koo	Data submitted	643
Europe				
Universitaire Ziekenhuizen Leuven, Leuven	Belgium	Ruben De Bosscher	Contributing	1200
Assistance Publique–Hôpitaux de Paris, Paris	France	Stefan Hatem	Contributing	20000
Cardiovista, Marseille	France	Alain Tavildari	Contributing	1000
Erlangen University Hospital, Erlangen	Germany	Mohamed Marwan	Data submitted	1993
Universitätsklinikum Ulm, Ulm	Germany	Meinrad Beer	Contractual	NA
Lefkos Stavros - The Athens Clinic, Athens	Greece	Georgios Benetos	Contributing	1000
University of Athens, Athens	Greece	Dimitrios Tousoulis	Contributing	2448
Semmelweis University, Budapest	Hungary	Pal Maurovich-Horvat	Data submitted	2191
University of Milan, Milan	Italy	Gianluca Pontone	Contributing	1000
Vrije Universiteit Medical Centre, Amsterdam	Netherlands	Ibrahim Danad	Contractual	NA
Zuyderland Medical Centre, Heerlen	Netherlands	Bas Kietselaer	Contractual	NA
University of Medicine & Pharmacy, Târgu Mureș	Romania	Theodora Benedek	Contributing	1000
Linköping University Hospital, Linköping	Sweden	Emmanouil Charitakis	Data submitted	192
Middle East				
Cleveland Clinic, Abu Dhabi	UAE	Ronney Shantouf	Contractual	NA
North America				
Cleveland Clinic, Cleveland	USA	Milind Desai	Data submitted	2246
Cornell University, Ithaca	USA	Leslee Shaw	Contractual	2000
Harvard University, Cambridge	USA	Ron Blankstein	Contractual	3000
Mayo Clinic, Rochester	USA	Bernard Gersh	Contractual	NA
Mount Sinai Hospital, New York	USA	George Dangas	Contractual	NA
National Institutes of Health, Bethesda	USA	Nehal Mehta	Data submitted	731
Scripps Research, San Diego	USA	Jorge Gonzales	Contractual	NA
South America				
Universidad Católica Argentina, Buenos Aires	Argentina	Gastson Rodriguez	Contractual	NA
Serfaty Clinicas Integradas, Rio de Janeiro	Brazil	Fabiano Serfaty	Contractual	NA
Oceania				
Flinders University, Adelaide	Australia	Joseph Selvanayagam	Contractual	2000
University of Sydney, Sydney	Australia	Gemma Figtree	Contractual	5000

*Sites listed in alphabetical order by continent, then country, then institute
†As at September 2022
‡ If known; where sites have status as 'Data submitted' patient number is confirmed. If Status is 'Contributing' or 'Contractual' patient number is an estimate or is yet to be determined.
NA: Not available; PI: principal investigator

3.4.11. Definitions of clinical risk factors and outcomes

In the ORFAN Arm 4, hypertension was defined based on the presence of a documented diagnosis or treatment with an antihypertensive regimen, according to the relevant clinical guidelines[166]. Similar criteria were applied for the definition of hypercholesterolemia and diabetes mellitus[167,168]. Valve disease was defined as the presence of any documented aortic or mitral valve stenosis or

regurgitation (of any severity), or previous valve repair and/or replacement procedure, including minimal invasive methods. Previous heart surgery was defined as any previous documented coronary artery bypass graft procedure or valve replacement or repair procedure (not including TAVI). Active smoking was defined as any documented or patient reported active smoking of any amount of tobacco products, including all tobacco products. AF was defined as documented AF on ECG, patient reported diagnosis, or being treated on anti-coagulant therapy for AF. All SCOT-HEART baseline variables have been described in further detail elsewhere, including in chapter 2 of this thesis[148].

3.4.12. Safety outcomes

Incidental findings

CCTA scans received at the ORFAN Core Lab as part of the ORFAN Arm 4 Study are all clinically indicated scans that have been requested and reported by specialist physicians at each participating site in the past (2010 to date). No formal clinical analysis, such as coronary stenosis and plaque assessment, is undertaken by the ORFAN Study team for express diagnostic or patient management purposes. This has already been performed by the local reporting and patient care team.

For this reason, incidental findings on the scan are presumed to have been managed internally by the patient's appropriately qualified referring physician or a member of their team at the time of scan acquisition.

Radiation dose

All CT scans raise concerns relating to exposure to ionizing radiation, however given that this study is not initiating any CT scans as part of the research, merely collecting previously acquired scans as part of clinical care, this is not a consideration for the Study.

3.4.13. Study monitoring

The ORFAN Study steering committee includes representatives from the grant applicants and trial management, as well as individuals not directly involved in the trial. A Data and Publications Committee to decide scientific priorities and publication rights will be convened. The study is observational in nature so a data and safety monitoring committee has not be convened.

3.5. Discussion

The ORFAN Study, once complete, will be the largest clinical CCTA cohort available for research purposes, with 250,000 patients included in Arm 4 from across the world, and 15,000 UK extensively phenotyped patients included in Arms 1-3. For the purpose of this thesis the study is only utilising CCTA, however the Study is not limited to CCTA. The protocol and ethical approvals allow the

collection of any CT data that patient's may have had for clinical purposes, including scans of the chest, abdomen, and pelvis. The possible future research that may occur within ORFAN expands beyond the current study aims and could include any research question relevant to the use of CT in clinical practice.

3.5.1. Context of the ORFAN Study

As elucidated in Chapter 1 of this thesis, cardiovascular disease remains the leading cause of morbidity and mortality despite major advances in risk assessment, prevention, and treatment[169]. Excess adipose tissue, manifesting as overweight and obesity, is a key risk factor for CVD including CAD, AF, and stroke. It is now well recognised that adipose tissue is a critical regulator of cardiovascular health, and exerts both protective and deleterious effects on the cardiovascular system[120]. Current understanding of the exact mechanisms and processes of the inter-play between adipose tissue and the cardiovascular system is lacking due to the varied pathobiology likely to be at play across different adipose tissue depots and tissues of the cardiovascular system[61]. For example, how peri-coronary artery adipose tissue impacts the clinical risk of CAD is likely mediated through different processes than how more remote adipose tissue deports, such as mediastinal adipose tissue or subcutaneous adipose tissue, impact such risk.

Due to its excellent spatial resolution and low cost, computed tomography assessment of the heart via CCTA has become the gold-standard clinical test for patients with chest pain[7,170]. Because of this, the improved analysis of CCTA is highly relevant for improved scientific understanding of the development of CAD, the assessment of CAD patient risk, and the improved management of patients with chest pain[120]. CCTA has therefore become the initial focus of attention within the ORFAN Arm 4 Study. As is outlined in Chapter 2, use of CT imaging enables the accurate and reproducible quantification of adipose tissue volumetric and textural characteristics in all depots by applying standardised CT attenuation thresholds and segmentation algorithms[140,171]. This unlocks the potential for the use of CT scans for the assessment of risk of conditions typically not considered with traditional CT reporting. Others have previously created clinical CCTA registries[146] and large cohorts for trials[147,148], but not on the scale or with the broad academic and technological aims that are driving the ORFAN Study.

3.5.2. Evidence and technology gap filled by ORFAN

The ORFAN Study was established to develop and utilise advanced data analysis techniques to close the scientific evidence gaps that exist in the relationships between adipose tissue and the cardiovascular system. This is achieved through the formation of longitudinal cohorts that will include CT imaging, blood samples, basic clinical data, and clinical outcomes data over multiple

decades. The key evidence gaps that will be addressed is an improved understanding of the role of adipose tissue imaging for the detection of dysfunctional adipose tissue. Adipose tissue in the clinical setting is limited to measurement of anthropometric indices, predominantly body mass index (BMI) or waist-to-hip ratio or waist-to-height ratio. Unfortunately, such simple indices do not capture important adipose tissue features, such as visceral distribution, or biology, such as the specific functional nature of the adipocytes and the implications for future disease risk for conditions such as AF and stroke. Initial disease targets, and those included in this thesis, within the ORFAN Study include the development of personalised risk assessment for conditions such as CAD events, AF risk, and stroke risk.

3.5.3. Strengths of the ORFAN Study

The strengths of The ORFAN study are myriad. A diverse cohort of CT patients with well described baseline characteristics from the time of the CT scans and linked clinical outcomes data from nationally collated data networks and registries has been created. The sample size is adequate for the implementation of artificial intelligence techniques such as deep-learning for CT analysis and both supervised and unsupervised machine learning approaches to data analysis of both clinical and radiomic data. The potential for the development of clinical practice changing technological advances in CT analysis, primarily for enhanced patient risk assessment, is significant.

3.5.4. Limitations of the ORFAN Study

Current limitations of the ORFAN Study include a lack of primary care data across all Arms. This is being addressed with pending applications for UK CPRD access for primary care data. Arm 1, 2, and 3 may be impacted by authorisation bias due to the nature of approaching patients for participation in research. This is alleviated somewhat by allowing patients to consent to images and data only, without having to consent to give a blood sample. The lack of international sites in ORFAN Arm 4 that are in less or least developed nations is a weakness, however, we are actively looking for collaborating sites in such regions.

3.6. Conclusion

The ORFAN Arm 4 Study will be the largest CCTA cohort available for researchers across the world. The cohort currently consists of some 88,000 patients with complete or near complete data for analysis, with approvals for expansion to 250,000 patients in place. The Study will be utilised to develop novel imaging technologies such as those within this thesis, and beyond.

3.7. Personal contribution

My contribution into the ORFAN Project is as follows:

- Drafting of the ORFAN Arm 4 protocol amendments to establish the study within the original ORFAN Study protocol
- Analysing pilot data and drafting the major successful grant application (BHS Translational Grant) to fund the ORFAN Study via BHF Translational Grant
- Writing the HRA CAG application for submission for NHS Act 2006 Section 251 support for the Study
- Contributing to the REC committee application for amendment of the study protocol on three occasions
- Planning and conducting the successful PPI initiative to assess suitability of the ORFAN Arm 4 data flow to patients within the NHS for the purposes of gaining study approvals
- On boarding over 12 new ORFAN participating sites from across the UK
- On boarding over six new ORFAN participating sites from outside the UK
- Managing the local clinical teams to support them in the creation and uploading of ORFAN Arm 4 patient lists to NHS Digital, NICOR and SSNAP at all ORFAN Study sites
- Managing the local clinical teams to support them to upload their Arm 4 patient lists to the NHS Digital National data Opt-Out check
- Managing the local clinical teams to support them to complete the CRF for each patient within ORFAN Arm 4 at their sites
- Contributing to the installation and management of CT scan portals at all UK ORFAN sites alongside Dr Yogesh Sohan
- Recruiting patients with ORFAN Arm 2 and 3 at the John Radcliffe Hospital, Oxford, UK. Including consenting, completing CRFs, taking blood samples and processing blood samples in the laboratory

4

Chapter 4. Automated segmentation of CCTA scans

Parts of this chapter are drawn from the publication West HW et al. 'Deep-Learning for Epicardial Adipose Tissue Assessment with Computed Tomography: Implications for Cardiovascular Risk Prediction. Journal of the American College of Cardiology: Cardiovascular Imaging. Accepted November 17 2022.

4.1. Introduction

This chapter examines the development of the novel CT imaging tools for automated segmentation of multiple discrete regions of the heart, specifically the whole heart within the pericardium, the left atrium (LA) and three segments surrounding the LA (peri-LA segments). The chapter includes an outline of the patient populations used for the development and testing of the novel segmentation tools and methods used to explore the diagnostic and prognostic value of the novel tools in clinical patients. The chapter also includes a description of the workflow in the Oxford Academic Cardiovascular Computed Tomography (OXACCT) Core Laboratory and the standard operating procedures (SOPs) that have been designed to govern this process.

To extract radiomic variables from CT images, segmentations of the area of interest must be performed. To automate segmentation via a deep-learning approach, many manual segmentations must be performed to train a deep-learning network. Any machine output must be compared to human expert 'ground-truth' to evaluate its performance. This process is described in detail within this chapter.

The scientific aims of the thesis that are achieved in this chapter include the automation of the quantification of EAT volume on clinical CCTA, and the automation of the segmentation of the LA and peri-LA tissues on clinical CCTA.

4.2. Methods

4.2.1. Summary of process for technical development

The overall development pathway for the automated segmentation tools on CCTA is summarised in Figure 4.1. Two parallel processes run simultaneously. The first was for the development of automated whole heart segmentation within the pericardium (green in Figure 4.1) and the second was the development of the automated left atrium (LA) and peri-LA segmentation (orange in Figure 4.1). The process of development was the same for both segmentation programs bar the initial steps necessary for the peri-LA segmentations. This is because scientific decisions based upon previous

literature were required to ensure the most meaningful segments were developed for the purposes of later analysis, followed by a period of creative experimentation with manual segmentation and testing for reproducibility.

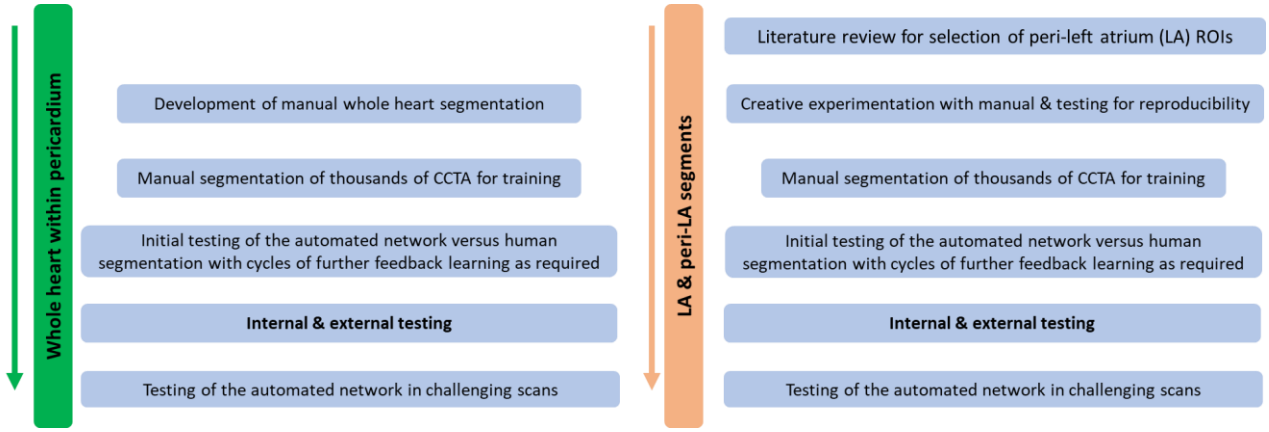


Figure 4.1. Summary of development steps for the automated segmentation tools. *Two parallel processes were actioned at the same time, the development of the whole heart segmentation (green) and the development of the LA and peri-LA regions segmentation (orange). The developmental steps were the same for each process bar the literature review and experimentation required for the initial development of the LA and peri-LA segmentations. CCTA: Computed tomography coronary angiography; LA: left atrium; ROI: region of interest*

4.2.2. Manual segmentation of CCTA

Manual segmentation was required for the whole heart within the bounds of the pericardium, the LA and the 3 peri-LA tissue regions of interest (ROI).

The goal was automated segmentation of the most biologically relevant ROI on CCTA for extraction of the radiomic features for AF & stroke risk prediction purposes.

The anatomical regions of the heart most relevant for AF and stroke risk were determined from the literature and translated into CCTA segmentations that capture the relevant tissues. These were:

- The wall of the left atrial appendage (LAA), also taking in the junctions of the left pulmonary veins
- The anterior LA wall
- Posterior LA wall (including junctions of the right pulmonary veins)
- The epicardial adipose tissue & connective tissues adjacent to the three the above regions.

These regions became the focus of the manual segmentation work to generate reproducible segmentations both internally to each reader, but also externally between readers.

The other segments we set out to automated were:

- Whole chamber of the LA

- Whole heart within the pericardium

4.2.3. Experimentation of peri-LA segmentation

A period of creative experimentation was necessary, driven by a need for reproducibility of all segmentations that would be carried forward into automated segmentations.

The free open-source image computing software package 3D Slicer[142] was utilised for all creative segmentation development. Dozens of different approaches to manual segmentation of the ROI were tested.

The process followed for this experimentation was as follows:

1. Perform potential manual segmentation individually and assess suitability.
2. Repeat segmentation on same scan and compare visually and with first order statistics (volume, mean attenuation, dimensions)
3. If acceptable, complete series of cases for comparison of intra-reader variability
4. If acceptable, complete series of cases for comparison of inter-reader variability between two human readers.

Multiple experiment segmentations were tested. These included segmentations of all the atrial wall and also defined regions of the wall and surrounding tissues. Each segment was performed for all voxels (no Hounsfield unit (HU) threshold applied), and typically would be a segment of the LA wall and the surrounding tissues immediately adjacent to the wall. The segments were never of tissue outside the pericardium.

Table 4.1 demonstrates various manual ROI segments that were deemed acceptable and carried forward to intra-observer testing, and those carried forward to initial inter-observer reproducibility testing.

Table 4.1. Manual CCTA segmentations for development of automated segmentation methods

Segment	Intra-observer CCC (n=20)	Inter-observer CCC (n=20)	Comments
Whole left atrial (LA) wall	0.76	0.74	Very time consuming (40min +)
Posterior LA wall	0.79	0.77	
Anterior LA wall	0.94	0.95	
Left lateral/posterior LA wall	0.79	Not performed	
Left atrial appendage wall	0.94	0.94	Time consuming (15min +)
Intra-atrial wall	0.97	0.96	
Pulmonary vein junctions	0.81	0.71	Technically challenging
LA chamber	0.99	0.98	Easy with HU thresholding

CCC: concordance correlation coefficient; HU: Hounsfield Unit; LA: left atrium

4.2.4. Selection of peri-LA segments

Ultimately a balance between difficulty of manual segmentation and scientific utility had to be struck for the manual segments carried forward into the manual training cohort to train the deep-learning model. The segments that were selected based upon their intra- and inter-reader correlation coefficients, and the time required for manual segmentation were:

1. The intra-atrial wall
2. The anterior LA wall
3. The left atrial appendage wall

The LA chamber itself was also carried forwards into en masse manual segmentation.

Figure 4.2 displays the segments carried forward into en masse segmentation for the purposes of training the automated segmentation DLN. The large-scale segmentation occurred on 3DSlicer[142] (version 4.11.0), as demonstrated in Figure 4.3.

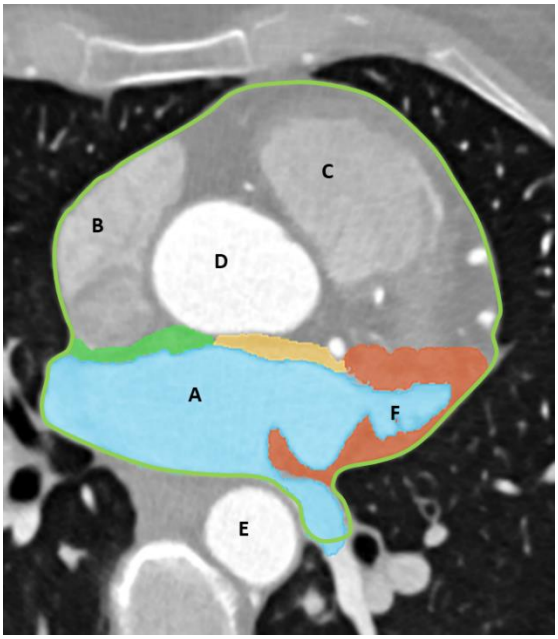


***Figure 4.2. CCTA segmentations selected for automation**. Single slice of a CCTA demonstrating the manual segmentations selected to carry into the full cohort for training of automated segmentation. Segmentation of all volumes is of all voxels regardless of HU density. Blue: Whole left atrium chamber; Green: Intra-atrial wall; Yellow: Anterior left atrium wall; Red: Left atrial appendage wall; Green line: Pericardium for whole heart within the pericardium A= Left atrium, B= Right atrium, C= Left ventricle, D= Aorta, E= Descending aorta, F= Left atrial appendage*

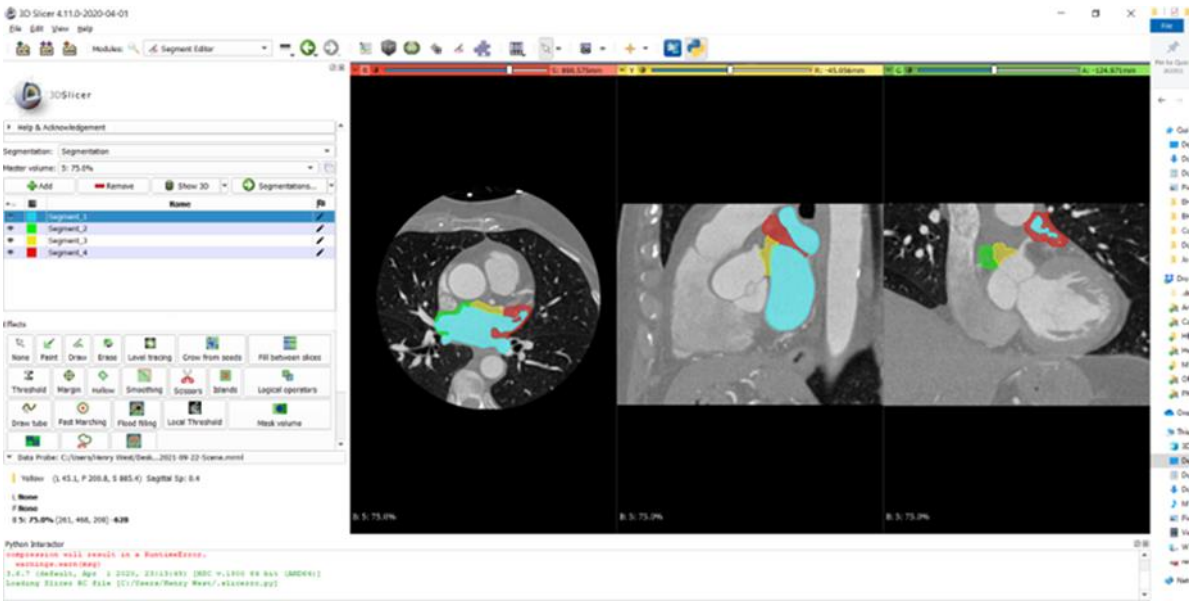


***Figure 4.3. Screenshot of 3DSlicer work package demonstrating a CCTA scan with manual segmentation on screen.** The screenshot from 3DSlicer (version 4.11.0) demonstrates the left atrium segment (blue), the inter-atrial wall segment (green), the left atrium anterior wall segment (yellow) and the left atrial appendage wall segment (red). The three traditional CT views are provided to give a scrollable and tethered 3D view of the segments for editing.*

4.2.5. Whole heart manual segmentation

The whole heart within the pericardium was also required to be manually segmented on all scans to train the automated whole heart segmentation model. The whole heart was manually segmented between the bifurcation of the pulmonary trunk superiorly and the apex of the pericardium inferiorly, as was already the SOP within the OXACT Core Lab.

As for the LA and peri-LA segments, 3DSlicer[142] (version 4.11.0) was used for all manual segmentation of the whole heart. Figure 4.4 demonstrates the workview of a CCTA manually segmented on 3DSlicer. All voxels within the pericardium including the visceral pericardium were included.

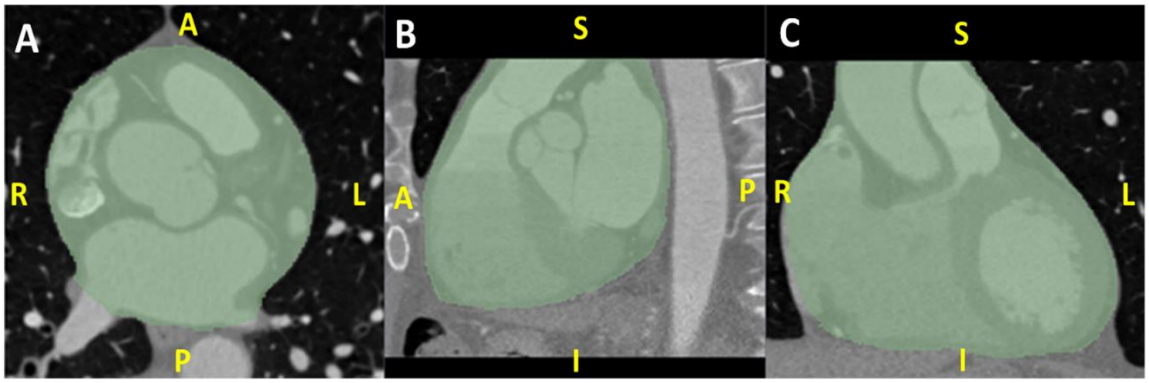


Figure 4.4. Manual segmentation of the whole heart. *The segmentation of the whole heart within the pericardium was initially undertaken on 3DSlicer. This view shows the traditional three CT views with a manually segmented whole heart (light green).*

Segmentation was shifted to CaRi-Heart® version 2.2.1 (Caristo Diagnostics Ltd., Oxford, UK) when this software became available within our research group. Figure 4.5 demonstrates the work view on CaRi-Heart for an automatically segmented whole heart volume.

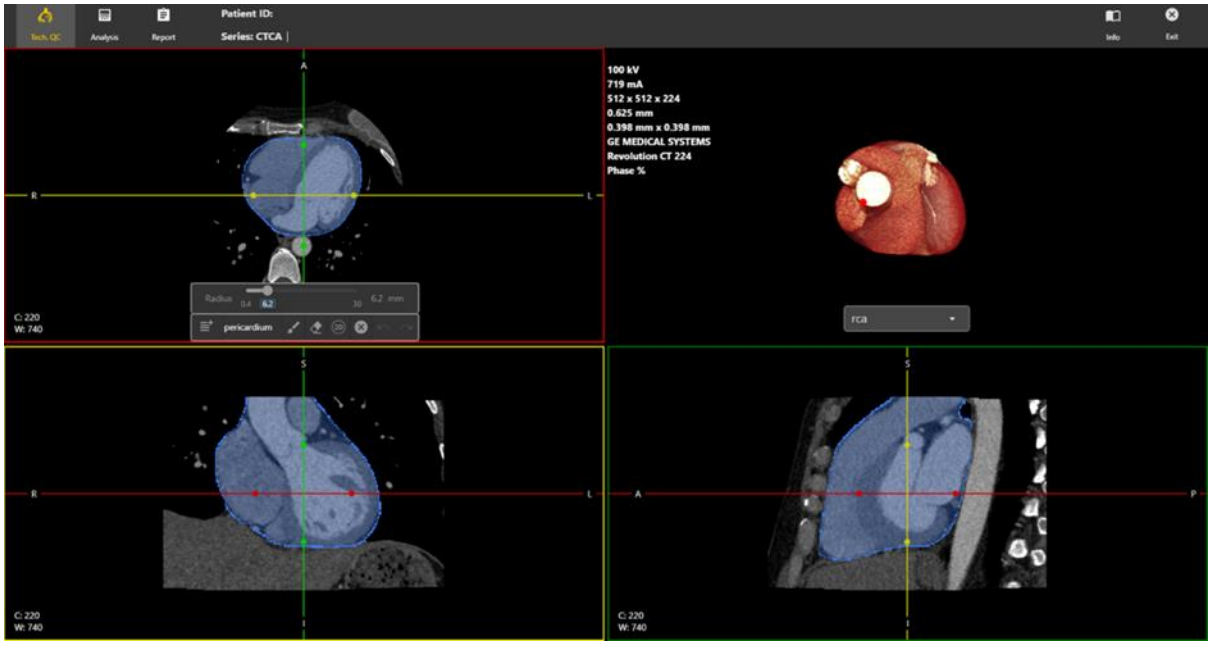


Figure 4.5. Image of CaRi-Heart user work view of manual whole heart segmentation. *The whole heart within the bounds of the visceral pericardium is manually segmented incorporating all voxels inferior to the bifurcation of the pulmonary trunk and superior to the most inferior portion of the apex of the heart. Image shows axial (red), coronal (yellow) and sagittal (green) views. Blue = segmented region.* ***Reproduced with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.***

4.2.6. The extraction of EAT volume from the heart segmentation

The extraction of EAT segments from the segmentation of the whole heart was achieved by automated selection of adipose tissue voxels. The HU range of -190 HU through to -30 HU was utilised, adhering to the conventional definition of adipose tissue density on CCTA[162,172,173]. The extraction of the EAT volume on a per-voxel basis from within the whole heart segment was fully automated in a bespoke module within the CaRi-Heart® platform, created by Dr Musib Siddique and me especially for the purpose of the quantification of EAT from CCTA heart segmentations. This was a streamlined version of the *RadExtractor* program developed by the same researchers (see section 4.1.9 below). An example of EAT segmentation is shown in Figure 4.6, with manual segmentations and automated segmentation shown independently, and merged.

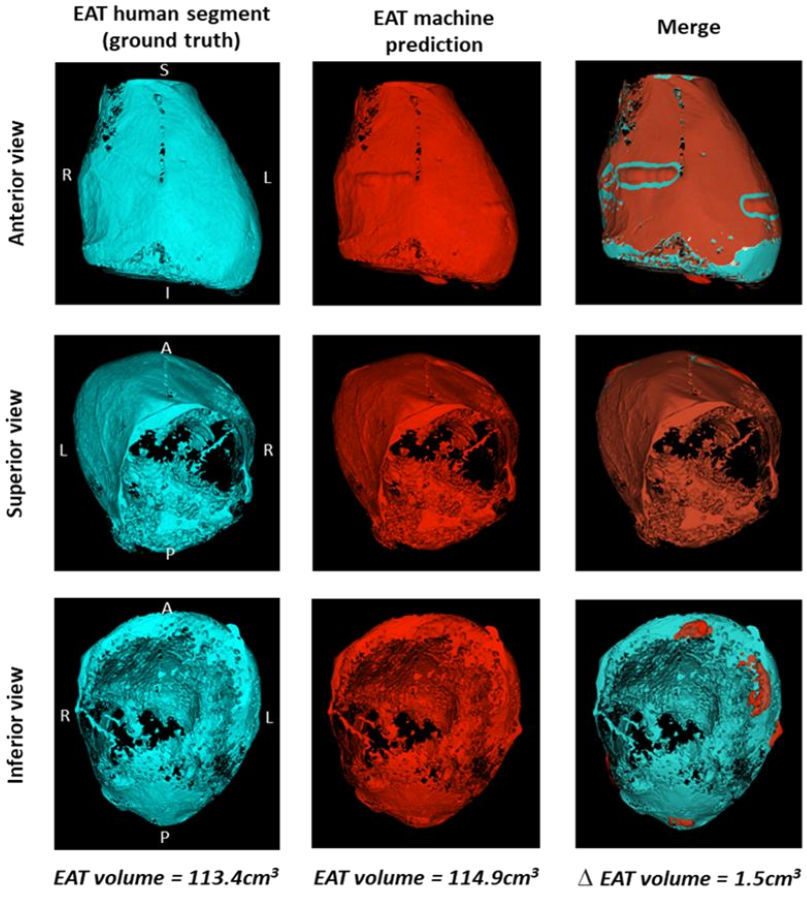


Figure 4.6. Automated segmentation of epicardial adipose tissue extracted from a whole heart segmentation. *A single CCTA from the ORFAN study demonstrating human expert segmentation as ground truth, automated machine segmentation and a merge.* ***Reproduced with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.***

4.2.7. Intra- and inter-observer variability testing for whole heart segments and epicardial adipose tissue volumes

Examination of EAT volume segments extracted from whole heart segments for intra-observer bias was excellent with both expert human readers having excellent intra-class correlation coefficient in a test of 50 blinded repeat reads of the same case (reader 1, ICC[95%CI] = 0.987[0.980-0.999] $p < 0.001$; reader 2, 0.984[0.976-0.998] $p < 0.001$; interval between reads of two weeks).

Inter-observer bias was assessed between two readers (myself and Ms Lucrezia Volpe). Both readers were blind to all patient scan characteristics at the time of their whole heart segmentation. Both readers were also blind to the result of the other reader at all times. 100 ORFAN Arm 4 CCTA cases were randomly selected for the analysis and both readers conducted their segmentation independently.

For whole heart segmentation volume the concordance correlation coefficient (CCC) was excellent between both readers (CCC=0.969) and there was non-significant bias on Bland-Altman analysis (p=0.88) (Figure 4.7).

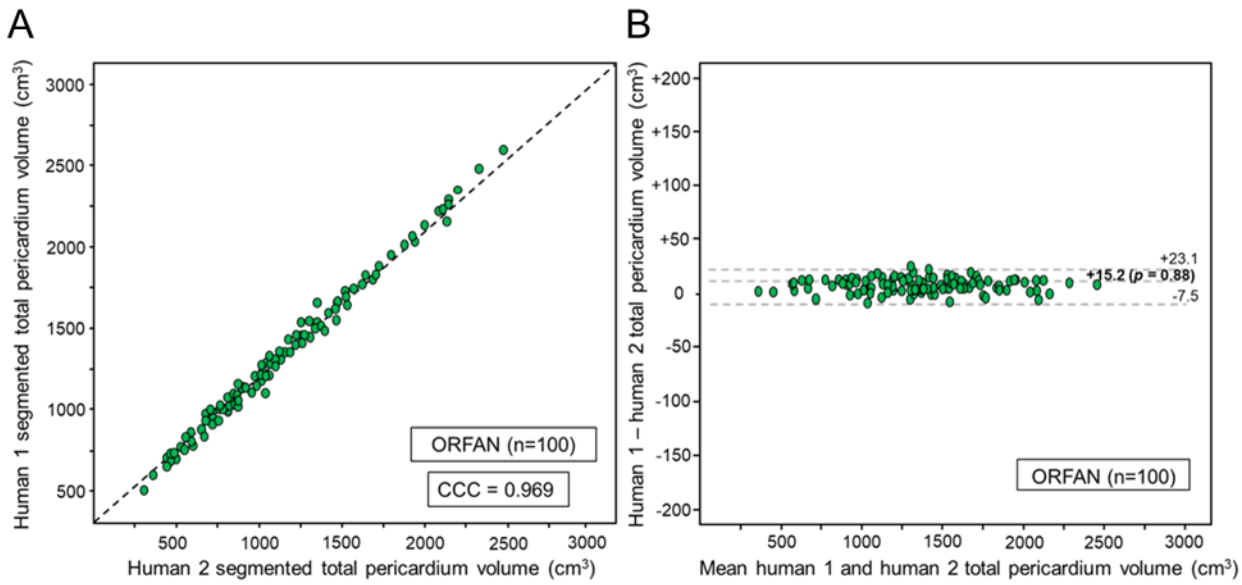


Figure 4.7. Inter-reader variability testing of the whole heart volume segmentation. Whole heart segmentation was tested between two independent and blinded human readers, with scatterplot analysis and concordance correlation coefficient (CCC) (A), and Bland-Altman analysis (B).

For the EAT volumes extracted from the same whole heart segments as already independently quantified by each reader, the CCC was also excellent at 0.970. Again, the bias on Bland-Altman Analysis was non-significant (p=0.74) (Figure 4.8).

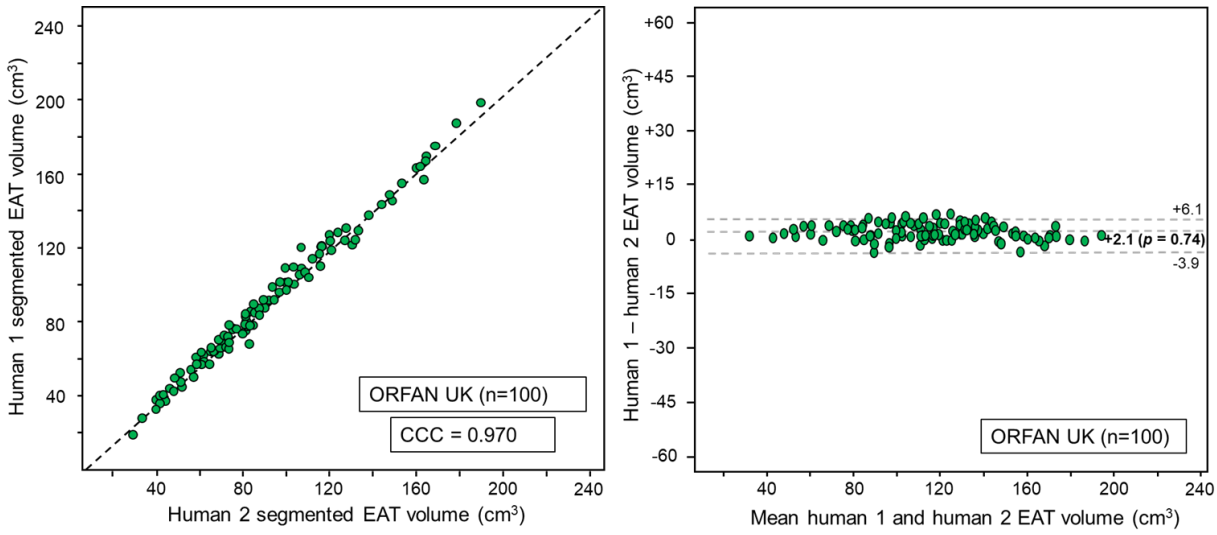


Figure 4.8. Inter-reader variability testing of the EAT volume segmentation. From within the whole heart segmentation, the EAT was extracted and its volume was tested between two independent and blinded human readers, with scatterplot analysis and concordance correlation coefficient (CCC) (A), and Bland-Altman analysis (B).

4.2.8. Inter-reader variability testing for left atrium and peri-left atrium segments

Inter-observer bias for the LA and peri-LA segmentations was assessed between the same two readers (myself and Ms Lucrezia Volpe) as the whole heart and EAT volumes. Both readers were blind to all patient scan characteristics at the time of their segmentations. Both readers were also blind to the results of the other reader at all times. 100 ORFAN Arm 4 CCTA cases were randomly selected for the analysis and both readers conducted their segmentation independently.

For the left atrium segment the CCC was excellent between both readers (CCC=0.978) and there was non-significant bias on Bland-Altman analysis (not-shown) (Figure 4.9A). For the intra-atrial wall segment the CCC was also excellent between both readers (CCC=0.962) and there was non-significant bias on Bland-Altman analysis (not-shown) (Figure 4.9B). For the anterior LA wall segment the CCC was excellent between both readers (CCC=0.958) and there was non-significant bias on Bland-Altman analysis (not-shown) (Figure 4.9C) and finally, similar results were found for the left atrial appendage segment with a CCC of 0.953 and non-significant bias on the Bland-Altman plot (not shown) (Figure 4.9D).

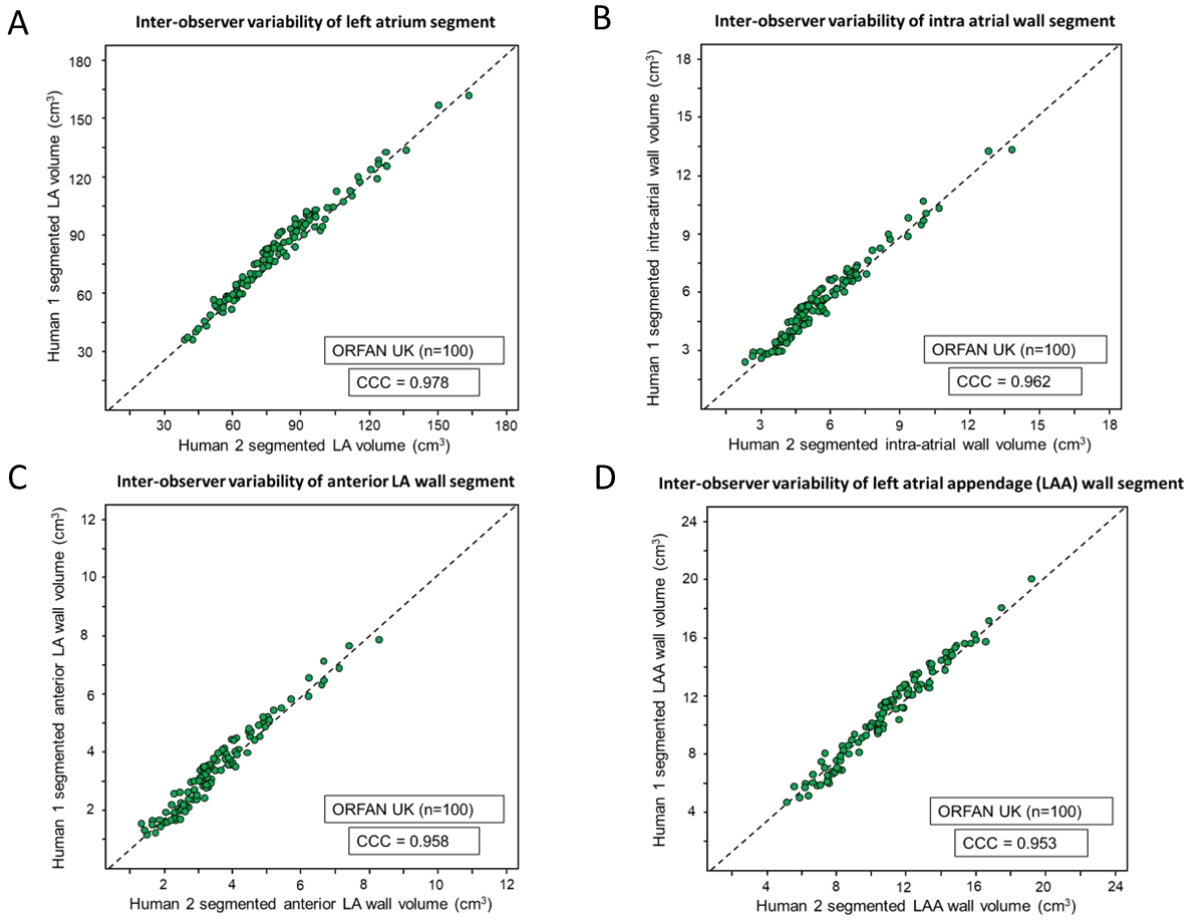


Figure 4.9. Inter-reader variability testing of the left atrium and peri-left atrium segmentations. Volumes for all manual segments including the LA (A), intra-atrial wall segment (B), anterior LA wall segment (C) and LAA wall segment (D) volume was tested between two independent and blinded human readers, with scatterplot analysis and CCC analysis presented for all segments. LA: left atrium, LAA: left atrial appendage; CCC: concordance correlation coefficient.

4.2.9. Formalisation of manual segmentation methods

To ensure clarity and reproducibility of the segmentation techniques developed for this project, I wrote two specific OXACCT Core Lab standard operating procedures (SOP) for the manual segmentation of a) the whole heart within the pericardium and b) the LA and peri-LA depots. Both SOPs were reviewed and approved internally by the Core Lab Director and Chief Radiographer.

To assess improvement in efficiency for quantification of EAT volume manual segmentation time was taken from the reported average of both human expert readers over 200 consecutive segmented cases each.

4.2.10. The extraction of radiomic features from segments

As outlined in chapter 2.8, the extraction of radiomic features from CCTA segments required an automated approach to avoid the manual work of extracting and saving the features segment by segment. The RadExtractor was utilised for all segmentations.

Given that our manual segmentation approach provided four independent ROI, the use of the RadExtractor facilitates the extraction of multiple unique sets of radiomic variables, as summarised in Table 4.2. Each independent ROI had radiomic features extracted (radiomic database 1-4), then there were two merged databases (radiomic database 5 and 6) of anatomically adjacent ROIs.

EAT specific databases were also created using the RadExtractor to limit the voxels from which the features were extracted to the -190 - +30 HU range. Only the peri-LA segments (segment 1-3) were relevant to this analysis.

Overall, this results in nine independent radiomic feature databases for analysis purposes.

Table 4.2. Radiomic feature databases for segmented CCTA regions of interest

Segment no.	***Segment 1***	***Segment 2***	***Segment 3***	***Segment 4***
Anatomical area	***Intra-atrial region***	***Anterior LA wall and surrounds***	***LAA wall and surrounds***	***LA chamber***
Radiomic database 1	X			
Radiomic database 2		X		
Radiomic database 3			X	
Radiomic database 4				X
Radiomic database 5	X			
Radiomic database 6		X		
EAT specific databases also extracted (x3)	X	X	X	

EAT: epicardial adipose tissue; LA: left atrium; LAA: Left atrial appendage

4.2.11. Study populations for technology development and validation

The ORFAN Arm 4 Study was utilised for the manual development and the validation of the deep-learning models for both the whole heart within the pericardium and the LA and peri-LA segmentations.

Table 4.3 provides the basic demographics of the patients utilised for each stage of the model development.

Initially, a random sample of 200 UK ORFAN patients were selected and sequestered from the population to serve as the internal validation cohort. A separate group of 100 were also sequestered at random to serve as a preliminary validation cohort to check progress of the model's accuracy compared to human segmentation prior to carrying out any cycles of direct feedback learning.

The 2200 patients for manual training were randomly selected from across the available ORFAN Arm 4 patient cohorts, at the time these were Oxford University Hospitals NHS Foundation Trust, Royal United Hospitals Bath NHS Foundation Trust, University Hospitals of Leicester NHS Trust, and Milton Keynes University Hospital NHS Foundation Trust.

The cohort for external validation was entirely separate, and came from the Cleveland Clinic, Ohio, USA, an international site of the ORFAN Arm 4 Study.

Given the cohorts are collated within the ORFAN Arm 4 study the patient populations are generally uniform in rates of risk factor exposure. The cohorts come from a variety of centres and so were scanned on a diversity of CT scanners, with no uniform enforcement of scan acquisition settings, as reported in Chapter 2.

Table 4.3 ORFAN Arm 4 study populations used for deep-learning model development and validation

		ORFAN *manual training cohort*	**ORFAN UK *feedback training cohort***	**ORFAN UK *internal validation cohort***	**ORFAN USA *external validation cohort***
Total number		2200	600	200	720
Age (years)		60 (50-70)	57 (49-64)	56 (48-62)	53 (43-62)
Male gender		1047 (47.6%)	315 (52.5%)	104 (52.0%)	389 (54.0%)
BMI (kg/m^2)		26.6 (23.6-29.9)	26.1 (23.1-30.0)	27.7 (24.8-30.7)	27.9 (24.5-32.2)
Active smoking		376 (17.0%)	93 (15.5%)	27 (13.5%)	N/A
Hypertension		711 (32.0%)	228 (38.0%)	73 (36.5%)	309 (42.9%)
Hypercholesterolaemia		896 (40.7%)	253 (42.2%)	76 (38.0%)	370 (51.4%)
Diabetes mellitus		286 (13.0%)	101 (16.8%)	25 (12.5%)	73 (10.1%)
Valve disease		139 (6.3%)	31 (5.2%)	13 (6.5%)	49 (6.8%)
Known CAD		188 (8.5%)	55 (9.2%)	21 (10.5%)	331 (46.0%)
Atrial fibrillation		98 (4.5%)	21 (3.5%)	9 (4.5%)	98 (13.6%)
Previous heart surgery		47 (2.1%)	19 (3.2%)	7 (3.5%)	16 (2.2%)
Scans					
Sites (scanner make & model)		Oxford (GE Medical Systems & Toshiba Aquilion One), Bath (Siemens Drive), Milton Keynes (Canon Aquilion Prime SP) & Leicester (Siemens Definition Flash)		Oxford (GE Medical Systems & Toshiba Aquilion One)	Cleveland, USA, (Philips Brilliance iCT & Siemens Definition Flash)
Tube voltage					
	120 kVp	2112 (96.0%)	567 (94.5%)	193 (96.5%)	463 (64.3%)
	100 kVp	81 (3.7%)	33 (5.5%)	7 (3.5%)	257 (35.7%)
	80 kVp	7 (0.3%)	0 (0%)	0 (0%)	0 (0%)
EAT volume (cm^3)		N/A	N/A	120.9 (95.1-156.2)	169.3 (111.6-241.7)

BMI: body mass index; CAD: coronary artery disease; EAT: epicardial adipose tissue; HU: Hounsfield Units; kVp: kilovoltage peak; N/A: not applicable.
Categorical variables are presented as numbers (percentages), while continuous variables are summarized as median [25th-75th percentile].

4.2.12. Challenging patient populations

Clinical CCTA commonly occurs in patients who make CT scanning a challenge for radiographers and radiologists. To ensure the automated segmentation models developed were compatible with these commonly occurring yet challenging patient groups the models were tested in external groupings of patients selected for their specific characteristics. The challenging characteristics were:

- Patients with recent cardiac surgery (<6 weeks post operation)
- Patients with BMI ≥ 40
- Patients with reported coronary artery calcium (CAC) score of ≥ 400
- Patients with significant metallic artefact within the pericardium (metallic valves and/or pacing wires)
- A combined patient group of recent open-heart surgery, BMI ≥ 30 & CAC ≥ 400.

Figure 4.10 demonstrates the scan artefacts that are commonly encountered in these populations. It is these artefacts that render the scans difficult to interpret and troublesome for automated segmentation algorithms such as those being developed in this thesis.

The patients utilised for these groupings came from three studies and were selected based upon their individual characteristic that rendered them relevant. None of the patients included in this analysis were utilised in training or validation of the models. The studies from where these patients were sourced included ORFAN Arm 4, the Oxford based AdipoRedOx Study of patients who had recent cardiac surgery and selected unseen patients from the UK SCOT-HEART Trial, as reported in Chapter 2.

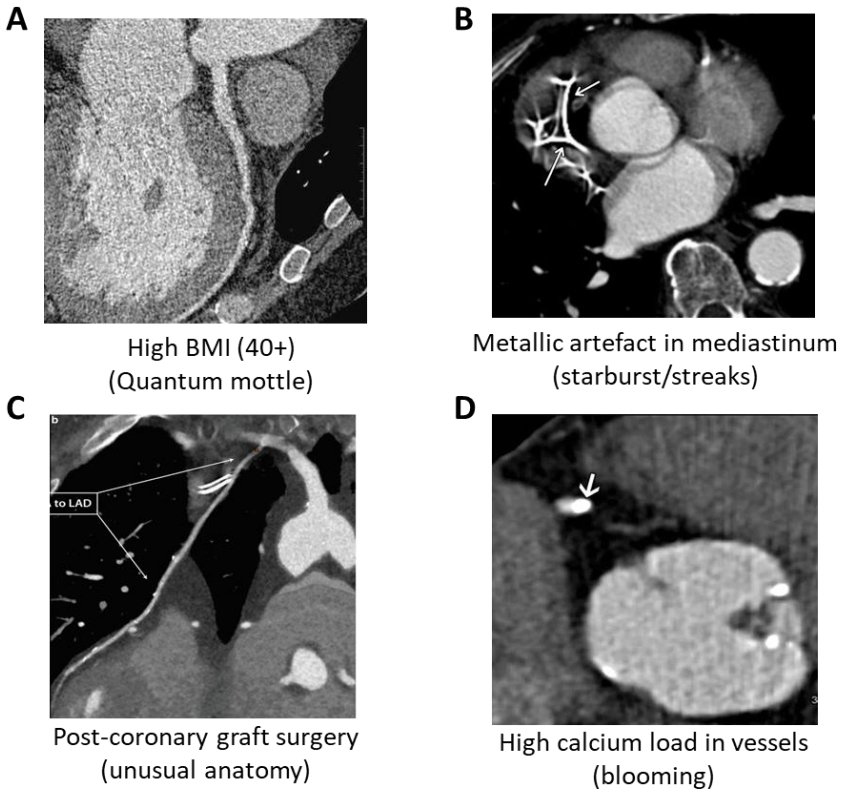


Figure 4.10. Commonly occurring scan artefacts encountered in clinical CCTA. Patients prone to these artefacts were selected for external validation analysis of the segmentation algorithms. Patients were drawn from the ORFAN Arm 4 Study, the AdipoRedOx Study and the SCOT-HEART Trial.

4.2.13. CCTA acquisition settings and exclusions

The CCTA acquisition setting utilised for the ORFAN Arm 4 Sites including international sites, the AdipoRedOx Study and the SCOT-HERAT Trial are all reported in Chapter 2.3.

For all sites, scans with profoundly poor-quality images, more than 5 missing slices, severe beam hardening or significant step artifacts which rendered basic clinical interpretation impossible were excluded from the analysis. Minor to moderate scan artefacts were acceptable and included in analysis, with a tendency to include rather than exclude scans.

4.2.14. The automated segmentation development process

Figure 4.11 summarises the overall development pathway employed for the training of all five automated segmentations. The developmental stage consisted of three major steps:

1. The manual segmentation of the 2200 ORFAN Arm 4 scans
2. A preliminary validation of the automated whole heart segmentation model in 100 sequestered cases

3. Three cycles of feedback learning to further improve the model's performance

The cycles of feedback learning were carried out in batches of 200 randomly selected ORFAN Arm 4 scans. These scans were fed into the automated model and the resulting segments were then manually corrected on 3DSlicer before being fed back into the model for enhanced learning. This process was repeated three times prior to a final deep-learning model being ready for validation testing. The feedback learning occurs in loops (also known as closed-loop learning) and leverages the output of the DLN and subsequent human annotation actions to retrain and improve the DLN over time. The AI-generated output (predictions of correct segmentation on CCTA) were compared against the final human segmentation (through manual editing of the AI prediction) and provided feedback to the model, allowing it to learn from its mistakes and its successes.

The validation stage consisted of three steps for all five segments:

1. Internal validation in 200 ORFAN Arm 4 UK scans specifically sequestered for this purpose
2. External validation in 720 ORFAN Arm 4 USA scans
3. External validation in selected cohorts of commonly occurring yet challenging patient groups

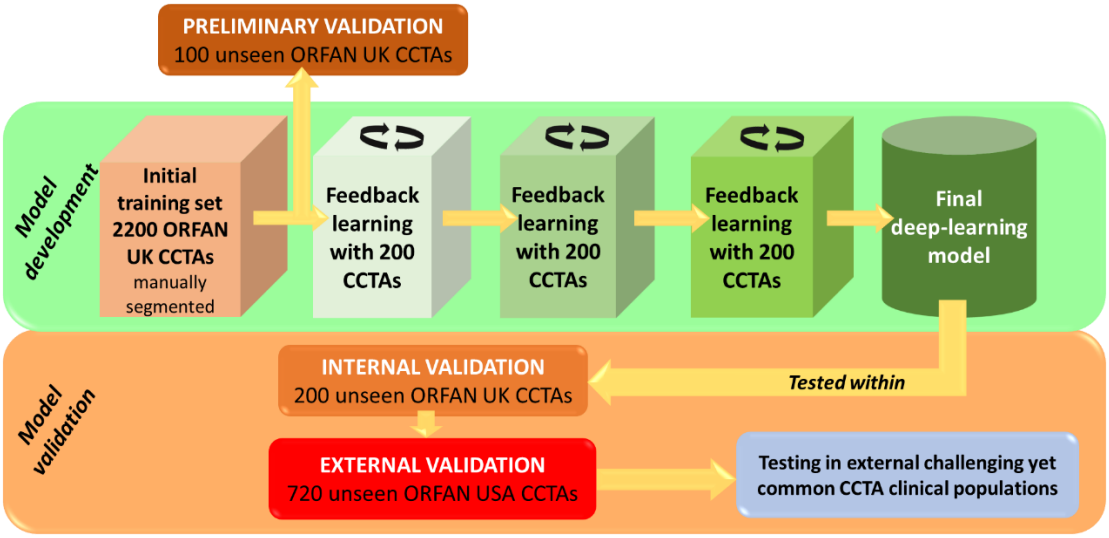


Figure 4.11. Schematic representation of the development pathway for automated segmentations. For all 5 segmentations (whole heart within the pericardium, the LA and the three peri-LA segments), the development for the automated segmentation DLN followed a uniform approach. The manual segmentation of 2200 CCTAs from the ORFAN Arm 4 Study, followed by an initial preliminary validation of the DLN in 100 unseen scans. Three cycles of closed loop feedback learning then occurred to fine-tune the DLN. The final models were then tested internally in 200 UK ORFAN scans, and externally in 720 USA ORFAN scans. The models were ultimately tested in external cohorts of challenging CCTA patients, in a similar manner to the original testing of the manual segmentations. CCTA: coronary computed tomography angiography. **Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.**

4.2.15. Manual segmentation of training dataset

All five developed manual ROIs (whole heart, left atrium and the three peri-LA segments) were segmented at the same analysis session on each CCTA scan. The scans selected for manual segmentation were from the ORFAN Arm 4 Study and comprised of a random assortment of CCTAs from a diversity of CT scanner manufacturers with variable scan acquisition settings (see Table 4.2 for full characteristics of the training dataset).

The average time of each case for analysis of all segments manually was 18.5 minutes. Each scan had to be loaded into 3DSlicer, quality assessed, segmented for each ROI and then saved in a named folder.

2200 CCTAs were initially segmented by me, with assistance from Research Assistance, Ms Lucrezia Volpe. This constituted over 17 weeks of full-time manual image analysis.

Segments were saved as discrete .nrrd files for feeding into the deep-learning training network in an off-line environment. Regular backups of saved data were made to ensure no loss of segmentations.

4.2.16. The deep-learning process

The training of the deep-learning models for automated segmentation of the whole heart within the pericardium, the left-atrium and the three peri-LA segments was conducted in the same manner.

For each segment the automated training was conducted separately but relied upon the same scans that had been manually segmented for all five ROI. Each of the five manual segments upon each CCTA scan were individually normalised to have zero mean and unary standard deviation, as on input it was found that the networks performed better in this intensity range. A convolutional neural network was used for segmentation by casting to a voxel-wise classification. The network was trained to predict whether the central voxel of a patch is or is not within the bounds of the desired segment depending on the content of the surrounding 3D patch.

A fully automatic method was employed using 3D Residual-U-Net neural network architecture[174] for 3D volumetric segmentation of each CCTA segment. A building block of a Residual-U-Net is called a residual block or skip connection. The activation of a layer is fast-forwarded to a deeper layer in the neural network in a residual block. This allows training much deeper neural networks without the issue of vanishing gradients. The architecture of this network resembles an auto encoder framework and consist of 4 down-sampling and 4 up-sampling blocks which are connected by a bridge block. Feature size is halved after passing through every down-sampling block by a max-pooling layer at its base. At the beginning of every up-sampling block the feature size is doubled using a transposed convolution layer with stride of size 2 for each dimension. Passing information through

interconnections was achieved by using zero padding. After every convolution operation the newly computed vector was filled with zeros to restore its original length and therefore both convolution layers use the same size of the input vector. During training, the parameters of the kernels were optimized using Adam optimizer, with the target of minimising the error between the predictions and the true labels. The network architecture was trained using Keras framework using Tensorflow 2.2 (Google Brain, Mountain View, CA, USA) as backend. The network used a binary cross-entropy as a loss function and Dice coefficient as the internal performance metric.

Figure 4.12 provides a visual representation of the deep-learning model employed for all five desired automated segmentations. The development approach allowed for the feeding of a CCTA DICOM file with the five manual segmentations into the network for training.

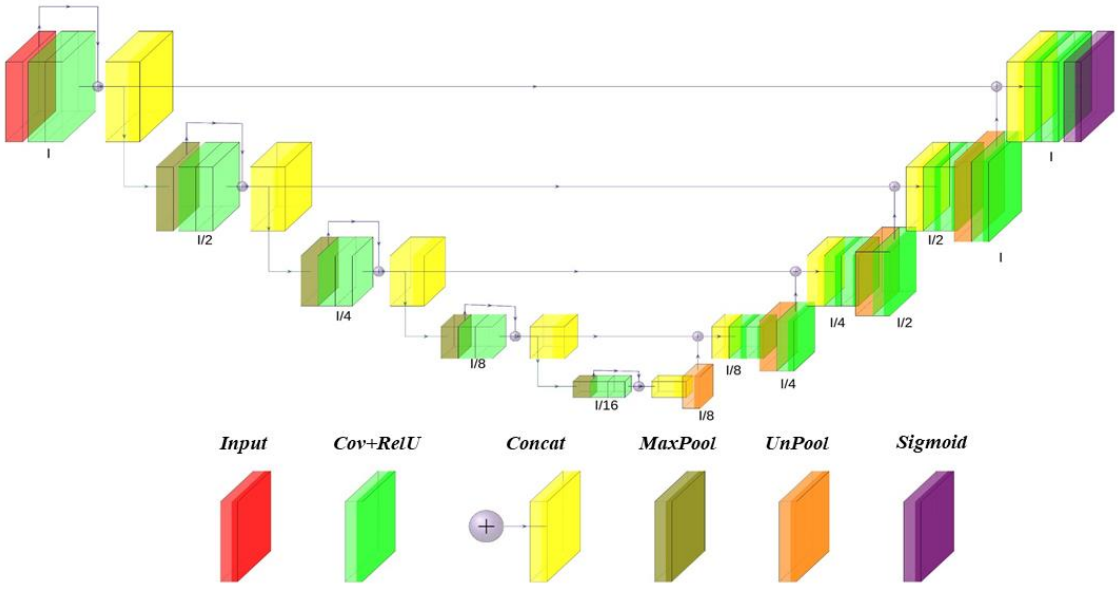


Figure 4.12. Block diagram of the residual U-Net based CNN architecture employed for all automated DLN. The input of the network is a full CCTA dicom scan, with the output being a voxel wise prediction of inclusion within the relevant segment. The light green boxes (Conv+RelU) represent a convolutional layer. These layers capture the spatial relationship between CCTA voxels to receive image features using small patches of input data by learning local connectivity patterns followed by RelU activation to ensure non-linearity, sparsity, and a reduced likelihood of vanishing gradient. Yellow boxes (Concat) represents residual interconnections to allow gradients to pass through the network directly by skipping over series' of convolution and non-linear activation functions to solve the degradation problem of the deep networks, and concatenation to add tensors of both connections. Max pooling layer is represented by dark green colour boxes (MaxPool) to extract sharp and smooth features. Orange boxes (UnPool) represents the up-sampling step to revert the effect of the max pooling operation. The purple box (Sigmoid) represents the output layer using sigmoid function. This outputs values in the range 0 to 1 representing the probability of a voxel being within the desired segmentation of the heart pericardium. The gray horizontal lines represent concatenation of decoder paths with the corresponding feature map from the encoder path. **Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.**

4.2.17. Statistical analysis

All segmentation values were compared as volumes in cubic centimetres. Comparisons between human and machine segmentations in preliminary validation, internal validation and external validation for all segmentation were performed blind to scan data and to the results of automated segmentation.

Comparisons to calculate correlation between human and machine segmentation volumes were made with scatterplots and the assessment of the concordance correlations coefficient (CCC). Comparisons for some segmentations between human and machine were also made via Bland-Altman analysis[175], with the bias (significant or non-significant) reported for each analysis.

4.3. Results

Validation results of the automated segmentation of the whole heart segmentation model and the EAT volume quantification are presented first, followed by the LA and peri-LA segments, alone and then with EAT volumes.

The mean analysis time for the deep-learning automated segmentation of the whole heart within the bounds of the pericardium was 12.4 seconds when computed on GPU compared to mean manual segmentation time of 18 minutes and 20 seconds. When performed on a regular desktop CPU the average automated segmentation time was 22 seconds (mean time of 50 random ORFAN Study cases).

4.3.1. Preliminary validation of whole heart segmentation

Following the manual segmentation of the first 2200 scans, a preliminary validation was performed to ensure the development approach was working for the whole heart within the pericardium. This was performed through automated segmentation of 100 sequestered cases with con-current blinded manual segmentation.

The results of this initial validation are shown in Figures 4.13 and 4.14. For both the whole heart segment volume and the EAT volume alone, the concordance between human and machine was excellent, with CCC of 0.969 for the whole heart (Figure 4.13A) and 0.960 for the EAT (Figure 4.15A). Bias was non-significant for both comparisons on the Bland-Altman (Figure 4.13B and 4.14B)

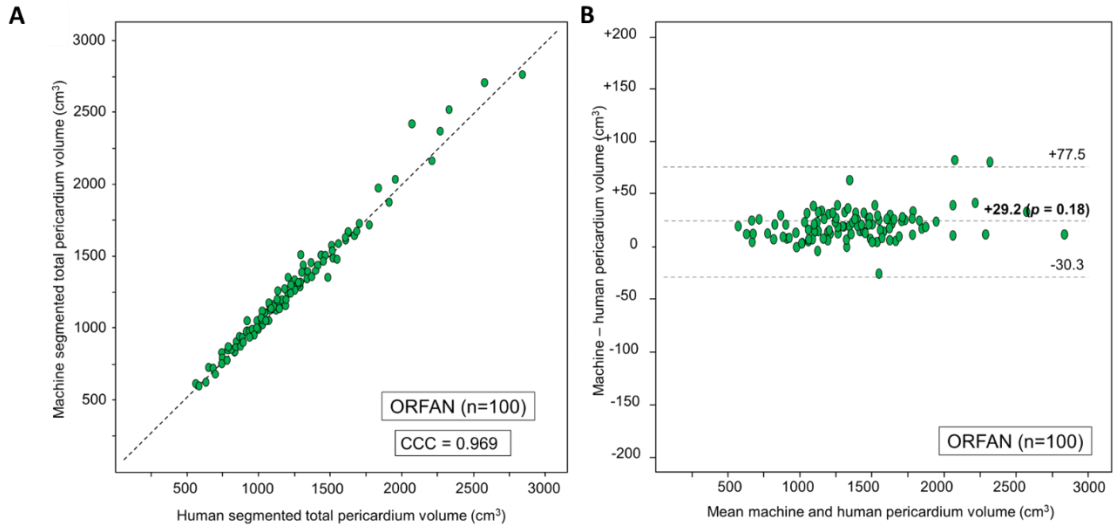


Figure 4.13. Preliminary validation of machine vs human whole heart segment volume assessment prior to feedback learning. Scatterplot (A) and Bland-Altman analysis (B) for comparison of human vs machine segmentation of the whole heart following training of the deep-learning network on manual segmentation but prior to any feedback loop learning. CCC= concordance correlation coefficient

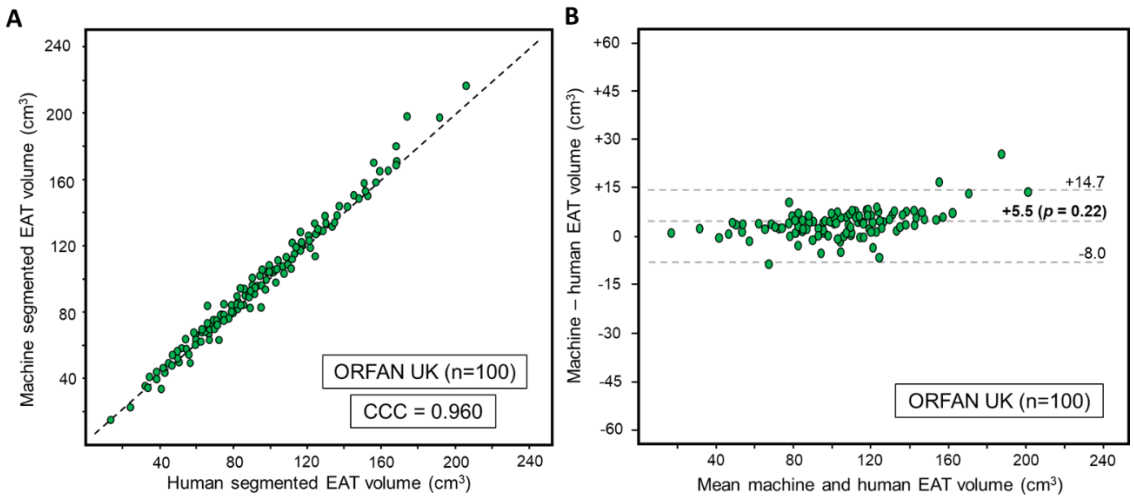


Figure 4.14. Preliminary validation of machine vs human EAT volume assessment prior to feedback learning. Scatterplot (A) and Bland-Altman analysis (B) for comparison of human vs machine segmentation of the EAT following training of the deep-learning network on manual segmentation but prior to any feedback loop learning. CCC= concordance correlation coefficient; EAT: epicardial adipose tissue

4.3.2. Internal validation of whole heart segmentation and EAT volume

Internal validation occurred following the three iterations of feedback loop learning to enhance the performance of the model. The median[IQR] EAT volume in internal validation was 120.9[95.1-156] cm^3. When the DLN was applied to the 200 unseen scans from the UK sites of the ORFAN Study for the whole heart volume within the pericardium the CCC was 0.971 (Figure 4.15A). The bias in Bland-Altman analysis (Figure 4.15B) was also non-significant at +20.7 [-27.0-65.3]cm^3, p = 0.11.

For the EAT volume within the pericardium the CCC between human and automated machine was 0.972 (Figure 4.16A). The bias in Bland-Altman analysis (Figure 4.16B) was also non-significant at 6.1[-11.1-15.7]cm^3, p = 0.19.

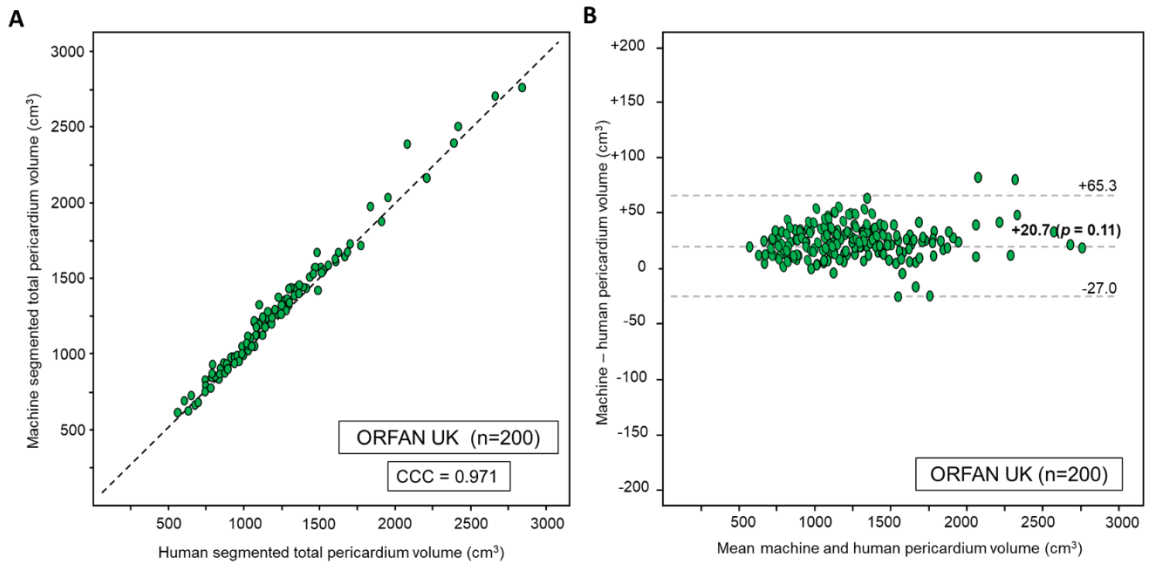


Figure 4.15. Internal validation of the whole heart segmentation deep-learning model. Internal validation of the automated whole heart segmentation model is demonstrated in 200 ORFAN UK cases in the scatterplot (A) and Bland-Altman plot (B).

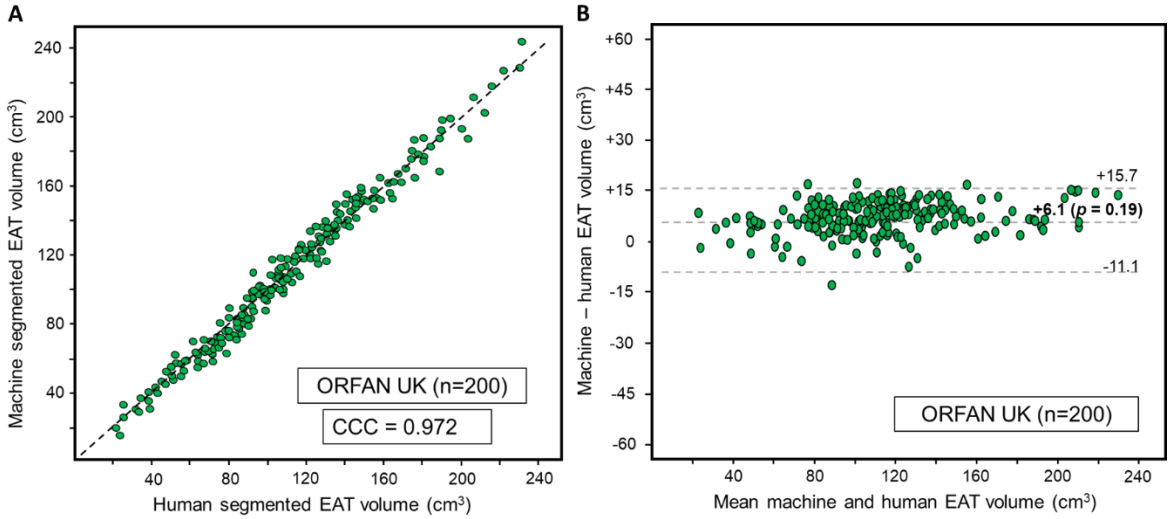


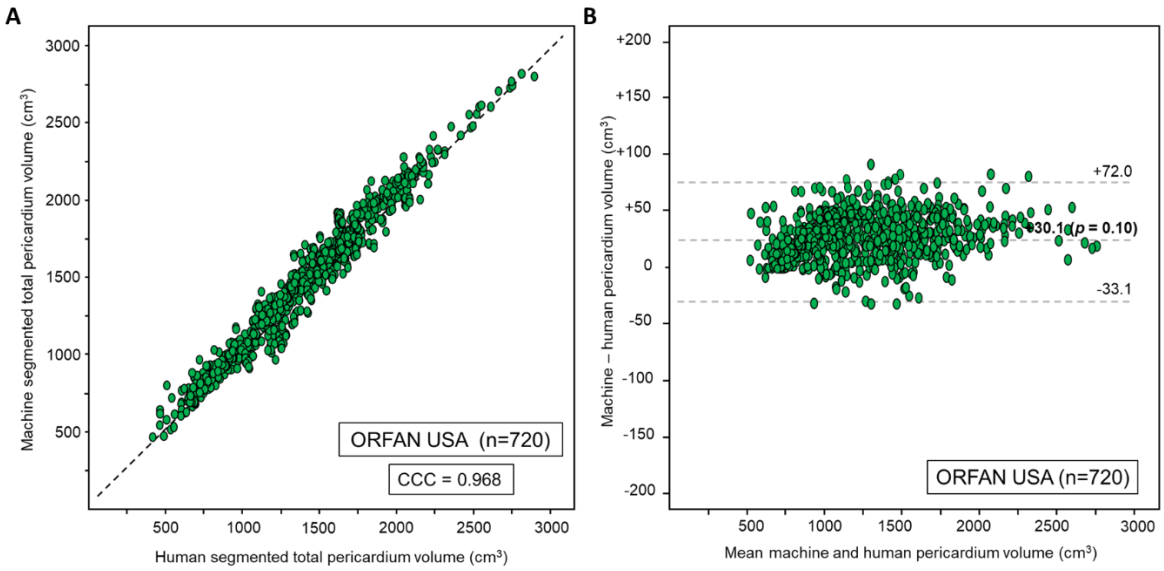
Figure 4.16. Internal validation of the EAT volume model. Internal validation of the automated EAT segmentation model is demonstrated in 200 ORFAN UK cases in the scatterplot (A) and Bland-Altman plot (B). ***Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.***

4.3.3. External validation of whole heart segmentation and EAT volume

The final deep-learning model was applied to the 720 unseen scans from the US sites of the ORFAN Study. The median[IQR] EAT volume in external validation was 169.3[111.6-241.7]cm^3.

For the whole heart volume within the pericardium in the external validation cohort the CCC was 0.968 (Figure 4.17A). The bias in Bland-Altman analysis (Figure 4.18B) was also non-significant at +30.1 [-33.1-72.0]cm^3, p = 0.10.

For the EAT volume within the pericardium the CCC for the automated deep-learning model versus human expert segmentation in the external validation cohort was excellent, at 0.970 (Figure 4.18A) and the bias in Bland-Altman analysis (Figure 4.18B) was non-significant at 3.2[-13.6-17.2]cm^3, p = 0.20.



4.17. External validation of the whole heart segmentation deep-learning model. External validation of the automated whole heart segmentation model is demonstrated in 720 ORFAN USA cases in the scatterplot (A) and Bland-Altman plot (B).

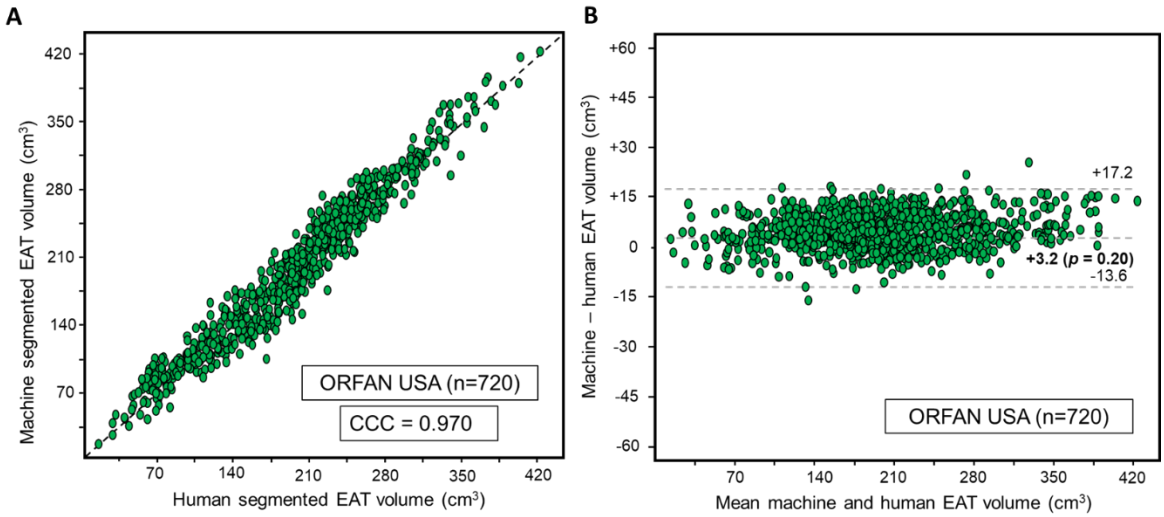


Figure 4.18. External validation of the EAT volume model. External validation of the automated EAT segmentation model is demonstrated in 720 ORFAN USA cases in the scatterplot (A) and Bland-Altman plot (B). **Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.**

4.3.4. Validation of the automated model for EAT volume quantification in challenging clinical populations

The automated deep-learning model for the quantification of EAT volume was applied to sets of unseen CCTA from commonly occurring yet challenging clinical populations for CT analysis. These groups included those who had cardiac surgery in the past (AdipoRedOx Study, UK, total n=253) and selected unseen patients from the SCOT-HEART trial (UK, total n=1558) with relevant characteristics. Excellent CCC for automated EAT segmentation versus human expert segmentation was achieved in all groups: Patients with reported coronary artery calcium (CAC) score of ≥ 400 (Figure 4.19A), CCC = 0.958; patients with BMI ≥ 40, CCC = 0.962 (Figure 4.19B); patients with recent cardiac surgery (<6 weeks post operation) CCC = 0.960 (Figure 4.19C); patients with significant metallic artefact within the pericardium, CCC = 0.955 (Figure 4.19D), and a combined patient group of recent open-heart surgery, BMI ≥ 30 & CAC ≥ 400, CCC = 0.955 (Figure 4.19E).

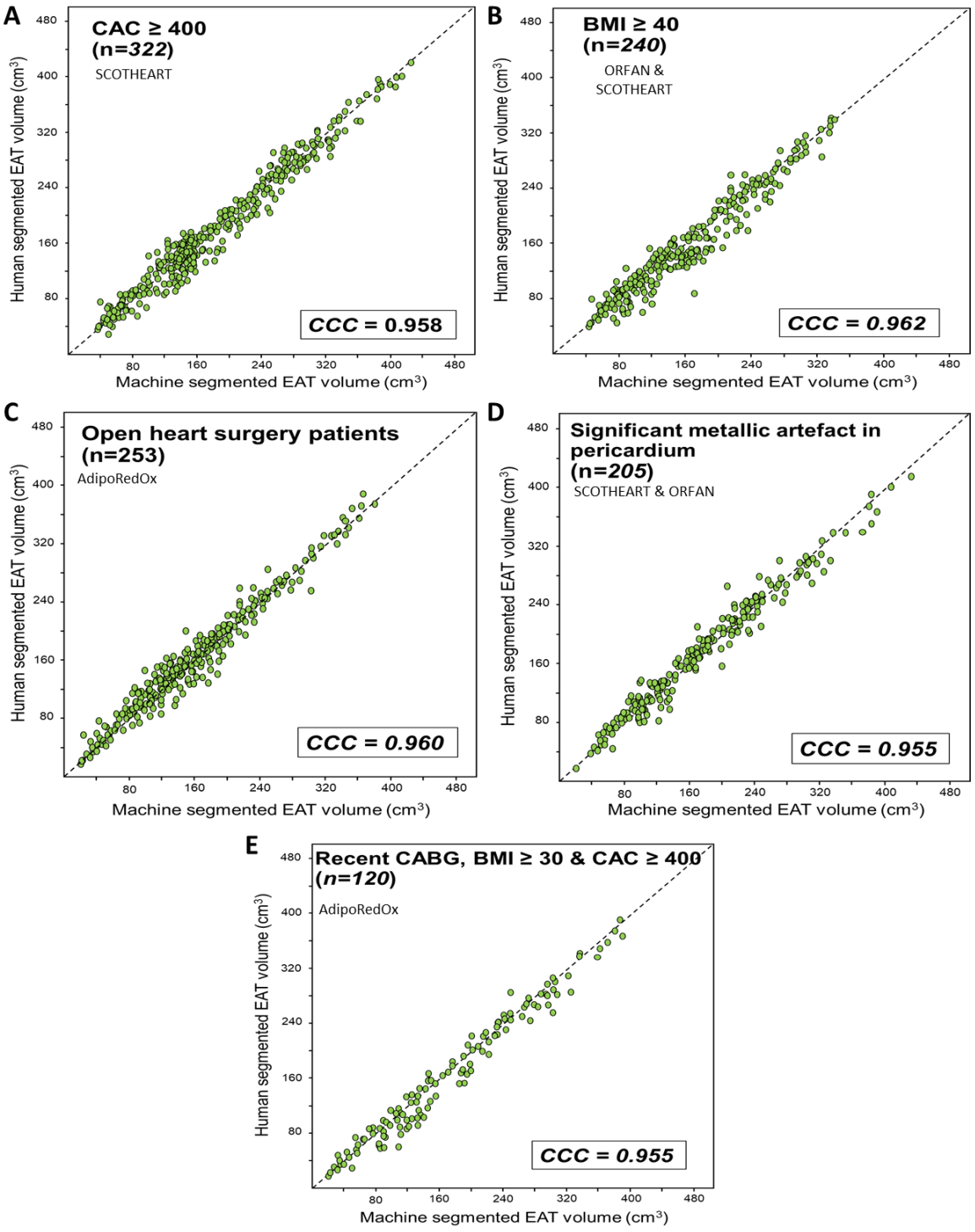


Figure 4.19. External validation of automated EAT segmentation in challenging clinical CCTA cohorts. The automated EAT volume quantification tool was applied to groupings of unseen CCTA from the ORFAN Study, the AdipoRedOx Study and the SCOT-HEART trial. Scatterplots are demonstrated for human vs machine segmentation for (A) patients with CAC ≥ 400; (B) patients with BMI ≥ 40; (C) patients who underwent open heart surgery (CABG) up to six weeks prior to CCTA; (D) patients with significant metallic artefact within the pericardium; (E) patients who underwent open heart surgery (CABG) up to six weeks prior to the CCTA, had BMI ≥30 and CAC ≥400. BMI: body mass index; CABG: coronary artery bypass graft surgery; CAC: coronary artery calcium score; CCC: concordance correlation coefficient; CCTA: coronary computed tomography angiography; EAT:

epicardial adipose tissue. ***Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.***

4.3.5. Internal validation of the automation of the LA and peri-LA segmentations

Internal validation of the LA and peri-LA segments, including the intra-atria segment, the anterior LA segment and the peri-LAA segment occurred in the same manner as the whole heart segmentation. This was in 200 patients from the UK arm of the ORFAN Study. First, the concordance between human and automated machine segmentation for the whole segment volumes was assessed, as shown in Figure 4.20. Excellent CCC were found for each segment. The CCC for the LA segment was 0.981 (Figure 4.20A), for the intra-atrial segment the CCC was 0.975 (Figure 4.20B), for the anterior LA segment the CCC was 0.97 (Figure 4.20C), and for the peri-LAA segment the CCC was 0.961 (Figure 4.20D). Bland-Altman analysis was conducted for each segment also, with all segments found to demonstrate non-significant bias (not shown).

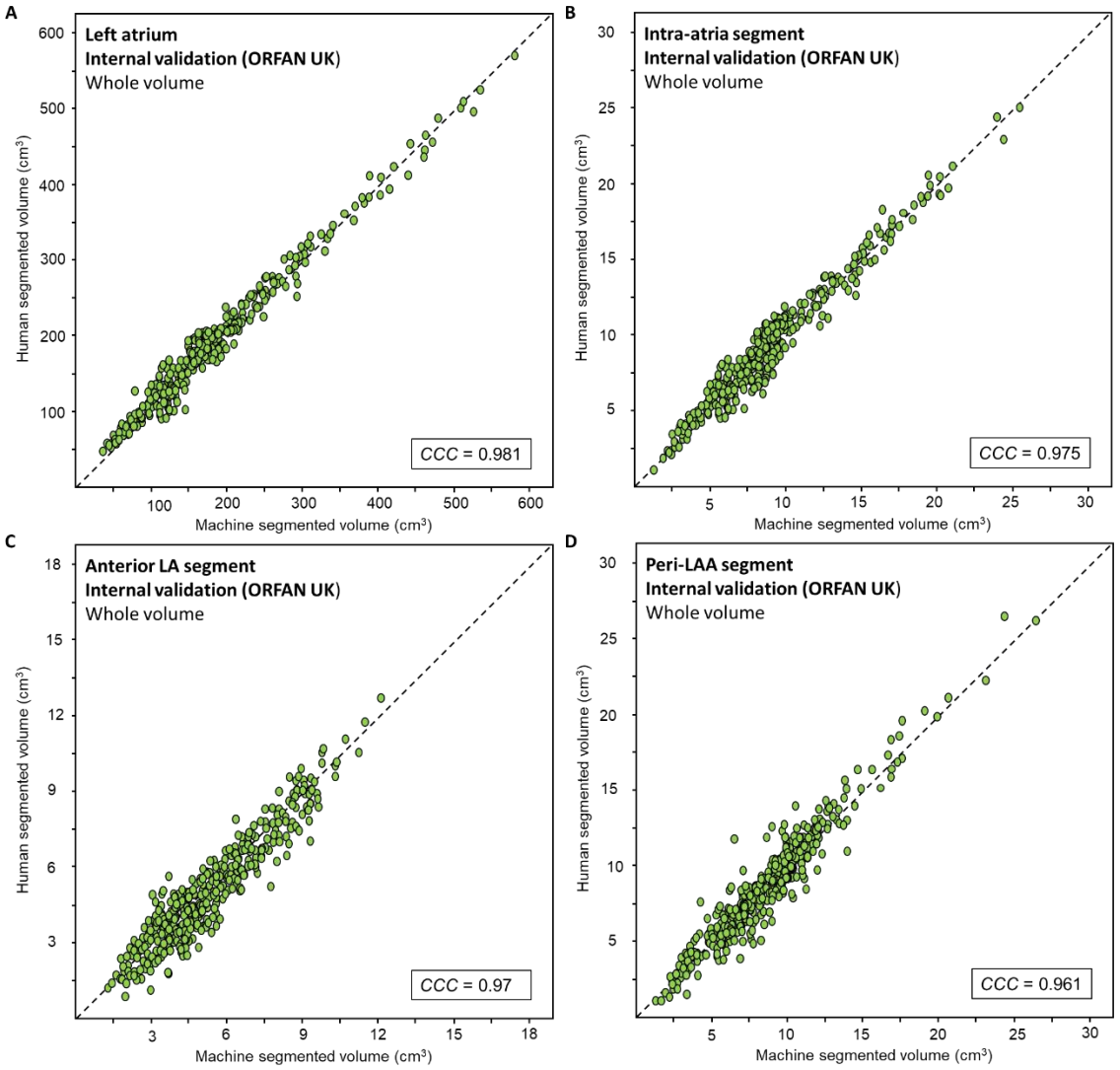


Figure 4.20. Internal validation of whole left atrium and peri-left atrium automated segmentations. Scatterplots demonstrating the internal validation of the whole LA segmentation model(A), the intra-atrial wall segmentation model (B), the anterior LA wall segment (C), and the peri-LAA segment (D). All tested in 200 ORFAN UK cases. CCC; concordance correlation coefficient; LA: left atrium; LAA; left atrial appendage.

Second, the concordance between human and automated machine segmentation for the EAT volumes was assessed for the three peri-LA segments, as shown in Figure 4.21. Excellent CCC were found for each segment. The CCC for the intra-atrial segment was 0.972 (Figure 4.21A), for the anterior LA segment the CCC was 0.969 (Figure 4.21B), and for the peri-LAA segment the CCC was 0.97 (Figure 4.21C). Bland-Altman analysis was conducted for each segment also, with all segments found to demonstrate non-significant bias (not shown).

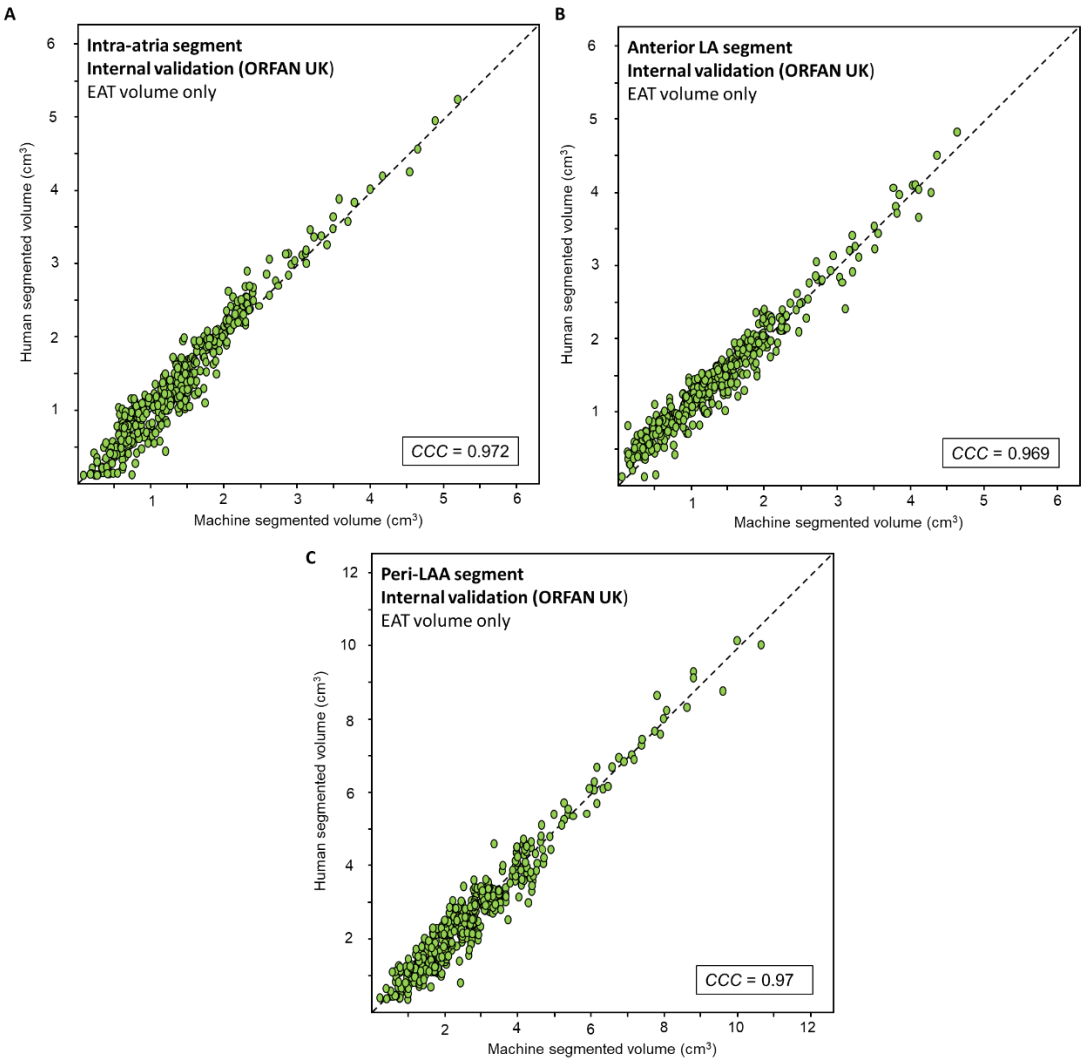


Figure 4.21. Internal validation of automated EAT volume segmentation for peri-left atrium segments. Scatterplots demonstrating the internal validation of the whole LA segmentation model(A), the intra-atrial wall segmentation model (B), the anterior LA wall segment (C), and the peri-LAA segment (D). All tested in 200 ORFAN UK cases. CCC; concordance correlation coefficient; EAT: epicardial adipose tissue; LA: left atrium; LAA; left atrial appendage.

4.3.6. External validation of the automation of the LA and peri-LA segmentations

The final LA and peri-LA segment deep-learning model was applied to the 720 unseen scans from the US sites of the ORFAN Study. For the whole segment volumes in the external validation cohort the CCC for all segments was excellent. The CCC for the LA segment was 0.971 (Figure 4.22A), for the intra-atrial segment the CCC was 0.969 (Figure 4.22B), for the anterior LA segment the CCC was 0.961 (Figure 4.22C), and for the peri-LAA segment the CCC was 0.95 (Figure 4.22D). Bland-Altman analysis was conducted for each segment also, with all segments found to demonstrate non-significant bias (not shown).

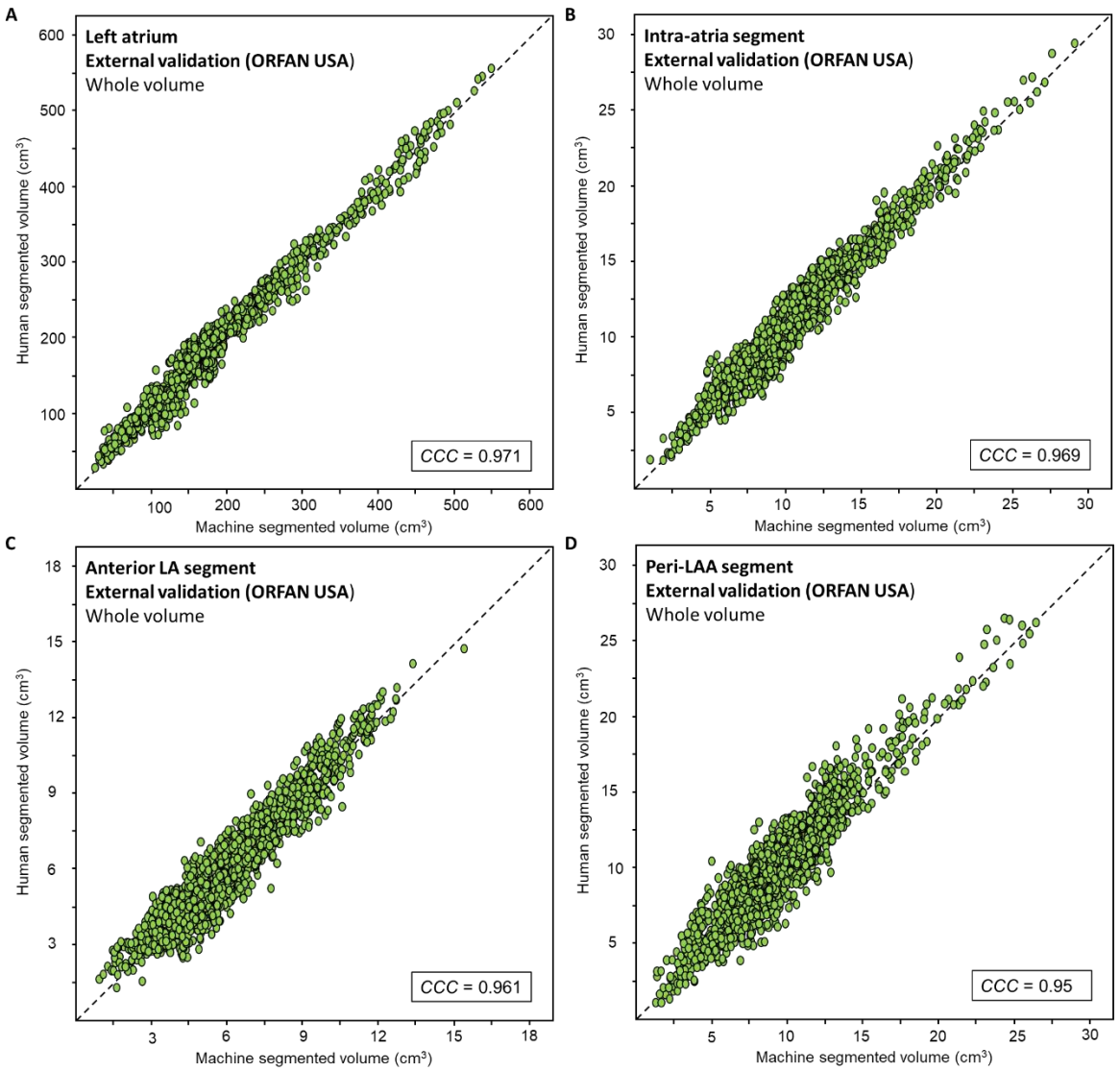


Figure 4.22. External validation of whole left atrium and peri-left atrium automated segmentations. Scatterplots demonstrating external validation of the automated LA segmentation model (A), the intra-atrial wall segment (B), the anterior LA wall segment (C) and the peri-LAA segment (D) in 720 ORFAN USA cases. CCC: concordance correlation coefficient; EAT: epicardial adipose tissue; LA: left atrium; LAA: left atrial appendage.

Subsequently, the concordance between human and automated machine segmentation for the EAT volumes was assessed for the three peri-LA segments, as shown in Figure 4.23. Excellent CCC were found for each segment. The CCC for the intra-atrial segment was 0.962 (Figure 4.23A), for the anterior LA segment the CCC was 0.957 (Figure 4.23B), and for the peri-LAA segment the CCC was 0.963 (Figure 4.23C). Bland-Altman analysis was conducted for each segment also, with all segments found to demonstrate non-significant bias (not shown).

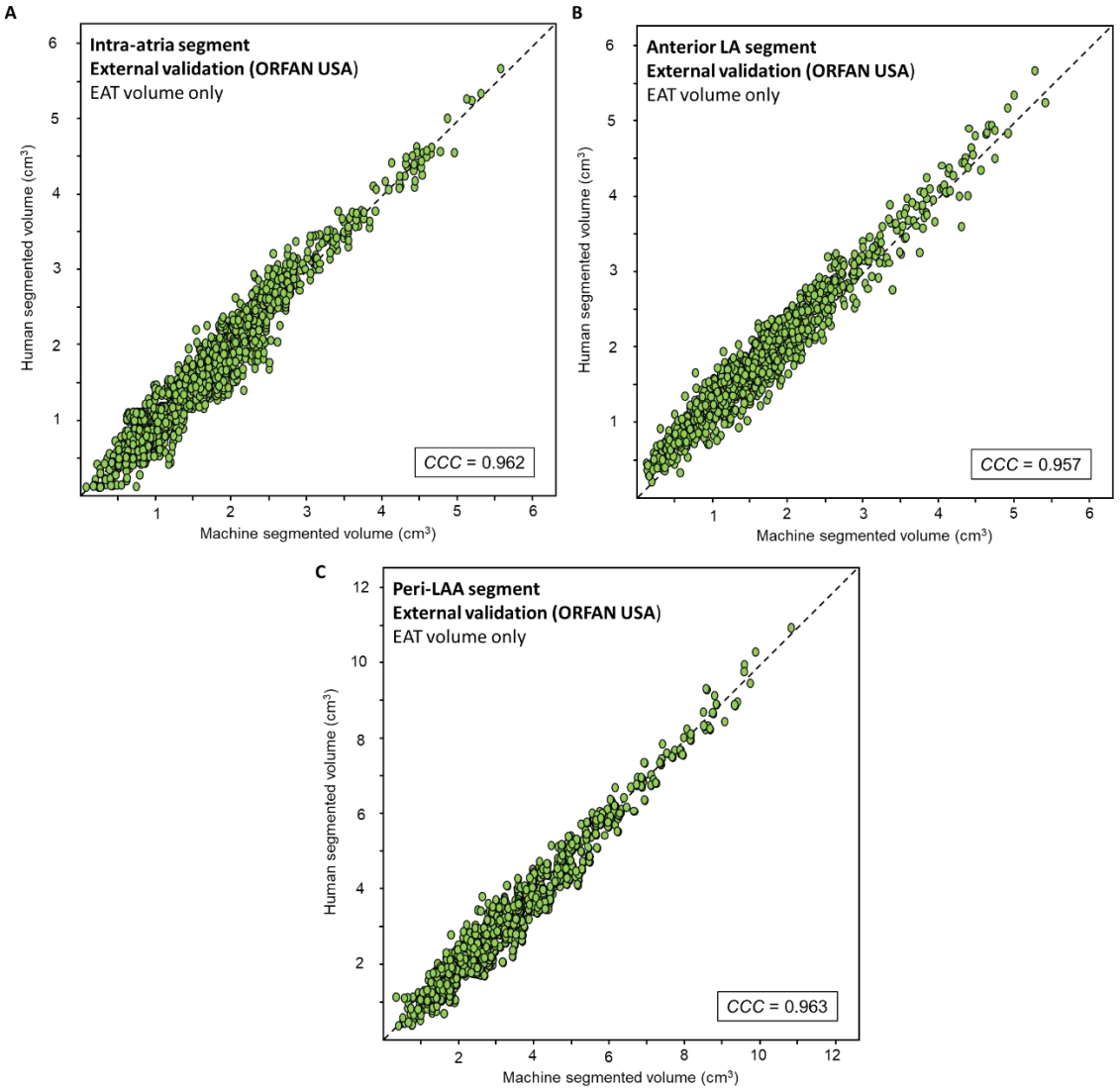


Figure 4.23. External validation of automated EAT volume segmentation for peri-left atrium segments. Scatterplots demonstrating external validation of the automated EAT segmentation model in the intra-atrial wall segment (A), the anterior LA wall segment (B) and the peri-LAA segment (C) in 720 ORFAN USA cases. CCC: concordance correlation coefficient; EAT: epicardial adipose tissue; LA: left atrium; LAA: left atrial appendage.

4.3.7. Validation of the automated LA and peri-LA segmentation in challenging clinical populations

All automated segments, the LA and the three peri-LA segments, were tested in the challenging patient populations. The LA segment was for the whole volume while the three peri-LA segments were for the EAT only. The CCC was calculated for all tests between human and the automated machine, and Bland-Altman analysis was also conducted to generate the mean difference and test the significance of that bias between the readers.

Table 4.4 shows the CCC values, the mean bias and statistical significance of that bias for each segment in each challenging population group. In no comparison between blinded human and

automated machine was there significant bias on Bland-Altman, and no CCC between the two readers was less than 0.95.

Table 4.4. Validation of left atrium and peri-LA automated segments in commonly occurring challenging patient groups

Segment	***Patient groups**** / ***Statistic***	**CAC ≥ 400**	**BMI ≥ 40**	**Recent cardiac surgery**	**Metallic artefact**	**Combined group**
Left atrium†	CCC	0.975	0.960	0.976	0.986	0.978
	Mean difference (cm^3)	+20.4	+26.7	+33.5	+11.1	+26.2
	Bias significance	0.43	0.20	0.13	0.48	0.18
Intra-atrial	CCC	0.97	0.90	0.980	0.975	0.974
	Mean difference (cm^3)	-2.3	+1.2	+2.4	-2.9	+1.9
	Bias significance	0.22	0.39	0.19	0.20	0.31
Anterior LA	CCC	0.98	0.973	0.978	0.971	0.973
	Mean difference (cm^3)	+0.9	+1.2	-1.9	-1.9	+0.8
	Bias significance	0.32	0.30	0.38	0.33	0.28
Peri-LAA	CCC	0.984	0.960	0.964	0.953	0.955
	Mean difference (cm^3)	+3.2	-1.9	+2.8	+4.5	+4.2
	Bias significance	0.21	0.19	0.23	0.11	0.10

*Groups include those with a reported coronary artery calcium (CAC) score of ≥ 400, patients with BMI ≥ 40, patients with recent cardiac surgery (<6 weeks post operation), patients with significant metallic artefact within the pericardium, and a combined patient group of recent open-heart surgery, BMI ≥ 30 & CAC ≥ 400.

†Assessed for whole volume, not EAT volume

BMI: Body mass index; CAC: coronary artery calcium; CCC: concordance correlation coefficient; LA: left atrium; LAA: left atrial appendage

4.4. Discussion

In summary, in this chapter using clinical CCTA scans from across three independent studies, I have demonstrated that:

a) The manual segmentation of the whole heart within the bounds of the pericardium, the LA and peri-LA tissues on CCTA is possible and highly reproducible between readers
b) The automation of reproducible segmentation of the whole heart within the bounds of the pericardium is possible
c) The automation of reproducible segmentation of EAT from within the whole heart segmentation is possible
d) The automation of reproducible segmentation of the LA and peri-LA tissues is possible

The automation of the segmentation of the whole heart, and consequentially EAT, is a necessary developmental step towards a real understanding of the clinical role that EAT plays in cardiovascular disease. Until very recently, studies reporting on the role of EAT in clinical disease outcomes have relied upon manual segmentation. The clinical implications and associations of EAT volume are explored in Chapter 5.

A deep-learning model utilising a single convolutional neural network for the fully automated and rapid quantification of each ROI, including EAT volume from CCTA has been developed. Previous automated models for EAT quantification have predominately been in small cohorts [176,177]. Most relevant is Commandeur et al[136] who developed a CNN capable of automated EAT segmentation and tested in the EISNER cohort. This network is yet to be tested in challenging real-world CCTA scans. The model developed here achieves accurate EAT volume quantification in technically challenging yet commonly occurring clinical populations – the first automated EAT model to do so. The need for any AI-based radiology approaches to be applicable for all-comers is fundamental to patient acceptance and the future uptake of such technology.

The DLN reduced EAT quantification time from an average of 18 minutes when performed manually, to an average of 12 seconds, rendering this tool usable in the clinical environment without adding workload to clinical teams. This time relies upon a GPU being available for analysis, however even on a regular CPU such as a standard desktop computer that is present in a hospital clinic environment the average segmentation time for EAT on a CCTA was on average 22 seconds (mean of 50 segmentations).

The deep-learning model employed for each automated segmentation functioned exceedingly well. The 3D residual U-Net architecture of the convolutional neural network is a relatively new form of three-dimensional image segmentation via deep-learning, but has gained popularity recently with increased clinical interest in automation of CT and MRI analysis[178,179]. U-Net architectures use an encoder/decoder structure which allows integration of multi-scale information with much improved gradient propagation during training. Had this automated CCTA analysis been attempted five years ago, it would have relied on 2D U-Net methodology, which uses only two-dimensional kernels to predict the boundaries of the ROU on a single slice basis within the DICOM. Because of this single slice approach, the 3D U-Net is a major advance, as the network merges segmentations across CT slices, and so enormous context specific to the nature of the ROI can be learnt and carried into ROI predictions within the volumetric DICOM inputs.

Limitations

The limitations of the automated segmentation models are few, as they were trained in diverse clinical cohorts from multiple different CT scanners and validated in true external cohorts. In regards to the use of the automated algorithms in all possible challenging clinical CCTA populations, other populations that were not tested for in this analysis do exist. Such populations include those with congenital cardiac conditions and those who have undergone heart transplant. For both these populations we lacked enough data for testing of the DLN. With the development of the ORFAN Arm 4 study, it will soon be possible to validate all models in these most challenging and more uncommon populations.

4.5. Conclusion

The automated segmentation of the EAT in clinical CCTA populations is possible and highly accurate via deep-learning means, reducing segmentation time from 18 minutes per case to <22 seconds on a regular desktop computer. The automated segmentation of the LA and peri-LA tissues for extraction of radiomic features is possible and highly accurate via deep-learning means.

4.6. Personal contribution

I undertook the planning and actioning of all transfers of data within the ORFAN Arm 4 Study. I also undertook the unification and cleaning of CCTA data and clinical datasets utilised in this chapter.

I developed and validated the manual segmentation techniques. I undertook the majority of manual analysis of all studies for segmentation of the whole heart, the LA and the peri-LA regions. Given the

time-consuming nature of segmentations the completion of this study would not have been possible without the contribution of Ms Lucrezia Volpe, Ms Katarina Dangas, and Ms Ria Desai who performed manual segmentations in ORFAN Arm 4 under my supervision. Radiomic and statistical analyses, including machine learning analyses and the training of algorithms in ORFAN Arm 4 were performed by me. Deep-learning model training and fine-tuning was overseen by Dr Muhammad Siddique. Qualitative and quantitative coronary artery analysis on CCTA scans utilised in the thesis were conducted by Drs Evangelos Oikonomou and Christos Kotanidis.

Also see Chapter 3 for details of my contribution to the development of the ORFAN Arm 4 study.

5

Chapter 5. EAT and cardiovascular disease outcomes

Parts of this chapter are drawn from the publication West HW et al. 'Deep-Learning for Epicardial Adipose Tissue Assessment with Computed Tomography: Implications for Cardiovascular Risk Prediction. Journal of the American College of Cardiology: Cardiovascular Imaging. Accepted 12 November 2022

Further, parts of this chapter were presented at the British Atherosclerosis Society Annual Meeting 2021 (Oxford, UK) in the abstract West HW et al. 'Deep learning detection of epicardial adiposity on cardiac CT predicts mortality and cardiac events', which won the Best Early Career Oral Presentation Prize.

Part of this chapter were presented at the British Cardiac Society Annual Meeting 2022 (Manchester, UK) in the abstract West HW et al. 'Automated deep-learning quantification of epicardial adiposity on cardiac CT predicts atrial fibrillation risk immediately following cardiac surgery and long term', which won the Best of the Best Clinical Abstract Prize.

5.1. Introduction

This chapter applies the automated segmentation technology developed in the previous chapter to explore the clinical applications of the whole heart deep-learning segmentation model for the automated quantification of EAT. The main objective of this study is to understand the clinical associations and correlations between EAT volume and major cardiovascular disease outcomes.

CCTA is increasingly used to evaluate CAD risk[180], with guidelines in Europe[22,181] and USA[170] recommending CCTA for assessment of patients with stable chest pain. The use of CCTA for detecting CAD is increasing world-wide and it is likely that further valuable information within CCTA scans that is currently not fully utilised in clinical practice could improve risk assessment and patient management. The volume of EAT is one such area of valuable information neglected if current clinical practice.

Adipose tissue is recognised as a key regulator of cardiovascular health and disease, exerting both protective and deleterious effects on the cardiovascular system[120]. EAT is a highly metabolically active deposit of visceral fat[61], and may be a feature of metabolically unhealthy obesity [113,182,183]. Indeed, EAT volume has been associated with multiple cardiovascular diseases, as was outlined in the introduction to this thesis, including CAD[184], AF[116,185,186] and heart failure[187,188]. CCTA provides the non-invasive gold standard measurement of EAT volume due to its excellent spatial resolution and

ease of examination. However, the process of manual quantification of EAT on CCTA is labour intensive, and it currently falls outside the scope of routine CCTA interpretation. If clinical utility of EAT volume could be demonstrated, then it could be possible that this measure could become part of standard of care by incorporating it into routine CCTA reporting, most likely through automated integration with radiology services.

As already explored in detail in previous chapters, deep learning can provide an automated solution for EAT volume measurement. We developed the automated deep-learning network necessary for the segmentation of EAT on routine clinical CCTA scans and validated it in all-comers. This validation included patients with commonly occurring yet challenging CT characteristics (see Chapter 4).

In the study presented in this chapter we applied the fully automated EAT quantification tool to investigate the clinical association of EAT volume with relevant cross-sectional and longitudinal disease outcomes including all-cause mortality, non-fatal myocardial infarction, non-fatal stroke, and the occurrence of post-operative AF.

The scientific aim of the thesis that is achieved in this chapter is to explore the ability of EAT volume to detect cardiovascular disease and predict risk of future cardiovascular disease.

5.2. Methods

5.2.1. Overall study design

The deep-learning model for automated whole heart segmentation and quantification of EAT volume was applied to the unseen external scans of the SCOT-HEART trial for real world evaluation of the prognostic value of EAT volume as a marker of metabolically unhealthy obesity. The model was also applied within the Oxford based AdipoRedOx Study to explore the prognostic value of EAT volume on the risk of post-cardiac surgery atrial fibrillation.

5.2.2. Study populations

For each study/trial (ORFAN, AdipoRedOx and the SCOT-HEART Trial) detailed study population descriptions, CCTA acquisition protocols, and scan retrieval processes are outlined in Chapter 2.

As discussed in Chapter 3, the ORFAN Study is an international multi-centre observational study which collects CCTA scans and patient clinical data from those who are undergoing (prospective arm) or have undergone CCTA since 2005. For this analysis, data was utilised from across four NHS sites in England, one in Scotland and one in the USA. The AdipoRedOx study involves patients who are

undergoing cardiac surgery as part of the study patients undergo CCTA shortly after their operation and are prospectively followed-up for clinical outcomes. The SCOT-HEART trial included clinical patients with suspected angina due to coronary heart disease, who were followed up for five years post-CCTA for clinical outcomes.

ORFAN Study

The ORFAN Study enrols adult patients who receive clinically indicated CCTA for all accepted indications at collaborating hospitals. Enrolment is through individual clinical sites and includes 13 NHS Trusts in England and two clinical sites in the USA. ORFAN scans utilised in this study were performed between January 2015 through to December 2020. For each patient included in the study the CCTA scan was pseudo-anonymised and delivered to OXACCT Core Lab for analysis. As reported in Chapter 4, 2200 ORFAN patients' scans were utilised for the manual training of the deep-learning model while 600 further unseen scans were utilised for feedback learning. 100 unseen UK ORFAN scans were utilised for preliminary evaluation of the model, and 200 unseen UK ORFAN scans were used for internal validation of the model. 720 ORFAN USA scans (Cleveland Clinic, OH) were used only for external validation of the model and remained sequestered throughout the development and internal validation process. No outcome analysis of the main ORFAN Arm 4 cohort was performed in this Chapter.

AdipoRedOx Study

The AdipoRedOx study enrols cardiac surgery patients at the Oxford University Hospitals NHS Foundation Trust, UK. The study is designed to investigate the role of inflammation and redox-state biology on heart disease through the collection of heart, vessel, and adipose tissue at the time of surgery. Relevant to this study, participants also undergo a CCTA within six weeks following surgery. Scans and relevant patient data from 253 participants were utilised in this study for the validation of the automated EAT quantification tool in challenging post-cardiac surgery patients. Outcomes were collected in hospital, including confirmed atrial fibrillation (AF) during the in-patient stay immediately following the surgery. Long term outcomes including new diagnosis of AF were collected from NHS Digital data via linkage with identifiers (NHS number and date of birth) to hospital episode statistics data for all regular admissions, outpatient episodes, emergency presentations and critical care episodes following the surgery.

SCOT-HEART trial

SCOT-HEART is an open-label, randomised, controlled, parallel-group trial that was performed at 12 centres across Scotland, commencing in 2009. Adult patients (n=4126) referred to outpatient cardiology clinics for stable chest pain were randomly assigned in a 1:1 ratio to standard care plus CCTA or to standard care alone. No standard-care patients were utilised in this study. The trial has

been described in detail previously[151,189]. The median duration of follow-up was 4.8 years. In this study, 1,558 eligible patients who had been randomised to the CCTA arm of the study were utilised.

5.2.3. Coronary computed tomography angiography technical parameters

The technical scan acquisition protocols for the ORFAN Study, the AdipoRedOx Study and the SCOT-HEART Study are included in Chapter 2.3.

5.2.4. Study ethics

All studies utilised in this chapter complied with the human research ethics committees of the coordinating institute. The ORFAN Study (15/SC/0545) and the AdipoRedOx Study (C11/SC/0140) are sponsored by the University of Oxford and are both conducted with approval of the South Central – Oxford C Research Ethics Committee, as part of the Oxford Heart Vessels and Fat (ox-HVF) programme. The SCOT-HEART trial was sponsored by the University of Edinburgh and was conducted with the approval of the South East Scotland Research Ethics Committee (NCT01149590). ORFAN USA sites research protocol was approved by all local institutional review boards, including material and data sharing agreements, with waiver of individual informed consent.

5.2.5. Statistical analysis

Participant demographics are summarised as numbers (percentages) or median (25th to 75th percentile) for categorical and continuous variables, respectively.

When applied in a cohort-wide setting in the AdipoRedOx and SCOT-HEART studies, all EAT volumes were standardised by patient body surface area using the Du Bois formula[190]. For cross-sectional analysis of disease risk conveyed by EAT volume, multiple-adjusted logistic regression was used for calculation of the odds-ratio of prevalent disease given increase in EAT by one standard deviation (SD). All analysis were adjusted for standard disease risk factors selected from the literature for each condition and are listed in the results and figure legends.

Longitudinal assessment of the prognostic value of EAT volume was performed by multivariable Cox regression models adjusted for standard disease risk factors for each outcome. Hazard ratios are reported per one SD increase in EAT volume. When utilised, coronary artery calcium (CAC) score was log-transformed before inclusion in regression models (ln[CAC + 1]). Receiver operator characteristic curves for the discrimination of prevalent CAD and non-fatal MI were plotted with a traditional risk factor models vs traditional risk factor model plus the addition of EAT volume. Area under the curve analysis was undertaken to compare the models for each outcome.

The optimum unified cut-off for EAT volume prognostication for all-cause mortality, non-fatal MI and non-fatal stroke in SCOT-HEART was selected by identifying the value that maximised Youden's J

statistic (sum of sensitivity and specificity) on time-dependent receiver operating characteristic (ROC) curve analysis for all-cause mortality, non-fatal MI and non-fatal stroke to ensure an optimum balance between sensitivity and specificity in our models. The influence of each outcome according to the number of events in the population was weighted and then tested the prognostic value of the high vs low EAT volume (as a dichotomous variable) by multivariable Cox regression analysis after adjustment for relevant disease risk factors as listed in the results. Statistical analyses were performed predominantly in STATA SE version 15 (Stata Corp, College Station, TX, USA), with some analysis in the *R* environment (*R for* windows 4.0.4). All tests were two-sided and α was set at 0.05, unless specified otherwise.

5.3. Results

The geographic location, demographics, clinical risk factors, and CCTA scan technical parameters for all ORFAN Study cohorts utilised in DLN training, validation and external testing were presented in Table 4.3.

The demographics and scan characteristics of the external clinical cohorts for which the AI EAT segmentation model was applied are shown in Table 5.1, these are the cohort for which results will be demonstrated herein.

The relevant clinical outcomes for the prospective clinical cohorts are presented in Table 5.2.

All results are reported with EAT volumes standardised for BSA. In all cohorts where EAT volumes were quantified, both manually and automatically, the values were normally distributed.

Table 5.1. Demographics and scan characteristics of external clinical cohorts

		AdipoRedOx *cohort*	**SCOT-HEART *cohort***
Total number		253	1558
Age (years)		67 ± 12.6	58 ± 9.8
Male gender		220 (87.0%)	887 (56.9%)
BMI (kg/m^2)		27.9 ± 5.4	28.7 ± 6.1
Active smoking		127 (50.2%)	298 (19.1%)
Hypertension		186 (73.5%)	541 (34.7%)
Hypercholesterolaemia		227 (89.7%)	810 (52.0%)
Diabetes mellitus		55 (21.7%)	173 (11.1%)
Valve disease		48 (19.0%)	144 (9.2%)
Known CAD		253 (100%)	158 (10.1%)
Atrial fibrillation		18 (7.1%)	32 (2.1%)
Previous heart surgery		253 (100%)	33 (2.1%)
Scans			
Site Scanner make & model		Oxford -GE Revolution GSI -GE Lightspeed VCT -Canon Aquilion One	Scotland -Philips Brilliance 64, -Siemens Biograph mCT -Canon Aquilion One
Tube voltage			
	120 kVp	253 (100%)	846 (54.3%)
	100 kVp	0 (0%)	712 (45.7%)
	80 kVp	0 (0%)	0 (0%)
EAT volume (cm^3)		141.9 ± 46.1	117.1 ± 42.9

BMI: Body mass index; CAD: Coronary artery disease; EAT: Epicardial adipose tissue; GE: General Electric.
Values expressed as n (%) or mean ± SD

Table 5.2. Outcomes from the prospective clinical cohorts		
Prospective follow-up		
	AdipoRedOx ***cohort of cardiac surgery***	**SCOT-HEART** ***outcomes cohort***
Duration (unique units)	6 ± 1.1 days in hospital 7.5 ± 1.1 years long term	4.8 ± 0.8 years
All-cause mortality (long term)	31 (12.3%)	35 (2.25%)
Cardiac mortality (long term)	20 (7.9%)	4 (0.26%)
Non-fatal MI (long term)	9 (3.6%)	35 (2.25%)
Non-fatal stroke (long term)	8 (3.2%)	8 (0.51%)
Post-operative AF in-hospital	97 (38.3%)	N/A
Post-operative AF long term	48 (19%)	N/A
Data are mean ± SD or n (%); AF: atrial fibrillation; MI: myocardial infarction		

5.3.1. Cross-sectional clinical correlations

At the time of the CCTA, application of the fully automated segmentation tool for quantification of EAT volume was found to be a significant independent predictor of the presence of AF at time of CCTA and obstructive CAD from CCTA (any one coronary vessel with ≥50% stenosis on CCTA), within 1,558 patients randomised to receive CCTA in the SCOT-HEART trial population.

When accounting for AF risk factors the odds ratio (OR[95%CI]) of AF at time of CCTA per 1 standard deviation (SD) increase of EAT was 1.20[1.06-1.45] p=0.03 (Figure 5.1A). When accounting for CAD risk factors the OR of obstructive CAD from the CCTA per 1 SD increase of EAT was 1.12[1.03-1.30] p=0.01 (Figure 5.1B). See Table 5.1 for EAT standard deviation values.

5.3.2. Longitudinal EAT volume clinical correlations

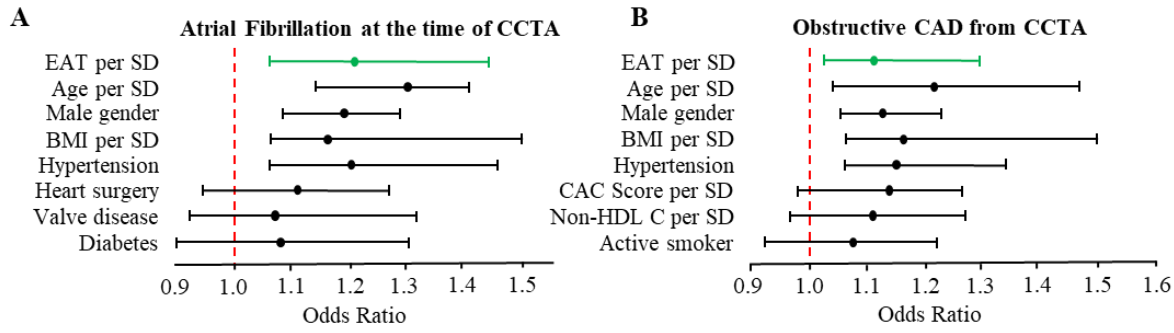
Median follow up for the 1,558 patients randomised to receive CCTA in SCOT-HEART which were analysed was 4.8 years. There were 35 deaths of all causes (2.25%) of which 4 (0.25%) were deaths related with coronary heart disease. There were 8 non-fatal strokes (0.51%) and 35 non-fatal myocardial infarctions (2.25%).

The hazard ratio (HR[95%CI]) of all-cause mortality per 1 SD increase of EAT was 1.24[1.08-1.34] p=0.03, after accounting for relevant general disease risk factors (Figure 5.1C). When adjusted for the same risk factors as Figure 5.1A, the HR[95%CI] of non-cardiac mortality per 1 SD increase of EAT volume was 1.14[1.02-1.37] p=0.04 (Figure 5.1D). This constitutes a ΔHR of -0.10, confirming that EAT is a measure of visceral adipose tissue related with multiple fatal pathologies, unrelated to CAD. When accounting for CAD risk factors the HR of non-fatal MI per 1 SD increase of EAT was 1.25[1.08-1.45] p=0.001 (Figure 5.1E). Finally, when accounting for stroke risk factors the HR[95%CI] of non-

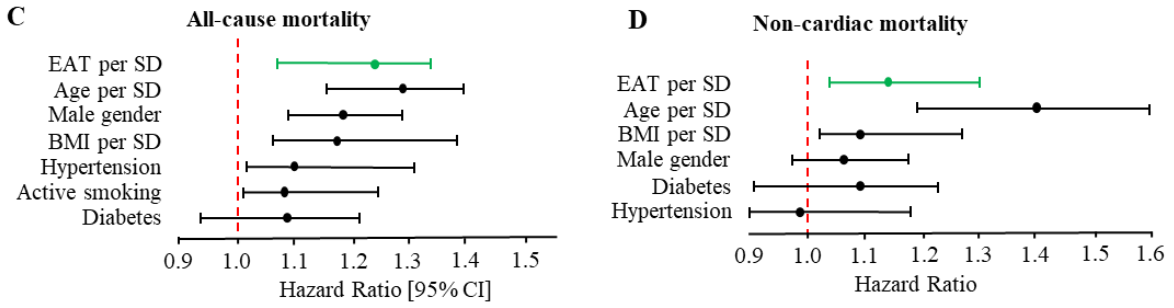
fatal stroke per 1 SD increase of EAT is 1.12[1.03-1.27] p=0.02 (Figure 5.1F). See Table 5.1 for EAT standard deviation values.

Adding EAT volume into a clinical model led to significant improvement in the ability to detect obstructive CAD on CCTA (any one coronary vessel with ≥50% stenosis on CCTA) cross sectionally (Figure 5.1G, ΔAUC 0.103 ($p<0.01$)) and improved the ability to predict future non-fatal MI longitudinally (Figure 5.1H, ΔAUC 0.06 ($P<0.01$).

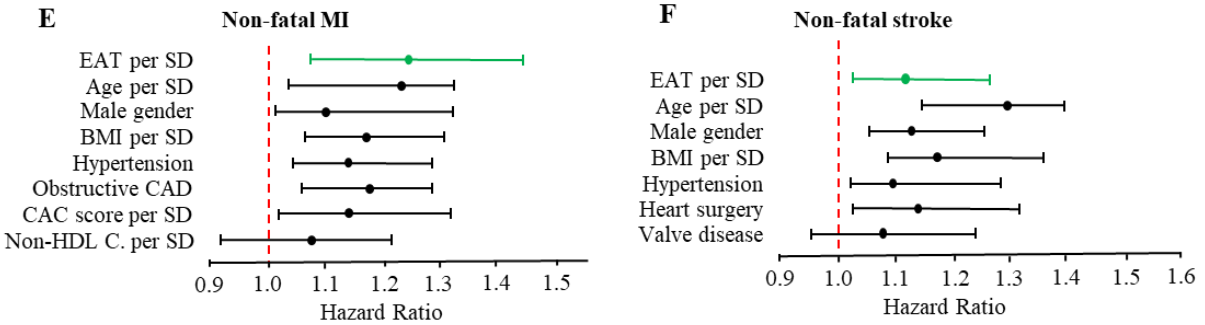
Cross-sectional evaluation of EAT volume SCOT-HEART



Longitudinal EAT volume and mortality in SCOT-HEART



Longitudinal EAT volume and non-fatal acute vascular events in SCOT-HEART



ROC curve analysis with EAT volume from SCOT-HEART

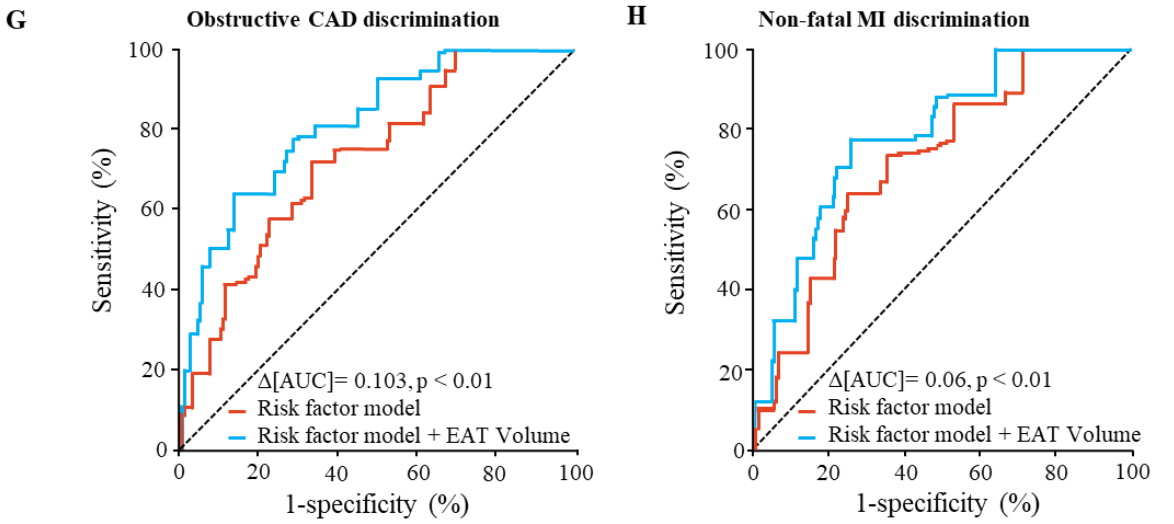


Figure 5.1. Cross-sectional and longitudinal associations between EAT volume and clinical outcomes in the SCOT-HEART trial. *Plots of cross-sectional adjusted risk models for diagnosed AF at the time of the CCTA, adjusted for age, gender, BMI, presence of hypertension, previous heart surgery, presence of valvular heart disease and presence of diabetes (A), and obstructive CAD from CCTA (any one coronary vessel with ≥50% stenosis on CCTA), adjusted for age, gender, BMI, presence of hypertension, CAC score, non-high density lipoprotein cholesterol, and active smoking (B). Odds ratio is shown per 1 standard deviation (SD) increase in EAT volume for 1558 patients randomised to receive CCTA in the SCOT-HEART trial. Plots of longitudinal hazard ratios per 1 SD increase in EAT volume in 1558 patients randomised to receive CCTA in the SCOT-HEART trial are shown for all-cause mortality (C), and non-cardiac mortality (D), both with adjustment for age, gender, BMI, presence of hypertension and diabetes, with further adjustment for active smoking in 6C. Non-fatal MI (E) is shown with adjustment for of age, gender, BMI, presence of hypertension, presence of diabetes, non-high density lipoprotein cholesterol, obstructive CAD from CCTA (any one coronary vessel with ≥50% stenosis on CCTA), and coronary artery calcium score, and non-fatal stroke (F) is shown with adjustment for age, gender, BMI, presence of hypertension, presence of valvular heart disease and previous heart surgery. ROC curves are shown for the discrimination of present CAD (G) and non-fatal MI (H). The risk factor model (blue) for each curve includes age, male gender, BMI, presence of hypertension, non-HDL cholesterol, presence of diabetes and CAC score. Red in both curve is the risk factor model with the addition of EAT volume.*
AF: atrial fibrillation; BMI: body mass index; CAC: coronary artery calcium score; EAT: epicardial adipose tissue; MI: myocardial infarction; Non-HDL C.: non-high density lipoprotein cholesterol.
Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.

5.3.3. A single predictive value of EAT volume

To generate distinct clinical groups based around a single EAT volume cut-point, the SCOT-HEART study population was dichotomised into high versus low EAT volume groups based on an optimum cut-off of 170.5cm^3 for the three primary outcomes of interest: all-cause mortality, non-fatal MI and non-fatal stroke. This cut-off was derived from a weighted Youden's J index analysis to generate a single value for application. Utilising this cut-point, there was significantly different adjusted hazard ratios for the low EAT group (<170.5cm^3) versus the high EAT group (≥170.5cm^3) for all outcomes. All analyses included multivariable adjustment for relevant risk factors including BMI.

High EAT volume values (≥170.5cm^3 vs <170.5cm^3) were associated with a higher prospective risk for non-fatal myocardial infarction (adjusted HR 1.88[1.23-3.97] p<0·01; Figure 5.2A), non-fatal stroke (2.21[1.04-4.89] p<0·01; Figure 5.2B), non-cardiac mortality (3.81 [1.47-14.13] p<0·001; Figure 5.2C), and all-cause mortality (4.98[2.87-10.77] p<0·0001; Figure 5.2D).

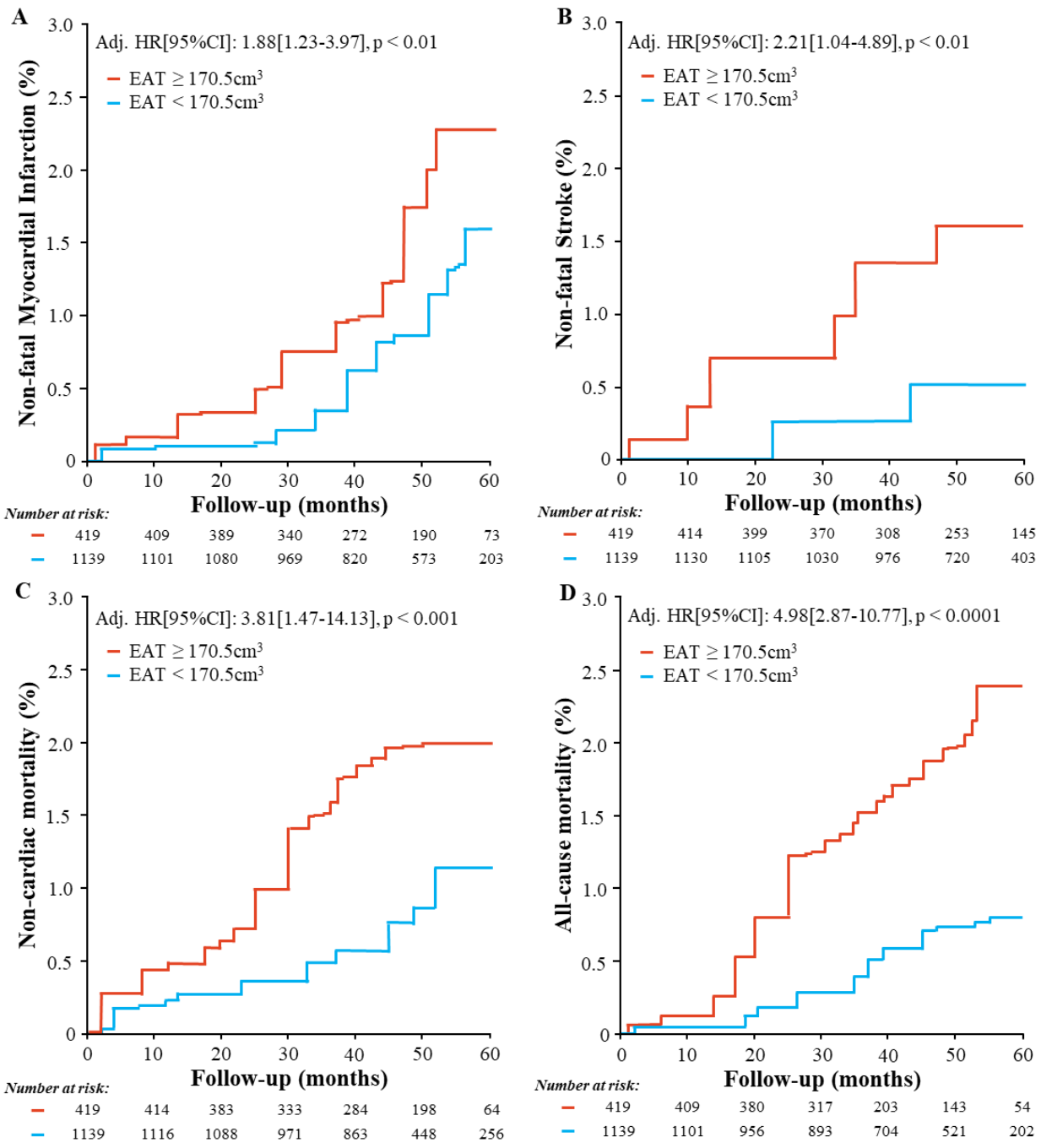


Figure 5.2. High EAT volume increases risk of major adverse events when assessed with a single cut-point. *Utilising a single cut-point for high-risk patients in SCOT-HEART (high risk = EAT ≥170.5cm3), Kaplan-Meier curves for (A) non-fatal myocardial infarction, (B) non-fatal stroke, (C) non-cardiac mortality, and (D) all-cause mortality, are demonstrated. All HR are adjusted for age, sex, BMI, hypertension, hypercholesterolaemia, diabetes mellitus, atrial fibrillation, smoking status and coronary artery calcium score (log-transformed). CI: confidence interval; EAT: epicardial adipose tissue; HR: hazard ratio.* ***Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.***

5.3.4. EAT volume and post-operative atrial fibrillation risk

Utilising 250 scans from patients in the AdipoRedOx Study, the longitudinal associations between EAT volume and in-patient post-operative AF (>30 seconds of AF on monitoring) and long term AF (paroxysmal, persistent, or chronic) following surgery were investigated.

Again, a weighted Youden's J index analysis was used to generate a single value to dichotomise the AdipoRedOx study population into high versus low EAT volume groups based on an optimum cut-off of 198.7cm^3 for the two primary outcomes of in-hospital post-operative AF and long term post-operative AF. To demonstrate the rationale for the differing EAT volume cut-offs in the primary prevention population (SCOT-HEART; as used in Figure 5.2) for prediction of cardiovascular events vs the heavily diseased CABG population for prediction of post-operative AF (AdipoRedOx) the distribution of the cohorts was plotted to capture the biological variability of visceral adiposity in these very different population samples (Figure 5.3). To be clear as to how the two populations were significantly different, the median EAT volume (IQR) in the SCOT-HEART Trial population was 130.1cm^3 (94.2-171.6) while in the severe atherosclerosis CABG population from the AdipoRedOx Study the median EAT volume was 161.9cm^3 (103.3-223.3). In practice, the SCOT-HEART EAT volume cut-off identifies (correctly) nearly all CABG patients as high-risk for all CVD outcomes (because they have confirmed CAD) and identifies very few "low risk" patients to allow for any meaningful comparison. Since the cut-off for prediction of cardiovascular events identifies most CABG patients as high risk, a new cut-off is needed to identify (within the high-risk population) the CABG patients who are at higher risk for post-operative AF compared to their peers.

Similarly, the Youden's J index cut-point of EAT volume at 198.7cm^3 (that identifies patients at risk for post-operative AF in the CABG population) identifies nearly all SCOT HEART Trial patients as low risk (correctly) for post-operative AF, as that population did not have CABG surgery.

Utilising the Youden's J index cut-point for the AdipoRedOx cohort, there were significantly different adjusted hazard ratios for the low EAT group (<198.7cm^3) versus the high EAT group (≥198.7cm^3) for both outcomes. All analyses included multivariable adjustment for relevant risk factors including LA volume and NT-proBNP.

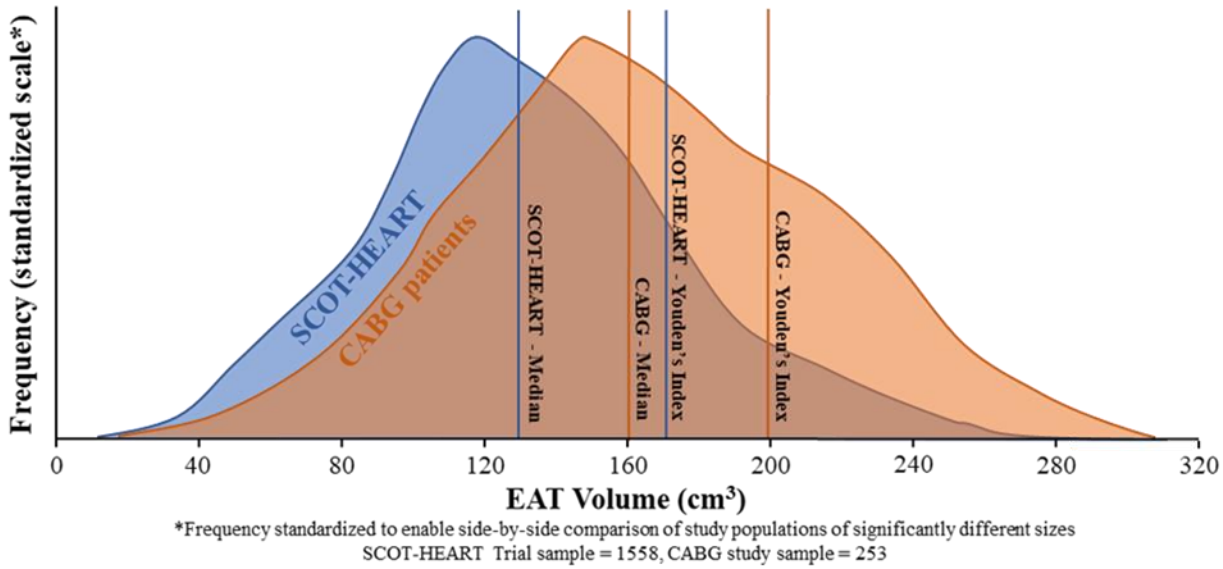


Figure 5.3. Overlayed distribution of SCOT-HEART and the Coronary Artery Bypass Surgery population in the AdipoRedOx study by EAT volume. *The SCOT-HEART (blue) and CABG (orange) population are shown in a smoothed distribution plot for the frequency of each EAT volume. The median values and Youden's J Index values for each cohort are demonstrated on vertical lines. The Youden's value for SCOT-HEART is trained on the longitudinal outcomes of all-cause mortality, MI and stroke, and the CABG population Youden's value is trained on in-patient post-operative AF risk and long term post-operative AF risk. EAT: epicardial adipose tissue*

There were 97 events of in-hospital post-operative AF (38.8%) and 48 cases (19%) of new-onset AF in nation-wide NHS Digital data for the AdipoRedOx cohort. The Kaplan-Meier curve for in-hospital post-operative AF following cardiac surgery for high (≥198.7cm^3) EAT volume versus low (<198.7cm^3). High EAT volumes were associated with a significantly greater risk for in-patient post-operative AF following adjustment for typical AF risk factors, with HR[95%CI] of 1.54[1.12-3.33] p<0.01, per 1 SD increase in EAT volume (Figure 5.4A). Equally, for long term new-onset AF following cardiac surgery high risk EAT volumes were associated with a significantly greater risk for long term AF following adjustment for typical AF risk factors, with HR[95%CI] of 1.34[1.10-2.87] p<0.01, per 1 SD increase in EAT volume (Figure 5.4B).

The addition of EAT volume into a clinical risk factors model, significantly improved the prediction of new-onset in-hospital AF in ROC curve analysis (Figure 5.4C) with ΔAUC of +0.108 (p<0.01) for risk factor model 1 with the addition of EAT volume, and ΔAUC of +0.221 (p<0.01) with the addition of the risk factor model plus EAT volume on top of CCTA derived LA volume alone. The same was found for new-onset long term AF (Figure 5.4D), with ΔAUC of +0.09 (p<0.01) for risk factor model 1 with the addition of EAT volume, and ΔAUC of +0.141 (p<0.01) over LA volume alone.

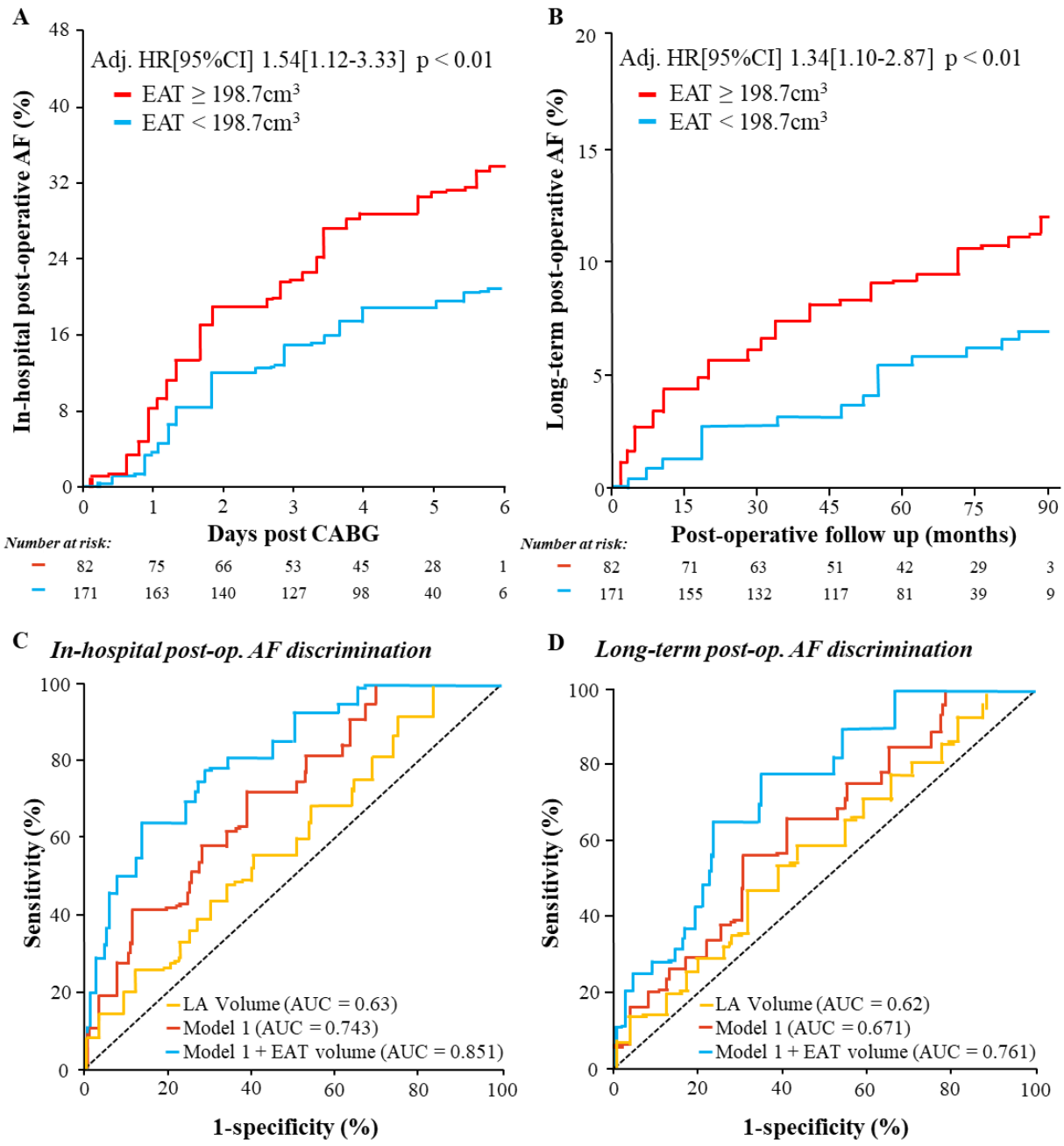


Figure 5.4. Prognostic value of EAT volume for post-operative AF. *Kaplan-Meier curve and adjusted HR for in-hospital post-operative AF (A) and long term post-operative AF (B), with sample dichotomised by Youden's J index derived cut-point of EAT volume (high risk ≥ 198.7cm3; low risk < 198.7cm3), expressed per 1 SD increase of EAT volume. Adjustment is made for age, gender, previous AF diagnosis, presence of hypertension, presence of heart failure, NT-proBNP, previous myocardial infarction, BMI, and LA volume. (C,D) Time-dependent ROC curves for discrimination of in-hospital post operative AF (C) and long term post-operative AF (D). CCTA derived LA volume (yellow) is shown alone; model 1 (red) consists of age, gender, previous AF, presence of hypertension, presence of heart failure, NT-proBNP, LA volume, previous MI & BMI. The addition of EAT volume into model 1 is demonstrated (blue). AF: atrial fibrillation; AUC: area under curve; BMI: body mass index; CI: confidence interval; EAT: epicardial adipose tissue; HR: hazard ratio; LA: left atrium; MI: myocardial infarction; RFs: risk factors.* ***Adapted with permission from West HW et al. JACC: Cardiovascular Imaging (2022) Accepted November 17 2022.***

5.3. Discussion

In this chapter I applied the deep-learning model for automated segmentation and quantification of EAT from clinical CCTA images for the investigation of the role of EAT in cardiovascular disease outcomes. The automated model was applied to the SCOT-HEART Trial CCTA cohort, demonstrating a clinically significant prognostic value of EAT volume for all-cause mortality and cardiovascular events, as a possible measure of unhealthy visceral obesity relevant to cardiometabolic dysfunction, regardless of whether EAT volume was used as a continuous variable or when used with a cut-off. Contrary to peri-coronary FAI, that captures the degree of coronary artery inflammation, and is predictive of cardiac (but not of non-cardiac) mortality, it is now demonstrates that EAT volume is predictive of non-cardiac mortality, confirming its role as a broader biomarker of visceral obesity, that affects survival in a broader sense. It is also demonstrated that this measurement has important prognostic value for post-operative AF in patients undergoing cardiac surgery, beyond known post-operative risk models including LA volume and NT-proBNP. Fully automated measurement of EAT volume incorporated into routine interpretation of CCTA promises to significantly improve the risk stratification of patients, across several important clinical outcomes.

EAT volume has previously been found to be associated with CVD metrics and outcomes including the presence of atherosclerosis [191] as well as coronary calcification progression[192]. EAT is a source of numerous pro-inflammatory mediators that circulate well beyond the microcirculation of the heart to exert paracrine and endocrine effects on the cardiovascular and endocrine systems[120]. We found EAT volume to be a significant predictor of all-cause mortality even with the exclusion of cardiac deaths. This suggests that the EAT may play a clinically significant role in broader metabolic diseases beyond atherosclerotic CAD. We propose that EAT volume should be treated as the gold standard for the detection of metabolically unhealthy visceral obesity and could form part of routine clinical interpretation of CCTA. This would shift the focus of CCTA examination as a purely structural assessment of the coronary arteries towards a more universal assessment of cardiovascular risk that considers a key visceral, metabolically sensitive, tissue depot.

EAT is in continuous bi-directional communication with the cardiovascular system[120]. When there is discordance between adipose tissue and the cardiovascular system the former is thought to shift function and exerts detrimental effects on the vessels and the heart muscle[193], which may pre-dispose the patient to adverse outcomes such as those which we investigated. We found that EAT volume is predictive of non-fatal MI and stroke independent of BMI and following adjustment for other relevant disease risk factors.

It is demonstrated that EAT volume is an independent predictor providing incremental value for post-operative AF regardless of patient BMI, LA volume, NT-pro-BNP, and other AF risk factors, indicating an important role for EAT in driving post-operative arrhythmogenesis. It is proposed that pro-inflammatory cytokines diffuse locally from dysfunctional EAT into the myocardium and contribute towards atrial myopathy[194], which drives the risk of AF.

Limitations of this chapter include that detailed mortality data were unavailable within the SCOT-HEART Trial population to investigate the exact causes of non-cardiac mortality that could be driving the finding of elevated risk of all-cause mortality conveyed by EAT volume. Other challenging clinical CCTA populations exist, such as those with congenital cardiac conditions, for which we lacked enough cases for testing of the deep-learning algorithm. Short-term AF risk data following cardiac surgery relied on inpatient monitoring only, without device monitoring following discharge.

5.4. Conclusion

In this chapter, I present the clinical implications of a new deep-learning model that allows accurate, reproducible, and rapid quantification of EAT volume on routine CCTA scans. I demonstrate that assessment of EAT volume can improve risk assessment for cardiovascular and non-cardiovascular outcomes independently of other risk factors including BMI. EAT volume is predictive of all-cause mortality and non-cardiac mortality, non-fatal MI and non-fatal stroke. EAT volume also conveys significant independent risk of post-cardiac surgery atrial fibrillation. Incorporating automated EAT volume quantification in CCTA reading protocols could improve global cardio-metabolic risk assessment and treatment planning for patients who undergo CCTA investigation, independently of the presence of coronary atherosclerosis.

6

Chapter 6. ORFAN Arm 4 Study descriptive results & stroke outcomes

6.1. Introduction

The ORFAN Arm 4 Study has become the largest observational cohort of patients who have undergone CCTA. The purpose of this chapter is to present descriptive results from the first tranche of clinical data extracted from participating sites in the Study, and to present the stroke event outcomes that are available to date for that cohort.

For the full context, aims, and methodology of the ORFAN Arm 4 Study, refer to Chapter 3 of this thesis. This Chapter is not an attempt to address the overarching aims of the ORFAN Study as outlined in Chapter 3.

It is important to consider the population that makes up the ORFAN Arm 4 Study. These are patients who were referred for CCTA imaging for clinical purposes. Most of the patients within the cohort should be regarded as intermediate risk for CAD given that the most common indication for CCTA is stable chest pain. It has previously been demonstrated in small cohorts that the presence of obstructive CAD on CCTA is associated with an increased risk of ischaemic stroke[195], however, there is no direct evidence that those referred for CCTA are at increased risk of ischaemic stroke regardless of the scan findings.

The outcome in focus within this chapter is stroke, as this is the outcome that is explored specifically via a novel risk prediction tool in the subsequent chapter of this thesis.

The scientific aim of this thesis that is achieved in this chapter is the aim to develop a major clinical CCTA cohort study for the development and testing of novel CT technologies such as radiomics based stroke risk assessment. This aim was first addressed in Chapter 3, with the methods of the ORFAN Arm 4 Study, and overall cohort descriptive results are presented herein.

6.2. Methods

The ethics, structure, and data acquisition methodology of the ORFAN Arm 4 Study are outlined in detail within Chapter 3 of this thesis and are not repeated here. The site-specific acquisition methods and statistical analysis methods for the data included within this chapter are outlined below.

6.2.1. ORFAN Arm 4 Sites

The first tranche of data received as part of the ORFAN Arm 4 Study came from eight NHS Trusts in England, UK. These were the first Trusts to join the study, and hence have had the longest period of time to collate the required patient lists, collect their local data, submit their patient lists to the third-party data holders, and transfer their patient's CCTA scans to the University of Oxford ORFAN Core Lab.

Table 6.1 presents the eight ORFAN Arm 4 sites whose data are included in this chapter, along with the period of inclusion for their first collection of clinical CCTA patients.

Table 6.1. ORFAN Arm 4 Sites included in first tranche of data

	Included ORFAN Arm 4 site	*Data period*
	Oxford University Hospital NHS Foundation Trust	01/01/2010 to 28/02/2022
	University Hospitals of Leicester NHS Trust	01/01/2010 to 31/01/2021
	Milton Keynes University Hospital NHS Foundation Trust	01/01/2010 to 31/05/2021
	Royal Brompton and Harefield NHS Foundation Trust	01/01/2010 to 30/04/2021
	Royal United Hospitals Bath NHS Foundation Trust	01/01/2010 to 28/02/2021
	Royal Papworth Hospital NHS Trust	01/01/2010 to 30/09/2021
	Guy's and St Thomas' NHS Foundation Trust	01/01/2010 to 30/04/2021
	Bart's Health NHS Trust	01/01/2010 to 30/09/2021

NHS: National Health Service

6.2.2. Data sources

The data presented here includes basic demographic and clinical risk factors which are sourced from the local sites and from NHS Digital Hospital Episode Statistics (HES) databases including admitted patient care, emergency care and outpatient care databases. Outcomes reported here come from NHS Digital (same databases) and the Sentinal Stroke National Audit Program (SSNAP) registry. All data variables accessed through these sources are listed along with definitions in Appendix 5.

6.2.3. Data availability

The data presented within this chapter are ORFAN Arm 4 Study data received at the Oxford ORFAN Core Lab up to the final week of November 2022. As the Study is ongoing in nature and data are still being gathered for transmission to the Core Lab the analysis presented here is preliminary in nature and will be updated with complete data once it is in hand throughout 2023.

This is particularly the case for clinical risk factor data from the time of the CCTA scan which rendered each patient eligible for inclusion in the Study. These data are collected by each participating site locally, and this process is time consuming. Where data was unavailable this is reported within the results.

6.2.4. Data assembly

Each site transmitted their local study patient list, with local CRF if completed, to the ORFAN Core Lab. Each list was then reviewed, and data was manually cleaned for eventual merger with data from all sites. Formats of date were unified, and all dates were standardised, and quality checks were performed. Some variables required transformation for ease of analysis and results reporting.

Age

Where full date of birth was available these were transformed into age in years. Some Bart's NHS Trust patients only had month and year of birth available, and the date of birth was standardised as the 15th day of the month for these patients before being transformed into age in years.

Ethnic category

For ease of presentation the detailed NHS Digital ethnic categories have been reduced to the four high level categorisations of White, Black, Asian and mixed, as in standard with Office of National Statistics reporting[196]. White includes British, Irish, Northern Irish, Scottish, Welsh, Cornish, Cypriot, Greek, Turkish, Italian, Irish Traveller, Gypsy/Romany, Polish, ex-USSR, Kosovan, Albanian, Bosnian, Croatian, Serbian, ex-Yugoslav, mixed white, white unspecific and any other white identifying background. Western and Northern European ethnic groups such as, but not limited to, French, German, Spanish, and Swedish which are not specifically mentioned are included in this definition of white ethnicity. Asian ethnicity includes all who identify as Indian, Pakistani, Bangladeshi, mixed Asian, Sri Lankan, Asian African, Tamil, Sinhalese, Chinese, Japanese, Filipino, Malaysian, Thai, Indonesian, Korean, and any other Asian or Central Asian ethnic group including Asian unspecified. Black ethnicity included any participant who identified as British Black, Black African, Black Caribbean, including Somali, mixed black, African American, and black unspecified. Mixed ethnicity includes those who identified as a mixed ethnicity, which includes mixed white and Black Caribbean, mixed white and black African, mixed white and Asian, and mixed Asian and Black (with all possible permeation). Patients who identified as 'other non-specified' were included as missing.

Risk factors

Risk factor data from the time of the CCTA is sourced from local site CRF in the first instance. If not available from the CRF at the time of writing the risk factors were sourced from NHS Digital HES data.

For both the CRF and NHS Digital data the diagnostic risk factors including hypertension, hypercholesterolaemia, diabetes mellitus, peripheral vascular disease, angina, acute myocardial infarction, chronic coronary syndrome, and atrial fibrillation were defined according to ICD-10 coding conventions. See Appendix 7 for the full table of ICD-10 codes extracted for each condition.

Smoking status, family history of heart disease, and other behavioural risk factor data were only available from the local CRF and was predominantly incomplete at the time of writing, so is not included in analysis.

Outcomes data

Stroke outcomes were sources from the SSNAP registry and NHS Digital HES data, with data collected from 2005 through 2021 inclusive. HES data was extracted and merged from the admitted patient care database, the emergency care database, and the outpatient database. Duplicated records were removed. SSNAP data relies upon NHS Trust self-report of stroke events, both ischaemic and haemorrhagic, as adjudicated by local stroke physicians, emergency physicians, and neurologists. The reports are made at the conclusion of a patient's hospital admission, either following discharge, transfer from the initial treatment site to another healthcare facility, or at time of death. NHS Digital HES databases were also searched for stroke events according to ICD-10 codes (see Appendix 7 for groupings of stroke codes). Where a patient was found to have had an acute stroke event in both or a single database, it was included in analysis.

There were exclusions from the SSNAP database. 98 participants with stroke events were excluded prior to analysis as they had been withdrawn from the study due to the UK national data opt-out policy and all data related to these participants was deleted. A further 127 participants with stroke events recorded in the SSNAP database were excluded from analysis as the SSNAP data did not include an admission or discharge date for the event, some of these patients were included through NHS Digital data that included admission dates for the same patients.

6.2.5. Statistical analysis

Descriptive data are reported as median (IQR) for continuous variables or number (%) for categorical variables. When shown, p values are for the comparisons between groups and are derived from the Mann-Whitney *U* test (continuous variables) or Pearson's χ^2 test (categorical variables). P values for trend are derived from Mann-Whitney *U* test (a Wilcoxon-type test for rank[197] performed with the *nptrend* action in Stata software). Where p-values for trend are shown they capture the full timescale show on the figure, unless specified in the text. Some denominators shown may differ because of missing data, however this will be reported in the text or table footnotes.

Descriptive statistical analysis was conducted in Microsoft Excel version 2211 (Microsoft, Redmond, USA) or in Stata version 15 (StataCorp, College Station, TX, USA).

6.3. Results

6.3.1. CCTA patients at each site

In total, 88223 unique patients who underwent clinical CCTA across the eight NHS Trust sites were included in this analysis. Rates of clinical CCTA significantly increased year on year, with marked reduction in 2020 due to the global COVID-19 pandemic.

Figure 6.1 demonstrates the overall number of clinical CCTA scans included in the study for each full year of data available. Unique patients who had at least one CCTA are included. The striking growth in CCTA scans across the preceding decade is obvious, with a p for trend of <0.001. Table 6.2 displays the annual number of clinical CCTA patients across each site that was included within the ORFAN Arm 4 Study first tranche of data collection. Between 2010 and 2020 there was a mean increase in CCTA scans performed across all sites of 19.5% per year. Between 2019 and 2020 there was a 23.19% decrease in CCTA scans, attributable to the COVID-19 pandemic impacting on access to radiological services. The full recovery of services following the most severe period of the pandemic (2020 and 2021) is not shown in these data as full data to the end of 2022 has not yet been collected across all sites.

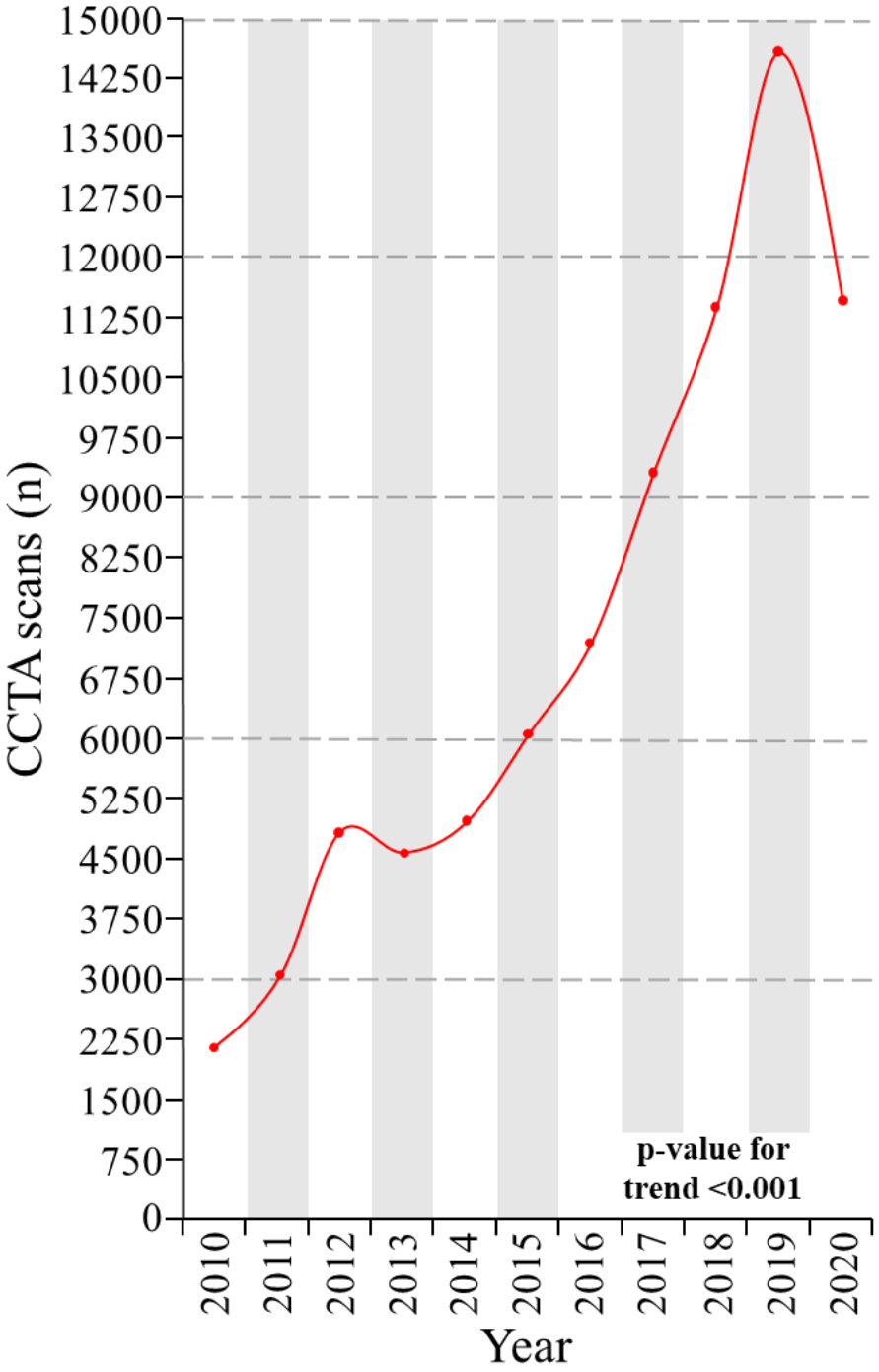


Figure 6.1. Clinical CCTA scans per year in the ORFAN Arm 4 Study. Each full year of data is included for unique patients who underwent at least one clinically indicated CCTA across the eight NHS Trust's included in the ORFAN Arm 4 Study in England, UK. CCTA: coronary computed tomography angiography

Table 6.2. ORFAN Arm 4 Study CCTA patients by site and year

	2010	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021***	**2022†**	**Total by site:**
Bath	29	55	142	109	134	99	98	254	556	1027	722	997	188	4410
Papworth‡	708	422	929	223	NA	NA	NA	NA	NA	2522	1715	1800	NA	8319
Guy's & St Thomas'	122	348	632	711	854	1010	1128	1637	2000	1553	1188	298	NA	11481
Leicester	94	90	104	104	561	831	799	1183	1278	1708	1566	80	NA	8398
Bart's	724	1218	1901	1984	2076	2498	3309	4203	4915	5246	3686	3103	NA	34863
Royal Brompton & Harefield	548	878	972	1129	1124	1322	1287	1288	1516	1527	1153	310	NA	13054
Oxford	0	5	150	262	370	402	535	633	891	1031	1103	1247	239	6868
Milton Keynes	0	0	10	4	44	56	92	110	133	111	207	63	NA	830
Total by year:	2225	3016	4840	4526	5163	6218	7248	9308	11289	14725	11340	7898	427	88223

*Incomplete annual data for 2021 presented for Papworth (date through end of September 2021), Guy's & St Thomas' (data through end of April 2021), Leicester (data thorough end of January 2021, Bart's (data through end of September 2021), Royal Brompton and Harefield (data through end of April 2021) and Milton Keynes (data through end of May 2021)
†Incomplete annual data for 2022 presented for Bath and Oxford (data through end of February 2022 for both).
‡Papworth did not include patients for the years 2014 through 2018. NA: Not available

6.3.2. Age at time of CCTA

The median(IQR) age of patients who underwent clinical CCTA was 57.8(49.1-67.6) years across the whole cohort. Figure 6.2 demonstrates the age histogram of the whole sample. No patients were missing age data from local site databases. NHS Digital age data was not included in this age only analysis as actual date of birth was not provided for the majority of patients from this source.

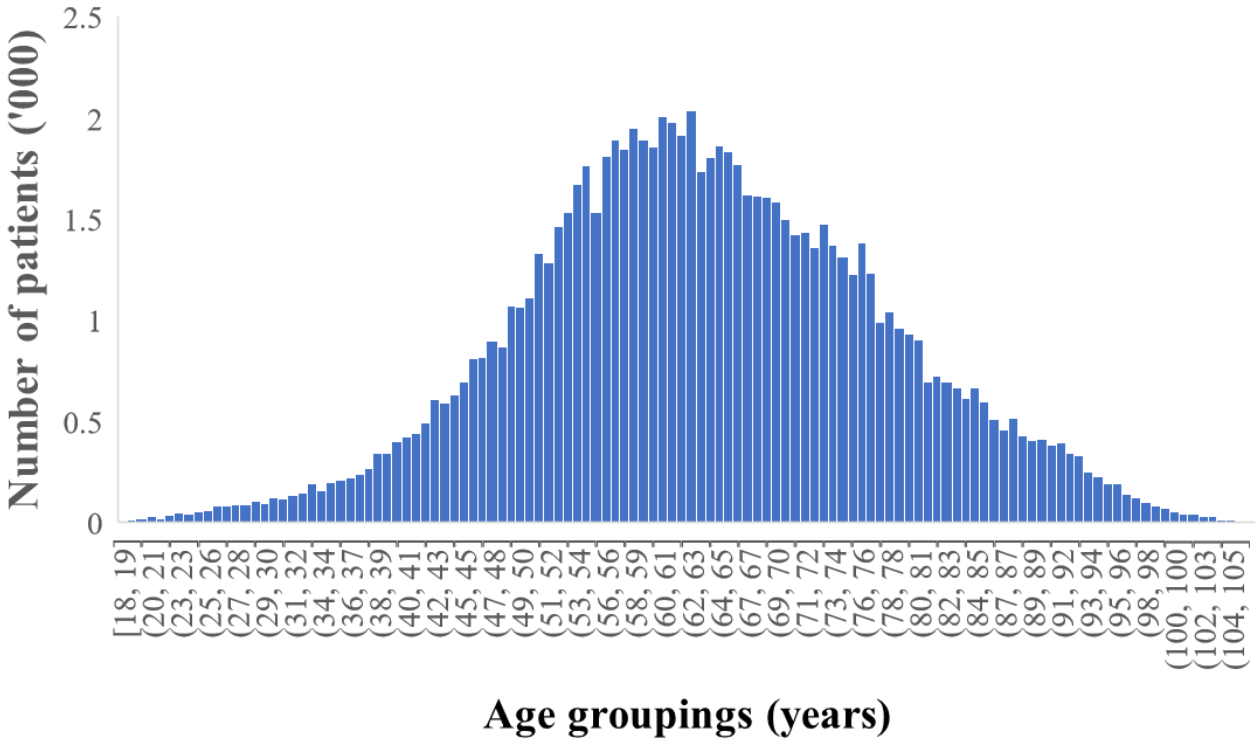


Figure 6.2. Age histogram of patients at time of first clinically indicated CCTA. *Age histogram of the entire cohort of 88223 patients. The age of clinical CCTA patients follows a normal distribution, with a slight right shift.*

Across the eight NHS Trusts who contributed patients the median age was lowest at Bart's (56.0 years) and The Royal Brompton and Harefield had the oldest median age (62.3 years). Table 6.3 lists each Trust with the median age and range for their clinical CCTA patients.

Table 6.3. Age of clinical CCTA patients by NHS Trust

Trust	Median age	IQR
Bath	60.7	52.2-69.9
Papworth‡	59.8	50.9-69.3
Guy's & St Thomas'	56.3	48.0-65.5
Leicester	57.2	49.3-65.8
Bart's	56.0	47.8-64.9
Royal Brompton & Harefield	62.3	51.2-74.2
Oxford	61.0	51.3-71.9
Milton Keynes	57.7	47.9-69.3
Whole cohort	*57.8*	*49.1-67.6*

The median age at time of clinically indicated CCTA per year is shown in Figure 6.3. All patients are shown, along with the break down for males and females. The median age for men was 57.0 (47.8-67.1) years and for women 58.7 (50.3-68.1) years.

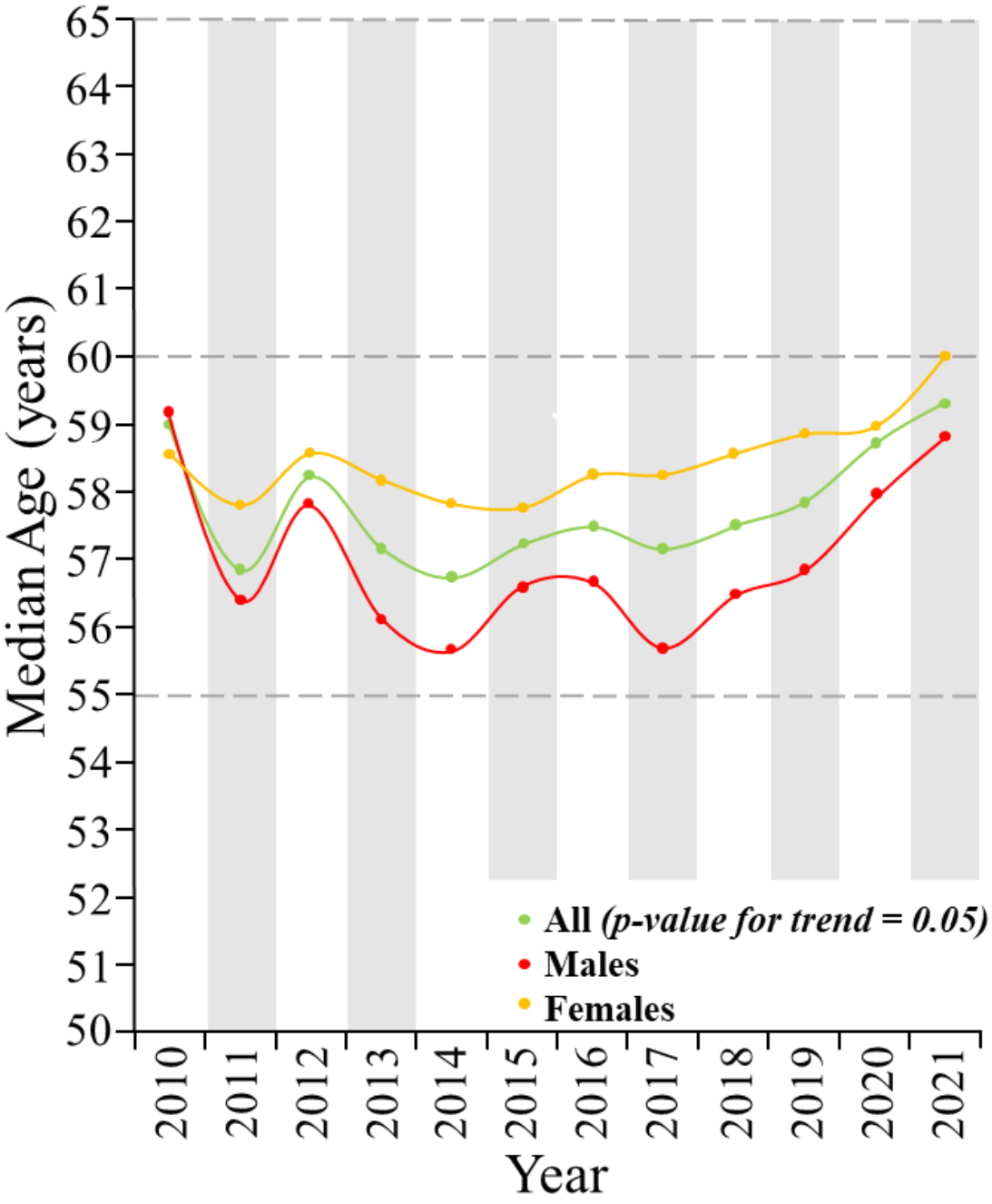


Figure 6.3. Median age at time of clinical CCTA within the ORFAN Arm 4 Study. *The median age for all patients referred for clinical CCTA is shown (green). The median age of males (red) and females (orange) is also displayed for each full year of data.*

6.3.3. Gender

Gender was missing for 34 patients (0.04%). All sites had the ability to input gender locally as male, female, transgender or intersex. No sites recorded any transgender or intersex patients. Similarly, NHS Digital data did not record any transgender or intersex patients for the cohort.

There were conflicting gender data between sources for 325 patients (0.36%). 164 patients were recorded in local site databases as being female while NHS Digital HES databases recorded them as being male. 161 patients were recorded in local site databases as being male while NHS Digital HES databases recorded them as being female. For all such contradictions the local site database was considered to be the most accurate source of data and was the value that was retained in analysis.

There were 46,690 (52.1%) men and 41,845 (46.7%) women in the sample from across the eight NHS Trusts.

The gender split across all eight Trusts is demonstrated in Table 6.4. Leicester was the only Trust with a majority female cohort of clinical CCTA patients, while The Royal Brompton and Harefield had the reported the lowest proportion of females, at 42.8%.

Table 6.4. Gender split of ORFAN Arm 4 CCTA patients across NHS Trusts

	Male	***% Male***	***Female***	***% Female***	***Missing***	***% Missing***
Bath	*2235*	*50.7*	*2175*	*49.3*	*0*	*0*
Papworth	4505	54.2	3814	45.8	0	0
Guy's & Thomas'	5925	51.6	5556	48.4	0	0
Leicester	3863	46.0	4508	53.7	27	0.3
Bart's	18064	51.8	16797	48.2	2	0.006
Royal Brompton & Harefield	7466	57.2	5588	42.8	0	0
Oxford	3820	55.6	3043	44.3	5	0.07
Milton Keynes	428	51.6	402	48.4	0	0
Total by gender	*46306*	*52.5*	*41883*	*47.5*	*34*	*0.04*

6.3.4. Ethnicity

The majority of patients identified as white, in keeping with the reported ethnicity breakdown of the whole population of England. The breakdown of ethnicity by each site is included in Table 6.5, with England's ethnicity data sourced from the Office for National Statistics included for comparison purposes. 5.1% of patients had missing ethnicity data from NHS Digital and from the ORFAN local site CRF.

Table 6.5. High level ethnic categories of the ORFAN Arm 4 CCTA patient cohort across NHS Trusts

	White	% White	Black	% Black	Asian	% Asian	Mixed	% Mixed	Missing	% Missing
Bath	3862	87.6	25	0.6	82	1.86	51	1.16	390	8.8
Papworth	7102	85.4	109	1.3	328	3.94	223	2.7	557	6.7
Guy's & Thomas'	7386	64.3	1578	13.7	1050	9.1	354	3.1	1113	9.7
Leicester	5809	69.2	203	2.4	1946	23.2	187	2.2	253	3.0
Bart's	16612	47.6	4055	11.6	9481	27.2	4708	13.5	7	0.02
Royal Brompton & Harefield	9888	75.7	422	3.3	1262	9.7	138	1.1	1344	10.3
Oxford	5533	80.6	78	1.1	319	4.6	111	1.6	827	12.0
Milton Keynes	672	81.0	38	4.6	66	8.0	16	1.9	38	4.6
Total	*56864*	*63.5*	*6508*	*7.3*	*14534*	*16.2*	*5788*	*6.5*	*4529*	*5.1*
*England ONS data**	*NA*	*81.7*	*NA*	*4.0*	*NA*	*9.3*	*NA*	*2.9*	*NA*	*NA*

**Whole of England percentages for high-level ethnic categories from the Office of National Statistics – Census 2021*[198]*

6.3.5. Cardiovascular disease risk factors

Cardiovascular disease risk factors were sources from local site databases and NHS Digital HES data for the entire cohort. Overall number of unique diagnoses for each cardiovascular disease risk factor are presented in Table 6.6, with breakdown of initial diagnosis prior to or after the patient's first CCTA.

The risk factors listed are cardiovascular diseases in their own right and constitute diagnostic risk factors relevant to cardiovascular disease outcomes such as CAD events and stroke.

Table 6.6. ORFAN Arm 4 cohort cardiovascular disease risk factors

Diagnosis*	n (%)	n (%) diagnosed prior to first CCTA	n (%) diagnosed after first CCTA
Hypertension	42392 (48.1)	31609 (35.8)	10996 (12.5)
Hypercholesterolaemia	18056 (20.5)	13196 (15.0)	4997 (5.7)
Diabetes mellitus	13532 (15.3)	10765 (12.2)	2767 (3.1)
Heart failure	8139 (9.2)	5568 (6.3)	2571 (2.9)
Atrial fibrillation	11002 (12.5)	8005 (9.1)	2997 (3.4)
Peripheral vascular disease	8525 (9.7)	5858 (6.6)	3667 (4.2)
Chronic coronary syndrome	18513 (21.0)	11825 (13.4)	6688 (7.6)

*See Appendix 7 for ICD-10 code groupings used for each condition

The accumulation of cardiovascular disease risk factors over time is demonstrated in Figure 6.4. Over the 16 years of complete data collection from NHS Digital and the local site CRF there were escalating diagnoses for each condition across the period. The impact of the global COVID-19 pandemic in 2020 is evident, with a sharp decline in all conditions diagnosed in 2020. The p-value for trend was significant for all risk factors in analysis until 2019 (all $p<0.05$), however trend analysis that included 2020 was only significant for hypertension (p-value for trend = 0.005), heart failure (p-value for trend = 0.01) and peripheral vascular disease (p-value for trend = 0.001).

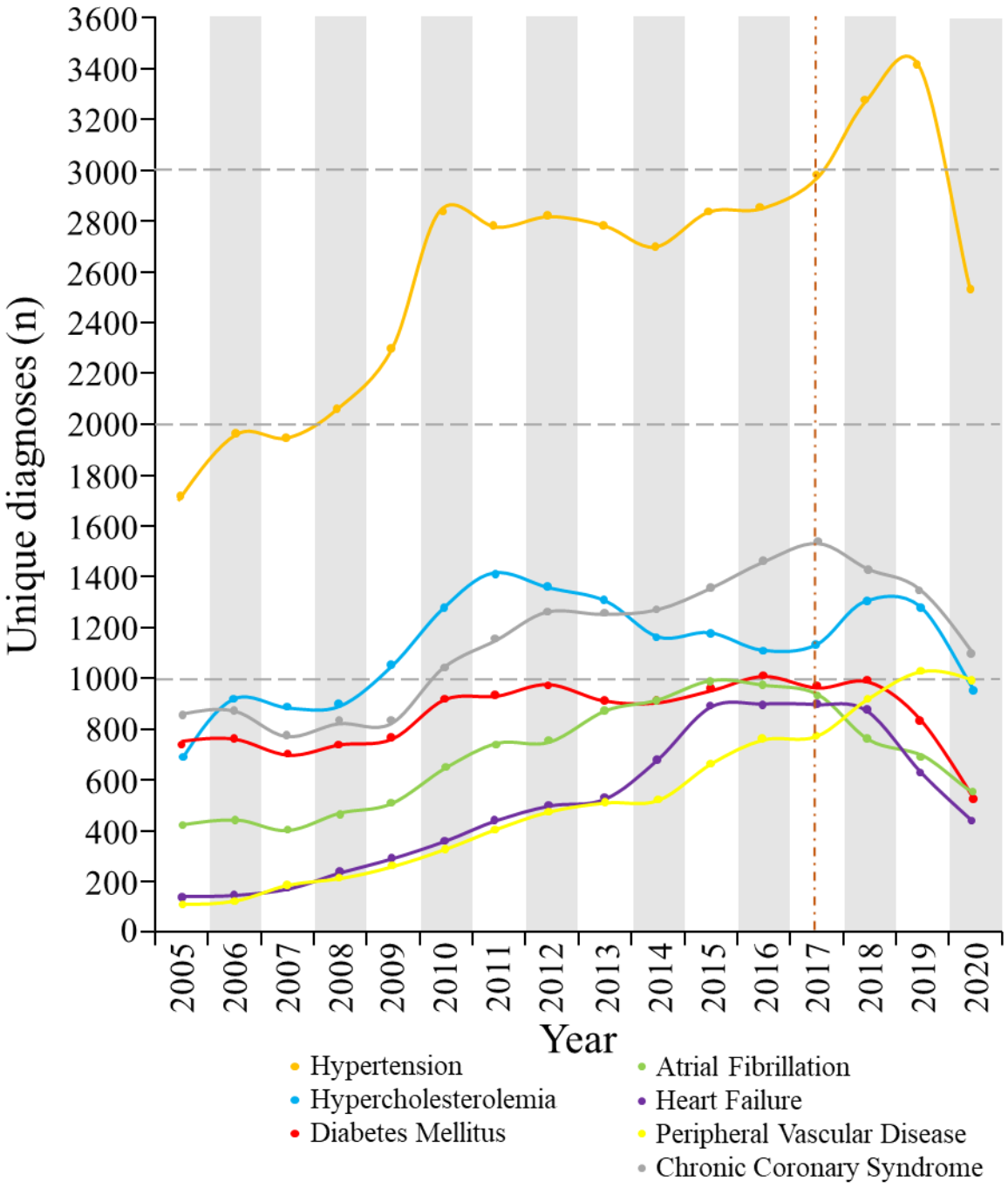


Figure 6.4. **Diagnosis of cardiovascular disease risk factors each year within the ORFAN Arm 4 Study.** Each full year of local site database and NHS Digital data collection is recorded from 2005 through 2020. The total number of unique first-time diagnoses from across the cohort are recorded. Trend lines are smoothed and do not represent true values beyond the fiducial marks for each year. The ochre dashed line represents the mean date of CCTA scan for the entire sample (June 2017). Definitions of each condition are listed in Appendix 7.

6.3.6. Stroke outcomes

Concordance between datasets & merging of data

There was non-concordance between NHS Digital HES data and SSNAP data for stroke events.

SSNAP data showed 1659 unique participants as having at least one stroke event either prior to or after the CCTA that rendered the patient eligible to be in the ORFAN Arm 4 Study. Of these, 1527 were ischaemic events and 132 were haemorrhagic.

NHS Digital HES data revealed 1315 unique participants as having at least one stroke event either prior to or after the CCTA that rendered the patient eligible to be in the ORFAN Arm 4 Study. Of these, 1167 were ischaemic events and 148 were haemorrhagic.

It is important to note that repeat stroke events are not included in these numbers, only unique patients with at least one stroke event in each database are reported.

Of the 1527 patients reported to have had at least one ischaemic stroke in the SSNAP database 794 (52%) were also included in the NHS Digital HES data as having an ischaemic stroke with the dates of the event matching.

Of the 148 patients reported to have had at least one haemorrhagic stroke in the NHS Digital database 42 (28%) were also included in the SSNAP database as having an haemorrhagic stroke with the dates of the event matching.

Data were merged to include all events from both SSNAP and NHS Digital data sources, with events that only appeared in a single database (SSNAP or NHS Digital) retained for analysis.

Stroke

There were 2228 unique participants (2.5% of the overall cohort) recorded as having stroke events (both haemorrhagic and ischaemic) in the merged NHS Digital and SSNAP database either prior to and/or after the CCTA that rendered the patient eligible to be in the Study.

1990 patients had at least one ischaemic stroke event, and 238 had at least one haemorrhagic stroke event. Three patients had at least one ischaemic and haemorrhagic stroke event.

138 participants had two separate stroke events, 23 participants had three separate stroke events, eight participants had four unique stroke events, and two participants had five unique stroke events. Figure 6.5 demonstrates the ischaemic and haemorrhage events that were recorded prior to and following the CCTA.

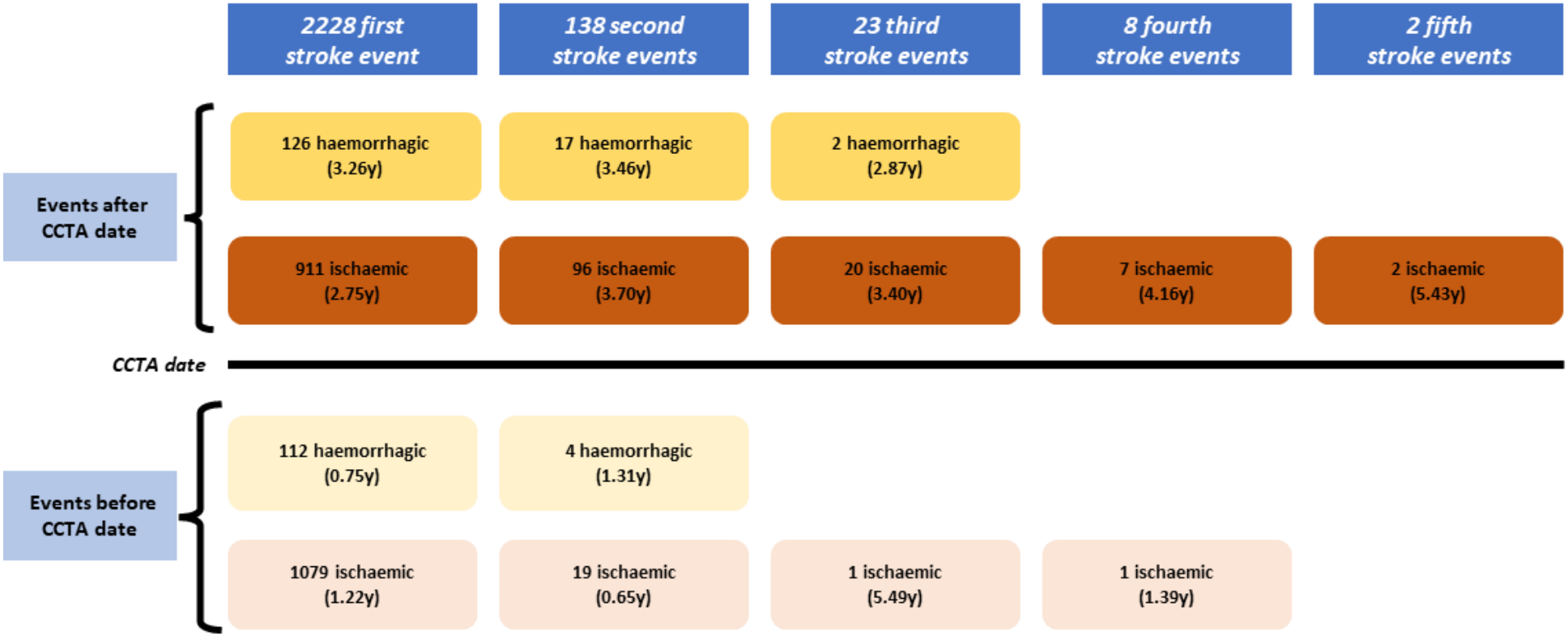


Figure 6.5. Stroke events prior to and following the clinical CCTA within the ORFAN Arm 4 cohort. *There were 2228 initial stroke events, with the breakdown of haemorrhagic and ischaemic events both after CCTA (above the middle line) and prior to the CCTA (below of middle line) reported. The median time to event is reported in years. For events prior to the CCTA this is the median time prior to the CCTA that the event occurred. Numbers of patients who had multiple subsequent stroke events are reported.*

Of the 2228 participants who had at least one stroke event recorded, 1054 (47.3%) of the events occurred following the CCTA scan that rendered the patient eligible for inclusion in the ORFAN Arm 4 cohort, with 1174 (52.7%) occurring prior to the CCTA. 911 first time unique ischaemic stroke events were recorded following the CCTA, while 1079 were recorded as having occurred prior to the CCTA.

As is presented in Figure 6.5, second stroke events occurred in 138 unique participants, with 96 ischaemic strokes and 17 haemorrhagic events occurring in patients who had already had an initial event following the CCTA. 19 participants had experienced two ischaemic strokes prior to the CCTA, and four patients had experienced two haemorrhagic strokes prior to the CCTA. A small group of patients had further events; 20 patients experienced three ischaemic events following the CCTA, with two experiencing three haemorrhagic events following the CCTA. One patient experienced four unique ischaemic stroke events prior to the CCTA. Seven participants experienced four ischaemic events following the CCTA, with two experiencing five ischaemic stroke events following the CCTA.

The basic demographics of the entire ORFAN Arm 4 cohort, split into those who had an ischaemic stroke event following the CCTA and those who remained stroke free are presented in Table 6.7. All risk factors bar most ethnic categories were found to differ significantly between the two groups.

Table 6.7. ORFAN Arm 4 cohort demographics and risk factors at time of CCTA

		Ischaemic stroke event(s) after CCTA	***Ischaemic stroke free after CCTA***	***Significance***
Number		911	87312	-
Follow-up after CCTA (years)		6.91	3.91	-
Age at time of scan		68.0 (56.5-78.9)	57.5 (48.6-67.4)	<0.001
Male gender		341 (55.1)	44817 (50.1)	<0.001
Ethnicity				
	White	450 (72.7)	55536 (63.0)	0.04
	Black	47 (7.6)	6360 (7.2)	0.16
	Asian	68 (11.0)	14239 (16.1)	0.04
	Mixed	23 (3.7)	5672 (6.4)	0.07
	Other/Missing	31 (5.0)	5796 (6.6)	0.09
Hypertension		434 (70.1)	30789 (34.9)	<0.0001
Hypercholesterolaemia		141 (22.8)	12854 (14.5)	<0.001
Diabetes mellitus		133 (21.5)	10463 (11.8)	<0.0001
Heart failure		115 (18.6)	5363 (6.0)	<0.0001
PVD		108 (17.4)	4675 (5.3)	<0.0001
Atrial fibrillation		207 (33.4)	7675 (8.7)	<0.0001
Chronic coronary syndrome		186 (30.0)	11451 (13.0)	<0.0001
CHA_2DS_2-VASc Score*		4.1 (2.6-6.0)	1.8 (0.2-3.5)	<0.0001

Data are median(IQR) for continuous variables or number(%) for categorical variables. p values are for the comparisons between groups and are derived from the Mann-Whitney *U* test (continuous variables) and Pearson's χ^2 test (categorical variables). Risk factor data is derived from NHS Digital HES records and local site databases. Risk factors are only included if they were recorded as present prior to or at the time of the CCTA.

*Stroke risk score developed for patients with atrial fibrillation[14]; 0-9 depending on risk factors: congestive heart failure (1 point), hypertension (1 point), age ≥75 years (2 points) or age 65–74 years (1 point), diabetes mellitus (1 point), previous stroke or TIA (2 points), vascular disease defined as peripheral vascular disease, previous myocardial infarction (1 point), female sex (1 point).

CCTA: coronary computed tomography angiography; PVD: peripheral vascular disease

6.4. Discussion

In this chapter, the first descriptive analysis of the ORFAN Arm 4 cohort for the first eight NHS Trusts to enter the Study is presented. The data presented are data available up to the end of November 2022, which included 88223 unique patients who had undergone clinically indicated CCTA between 2010 and 2022, inclusive, with most data available consisting of the years 2011-2021.

I report the following descriptive findings for the cohort:

a) A significant upwards trend in the usage of CCTA in clinical practice in England from 2010 through 2019
b) A significant decline in the usage of CCTA in England during the COVID-19 pandemic
c) The median age for first use of clinical CCTA services across England is stable
d) Between 2010 and 2020, men were significantly more likely to have received a CCTA than women
e) A larger proportion of non-white identifying patients than the national average for their respective ethnic categories underwent clinically indicated CCTA
f) Patients who underwent CCTA had high levels of exposure to cardiovascular disease risk factors, with the majority exposed to the risk factor prior to the CCTA

I also report the following outcomes for ischaemic stroke from within the cohort:

a) 2.2% of the clinical CCTA cohort had experiences a stroke event prior to or following their CCTA
b) 0.7% of the cohort had an ischaemic stroke event following the clinical CCTA up to March 2022 and 1.3% had experienced an ischaemic stroke prior to the clinical CCTA
c) Multiple ischaemic stroke events were rare, with only 0.2% of the cohort having more than one ischaemic stroke following the CCTA
d) Those who did have ischaemic stroke events following the CCTA has significantly greater levels of exposure to cardiovascular disease risk factors at the time of the scan compared to those who remained stroke free
e) Stoke risk as assessed by the CHA_2DS_2-VASc score was a significant predictor of future stroke regardless of AF status

These are descriptive results that provide significant insights into the ORFAN Arm 4 Study cohort. Yet still there is a large amount of data that could be presented descriptively for this cohort which remains unreported. I have attempted to select the most pertinent data for each basic demographical variable or group of variables to reveal key insights into the cohort. For example,

future analysis will provide further breakdown of changes in ethnicity of patients referred for clinical CCTA across different regions of England.

The rate of stroke events within this population sample was found to be 2.2% across a 16-year period. This is in line with UK national stroke prevalence rates of 1.6-2.4% across the counties of England according to GP practice data[5]. Given the selected nature of this cohort as being those sent for clinically indicated CCTA, a slightly higher stroke rate could be expected. Those who suffered from an ischaemic stroke event following the CCTA were found to have been exposed to cardiovascular disease risk factors significantly more than those who remained stroke free. This was a striking finding and one that indicates the importance of holistic patient care at all clinical interactions. Many of these patients who went on to have a stroke following the clinical CCTA likely had CAD and stroke risk well beyond residual effects, and were likely receiving non-optimum medical therapy given the profound impact that their risk factors had on the future risk of disease.

The stroke data presented here represent a single outcome selected from scores of other clinical outcomes highly relevant to cardiovascular health and beyond. The power of the ORFAN Arm 4 Study as a tool for research into cardiovascular risk and disease outcomes is apparent.

A key strength of the ORFAN Arm 4 cohort is the ability to sources clinical outcomes data from multiple sources. As reported in this chapter for ischaemic stroke outcomes, the NHS Digital HES data sourced from all major HES databases was found to be significantly different to the UK national stroke registry (SSNAP). We believe it is likely that all events within each database are accurate but that each database has data collection limitations that mean many events are missed altogether or are mis-categorised. The approach of adjudicating events within the ORFAN Arm 4 study is not complete for the overall study. When complete it will ensure a more accurate representation of clinical events than simply selecting a single data source for analysis. Non-concordance between NHS Digital hospital records and event based record has been noted previously[199], however not specifically between the SSNAP database and HES stroke event data.

Other limitations relate to technical considerations. Many software systems relevant to the ORFAN Study limit gender options to male and female, this is the case for the Computerised Radiology Information System (CRIS) software (Wellbeing Software, Mansfield, UK) that all Trusts utilised to generate their initial patient lists and for the extraction of many basic demographical data variables. This restricts a broader definition of gender which should include transgender patients despite the inclusion of transgender as an option in the ORFAN Arm 4 local site databases.

Data soon to be available for analysis will include patients who underwent CCTA for the remainder of 2021 and the whole for 2022. These data will reveal the full impact of the COVID-19 pandemic on this important clinical service and any lingering impact on patient access to CCTA.

6.5. Conclusion

The ORFAN Arm 4 Study has become the largest clinical CCTA study available for research. This chapter set out to describe the basic characteristics of the cohort and demonstrate the research potential with a single outcome – that of stroke.

Early descriptive analysis has revealed a significant increase in CCTA rates across the decade from the initial piloting of CCTA in 2010 through to 2020, with a marked decline in service provision due to the COVID-19 pandemic. The most common patient to receive a CCTA across the whole cohort was a 57-year-old white male, however it is also shown that non-white identifying persons were more likely to receive CCTA when compared to national background rates of their ethnic category. As is expected, the clinical CCTA population is one highly exposed to cardiovascular disease risk factors, with growing exposure over time.

Finally, it was demonstrated the rates of stroke events within this cohort, showing that some two percent of the cohort has experienced a stroke event across the whole study period, with 619 patients (0.7%) having an ischaemic stroke event following the CCTA that rendered them eligible for the ORFAN Arm 4 Study. Given these event numbers and the data available and forthcoming for this cohort, the potential utility of this study for the development of novel disease risk scores and imaging technologies is significant.

7

Chapter 7: Pilot radiomic stroke risk algorithm

7.1. Introduction

In this chapter, I demonstrate the utility of the ORFAN Arm 4 Study cohort for the development of novel ischaemic stroke risk assessment tools. The ability to combine transcriptomic data from myocardial and EAT samples with CT radiomics for the detection of the atrial myopathy could revolutionise the assessment of AF and stroke risk and pave the way for non-invasive detection of disease processes previously only assessable through invasive tissue biopsy. Peri-atrial radiotranscriptomics may be able to detect inflammation, peri-atrial lipolysis, oedema, angiogenesis and fibrosis of the atrial wall causing respective peri-atrial changes including EAT. Such tools could be used for the assessment of disease risk for conditions directly related to the atrial myopathy such as AF and stroke, as is the case in this study, and other conditions such as heart failure and the risk of relapse following invasive catheter ablation therapy for AF. Radiotranscriptomic assessment of the peri-LA region may also be used to help detect responders to treatments for the prevention of AF and stroke, such as traditional and novel anti-inflammatories.

Specifically in this chapter, I present a pilot version of the Atriomic Risk Algorithm that relies exclusively upon radiomic data only – this is a preliminary and exploratory analysis. The risk assessment tool is the application of the automated LA and peri-LA region segmentation tools developed as part of Chapter 4 of this thesis. The data presented here are preliminary in nature, as the full Atriomic Risk Algorithm will be validated with the input of transcriptomic ground truth data to constitute a true radiotranscriptomic risk signature, as opposed to an exclusively radiomic risk signature as is presented herein.

As outlined in the introduction to this thesis peri-LA EAT has a wide range of bi-directional interactions with the myocardium of the atrium. There is excellent evidence that peri-myocardial EAT is structurally influenced by inflammatory paracrine signalling from the heart myocardium. This can supress adipogenesis, activate lipolysis, increase oedema, and drive inflammatory processes, culminating in adipocytes of a reduced size and lipid-content adjacent to the diseased atrial tissue. If this inflammation persists, fibrosis and increased neoangiogenesis within the EAT and adjacent connective tissues will likely ensue. These tissue level changes around the atria are the substrate that is targeted through the radiomic assessment of the peri-LA regions. The CT texture shifts in the

peri-LA segments are what is identified through radiomics. These texture shifts are impossible to see with the human eye and their detection is only possible through data analysis of extracted radiomic features. The texture shifts are harnessed to build a radiomic signature that can quantify the extent of the atrial myopathy and be used as a marker of risk for future ischaemic stroke events – regardless of the AF status of the patient at the time of the CCTA scan.

The scientific aim of the thesis that is achieved in this chapter is to explore the role of radiomics for risk stratification for ischaemic stroke utilising LA and peri-LA segmentations from CCTA.

7.2. Methods

7.2.1. Study design

Cohort were selected from the ORFAN Arm 4 Study for inclusion in this preliminary analysis based upon the availability of CCTA data. ORFAN sites were assigned as training and internal testing (group 1) or external validation sites (group 2). The Royal Brompton and Harefield NSH Foundation Trust and The Oxford University Hospitals NHS Foundation Trust sites of the ORFAN Study were used as group 1, while The Royal Papworth Hospital NHS Foundation Trust was used as group 2.

Group 1 was split randomly into a training set (80%) and an internal testing set (20%). Group 2 was used entirely for external validation testing.

7.2.2. Outcome data

Ischaemic stroke event data was utilised from pooled records from the NHS Digital HES databases (as described in Chapter 3 of this thesis) and the SSNAP databases. These data were available for all outcomes from across the NHS from 2005 through 2021. Stroke events from NHS Digital data were defined as any event assigned ICD-10 code I63 (any subcode I63.0-I63.9), while SSNAP stroke events are defined and reported by the treating clinician at each NHS Trust where the event occurred.

Follow up was set at 5 years from the date of the CCTA scan, with all patients censored at this time point, or if a stroke had occurred <5years following the CCTA the patient was censored at that time point. If the ischaemic stroke was fatal, this event was included in the analysis. Patients who died of conditions other than ischaemic stroke were included in analysis but were censored at time of death. Other exclusions included those with unreadable images, those who had previously had a heart transplant and those with major anatomical anomalies.

7.2.3. Automated CCTA analysis

The full description of the development, internal and external validation testing of the automated CCTA LA and peri-LA segmentation volume of interest (VOI) is provided in Chapter 4 of this thesis and is not reproduced here. The VOI segmented in this analysis were as described in Chapter 4: the

LA chamber, the intra-atrial wall and surrounding tissues, the anterior LA wall and surrounding tissues, and the peri-left atrial appendage wall and surrounding tissues (see Figure 4.2). The 3D Residual-U-Net convolutional neural network analysis was applied to all CCTA scans from group 1 and group 2. Radiomic features were extracted for all segments individually, with no thresholding by Hounsfield unit values. The extraction of radiomic features was performed automatically utilising our own *RadExtractor* (see section 2.7) and allowed analysis of merged VOI. Scans that were flagged as unsuccessful segmentations (flagged through automated quality control when any VOI volume was >2 standard deviations from the mean) were removed and manually segmented, and a further random sample of 1% of all scans were manually reviewed for quality control.

7.2.4. Radiomic characterisation and machine learning modelling

In the pilot analysis presented here, we combined all peri-LA segments into a single VOI and extracted the radiomic features for the overall VOI.

To limit our analysis to radiomic features that could be of value as biomarkers, we undertook a series of filtering steps, to exclude features that are not stable in test–retest analyses and features that are highly correlated with each other to retain only features that predict the outcome variable of ischaemic stroke, this is summarised graphically in Figure 7.1.

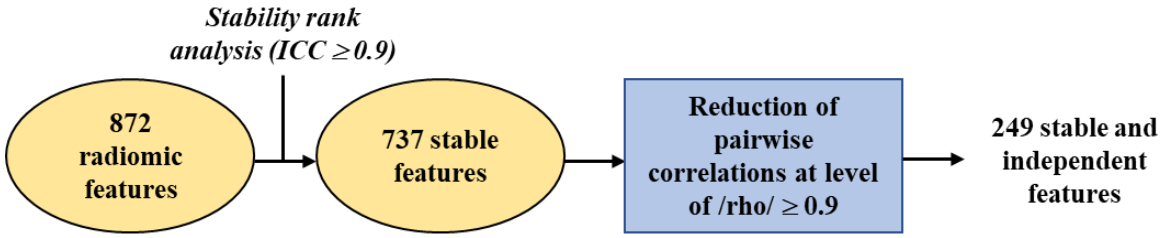


Figure 7.1. Radiomic feature selection process overview. A total of 737 out of 872 (84.5%) features were found to be stable following intra-observer analysis performed within the ORFAN Arm 4 cohort, as defined by an intraclass correlation coefficient (ICC) of ≥ 0.9. Further reduction of pairwise correlations at the level of /rho/≥0.9 resulted in 249 independent radiomic features for inclusion in analysis.

Firstly, stability assessment of all radiomic features was performed to ensure that only features that were stable between assessments were included in analysis. This was tested on 50 CCTA scans from the ORFAN Arm 4 cohort that had repeat automated segmentation of the relevant VOI completed. Of the total of 872 features extracted, 84.5% were found to be stable following ranking on descending order based on their ICC. A total of 737 features had an ICC equal or greater than 0.90 on inter-observer analysis and were included in further analysis (Figure 7.2).

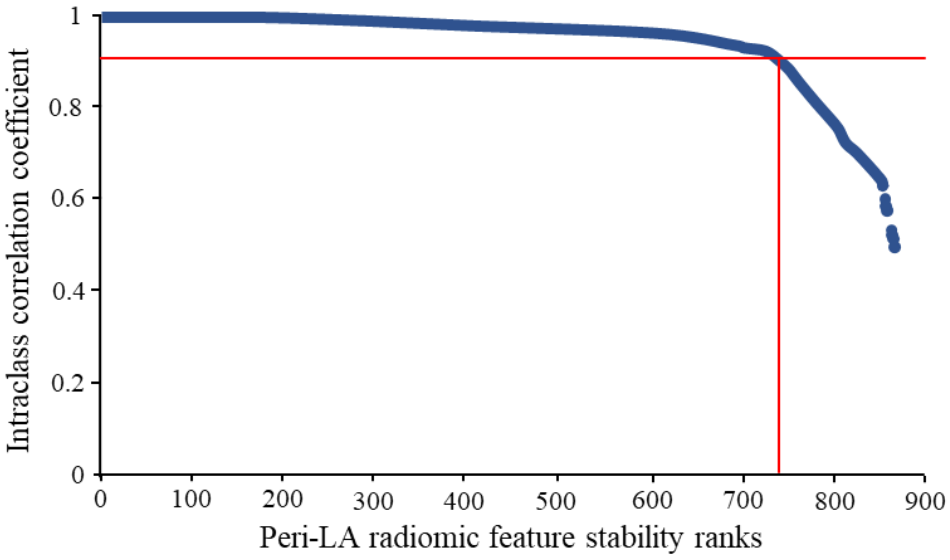


Figure 7.2. Stability analysis of LA and peri-LA volume of interests radiomic features. Plot of the inter-observer intraclass correlation coefficient of all 872 radiomic features measured in the LA and peri-LA volumes of interest. Radiomic features are ranked on descending order based on their ICC. A total of 731 radiomic features were found to have an ICC equal to or greater than 0.90 on inter-observer analysis and were subsequently selected for inclusion in further analyses.

Following the extraction and filtering of the radiomic features that were found to be unstable in test–retest analyses, features that were highly correlated with each other were identified and removed from the pool of features. This reduction of pairwise correlations at the level of |rho|≥0.9 resulted in a final pool of 249 radiomic features for inclusion in analysis for the generation of the ischaemic stroke risk prediction signature.

Recursive feature elimination with a random forest algorithm and repeated five-fold cross validation showed a peak in the accuracy of the trained model with 10 selected radiomic features (Figure 7.3).

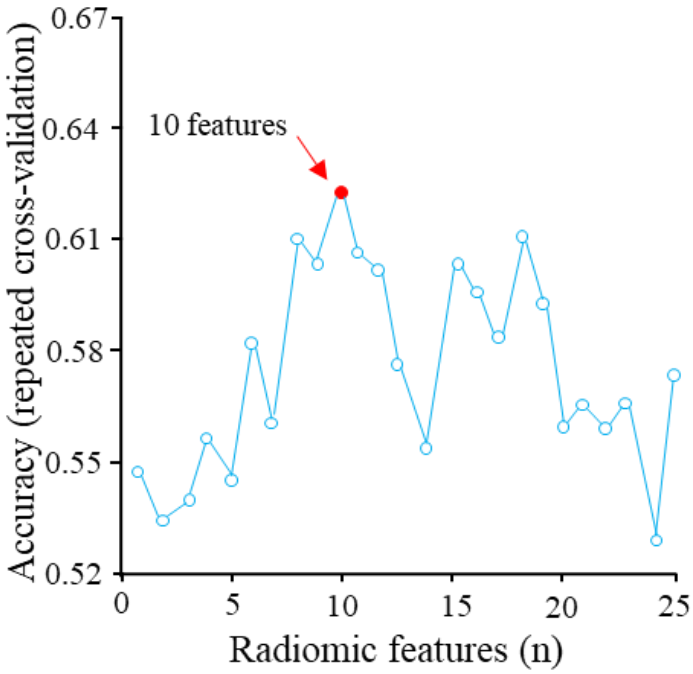


Figure 7.3. Recursive feature elimination. Recursive feature elimination with a random forest algorithm and repeated five-fold cross-validation showed a plateau in the accuracy of the trained model with a maximal number of 10 selected features.

To avoid overfitting we tuned the hyperparameters by optimising step size shrinkage, the L2 regularization parameter, and the learning rate. We used the early stopping technique that works by monitoring the performance of the model that is being trained on a separate test dataset and stopping the training procedure once the performance on the test dataset has not improved after a fixed number of training iterations. This avoids overfitting by attempting to automatically select the inflection point where performance on the test dataset starts to decrease while performance on the training dataset continues to improve as the model starts to overfit. The 10 features were used to train an extreme gradient boosting (XGBoost) algorithm using decisions trees to identify patients at risk of future ischaemic stroke events. The raw product of the algorithm was named the pilot Atriomic algorithm. The features that made up this signature and their importance in the model (gain value) are shown in Figure 7.4.

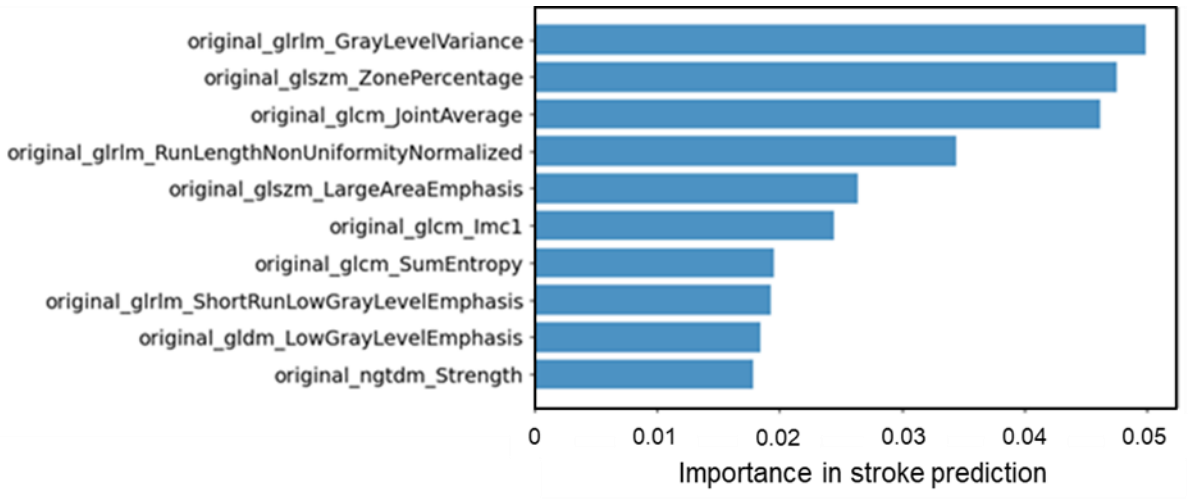


Figure 7.4. Radiomic features included in pilot radiomic stroke risk signature. Following the filtering of radiomic features and recursive feature elimination a plateau in the accuracy of feature risk prediction for ischaemic stroke was found at 10 features, each listed above. These are derived from the final random forest model and represent the features utilised in the risk prediction signature for ischaemic stroke. Gain values as shown here represent the relative contribution of each radiomic feature in the model. A higher gain value when compared to another feature implies higher importance for generating a prediction value.

7.2.5. Calculation of the CHA_2DS_2-VASc Score

The CHA_2DS_2-VASc score[14] was calculated for each patient by assigning 1 point each for age 65–74 years, hypertension diagnosis, diabetes mellitus diagnosis, congestive heart failure diagnosis, vascular disease diagnosis, and female sex and 2 points for a previous stroke or TIA and age ≥75 years. All conditions were defined according to ICD-10 definitions as outlined in Appendix 7, with data sourced from NHS Digital HES datasets.

7.2.6. Statistical analysis

Participant demographics are summarized as numbers (percentages) or median (IQR) for categorical and continuous variables, respectively. Hierarchical clustering of radiomic features was performed using Ward's method and Minkowski distance with *p* set to 10. Radiomic features that were retained after filtering, were scaled and fit in an extreme gradient boosting algorithm (*XGBoost*) with decision trees, with the eta set to 0.5, number of rounds 100, and maximum tree depth 50. Prognostic value for follow-up ischaemic stroke was assessed in receiver operating characteristic (ROC) curves for both the internal and external testing cohorts, with different risk scores compared via the delta of the area under the curve (AUC). Statistical analysis was performed in STATA SE version 15 (Stata Corp, College Station, TX, USA).

7.3. Results

Figure 7.5 demonstrates the flow chart of patients within this pilot study for the training, internal testing, and external testing of the radiomic signature for stroke risk, including exclusions, median follow up time, and ischaemic stroke outcomes following the CCTA. Between 2010 and 2021, 19,922 individual patients underwent clinical CCTA at either The Royal Brompton and Harefield NHS Foundation Trust or the Oxford University Hospitals NHS Foundation Trust (Group 1). A large proportion of exclusions were necessary to limit the training and internal testing cohorts to relevant patients who had at least 5 years of follow-up following their clinical CCTA, and had never had an ischaemic stroke prior to the CCTA. In total 6266 patients from across each site were excluded, resulting in 11656 patients with diagnostic CCTA being available for inclusion in group 1. Consequentially, 9325 patients were utilised for training of the algorithm (80%) and 2331 were utilised for internal testing (20%). The external testing cohort was drawn from the Papworth NHS Foundation Trust. Following exclusions, 4523 patients were available with at least 5 years follow up following their CCTA.

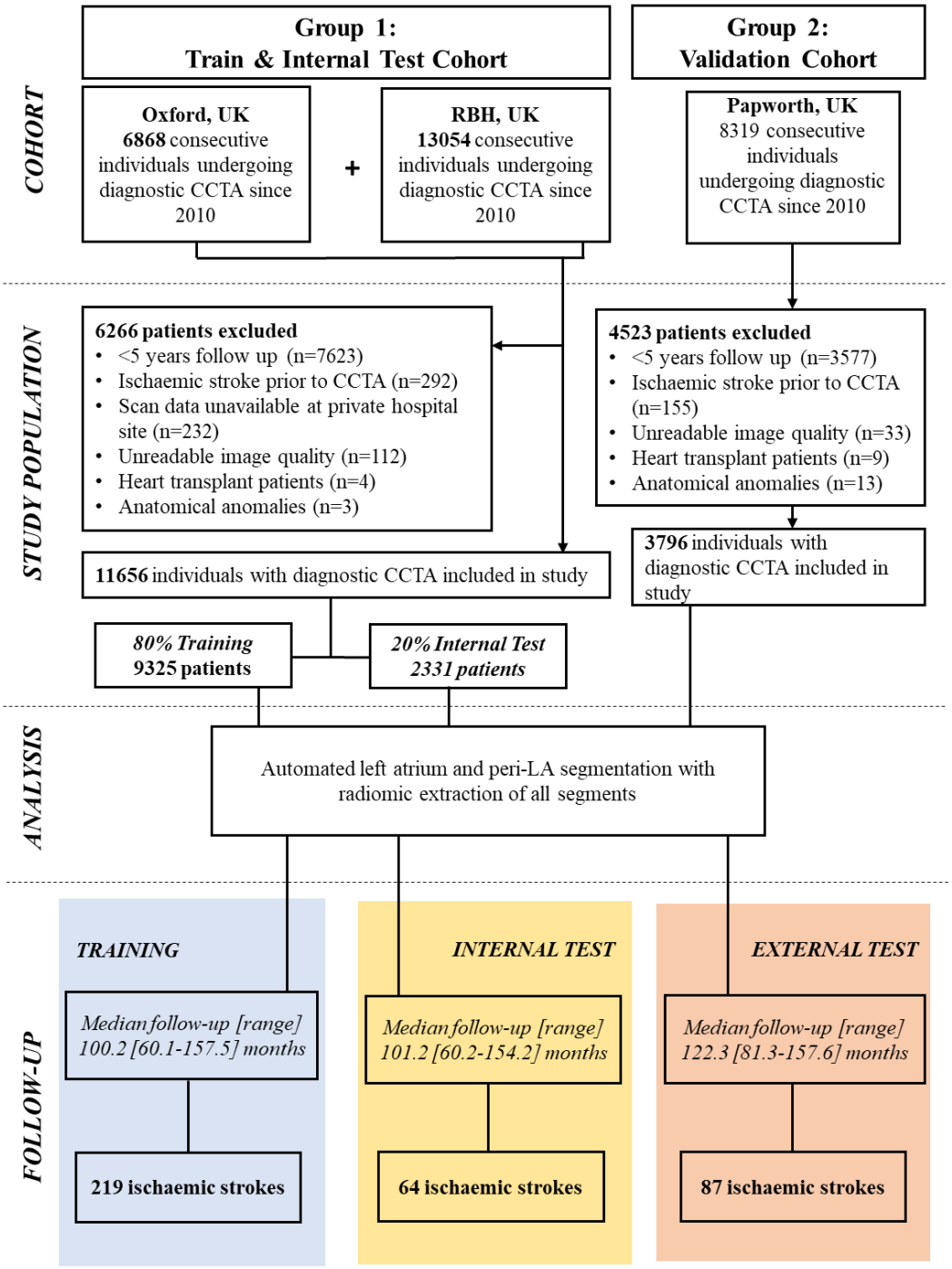


Figure 7.5. Study design and flowchart. The ORFAN Arm 4 Study cohorts utilised in this pilot development of a CT radiomic risk signature for ischaemic stroke are outlined. In total, 19,922 patients were available for inclusion in group 1. Following exclusions, the total number of patients available for the training cohort was 9325 and the total number of patients available for the internal testing cohort was 2331. External validation was undertaken in 4523 number of patients following exclusions.
CCTA: coronary computed tomography angiography; RBH: Royal Brompton and Harefield NHS Foundation Trust

The demographics and technical scan parameters of the cohorts utilised for training, internal validation, and external validation at the time of the CCTA scan are shown in Table 7.1.

During follow-up of the training cohort there were 219 confirmed ischaemic stroke events. In the internal testing cohort there were 64 events, and in the external testing cohort there were 87 events.

Table 7.1. ORFAN Arm 4 cohort demographics for the pilot radiomic risk score for ischaemic stroke

	Training cohort	**Internal test cohort**	**External test**
Number	9325	2331	4523
Age in years	62.2 (52.4-76.9)	61.2 (53.1-76.6)	61.3 (52.1-68.9)
Male gender	5194 (55.7%)	1326 (56.9%)	2429 (53.7%)
White ethnicity	7162 (76.8%)	1872 (80.3%)	3781 (83.6%)
Hypertension	4308 (46.2%)	1070 (45.9%)	2198 (48.6%)
Hypercholesterolaemia	2117 (22.7%)	487 (20.9%)	900 (19.9%)
Diabetes mellitus	1567 (16.8%)	405 (17.4%)	728 (16.1%)
Valve disease	336 (3.6%)	98 (4.2%)	294 (6.5%)
Peripheral vascular disease	849 (9.1%)	233 (10.0%)	371 (8.2%)
Chronic coronary syndrome	2247 (24.1%)	520 (22.3%)	1103 (24.4%)
Atrial fibrillation	1212 (13.0%)	270 (11.6%)	597 (13.2%)
CHA_2DS_2-VASc Score*	2.5 (0.8-4.4)	2.2 (0.4-4.2)	2.1 (0.4-3.9)
CCTA scans			
Sites (scanner make & model)	Oxford (GE Medical Systems & Toshiba Aquilion One), RBH (Siemens Drive)		Papworth (GE Medical Systems & Toshiba Aquilion One)
Slice thickness *0.625mm*	7879 (84.5%)	1939 (83.2%)	4523 (100%)
0.5mm	1446 (15.5%)	392 (16.8%)	0 (0%)
Tube voltage *120 kVp*	8859 (95.0%)	2175 (93.3%)	3899 (86.2%)
100 kVp	345 (3.7%)	135 (5.8%)	624 (13.8%)
80 kVp	121 (1.3%)	21 (0.9%)	0 (0%)
Prospective follow-up			
Duration in months (range)	100.2 (60.1-157.5)	101.2 (60.2-154.2)	122.3 (81.3-157.6)
Ischaemic stroke	219 (2.3%)	64 (2.7%)	87 (1.9%)

*Stroke risk score developed for patients with atrial fibrillation[14] 0-9 depending on risk factors: congestive heart failure (1 point), hypertension (1 point), age ≥75 years (2 points) or age 65–74 years (1 point), diabetes mellitus (1 point), previous stroke or TIA (2 points), vascular disease (1 point), female sex (1 point). GE: General Electric Corporation; kVp: kilovoltage peak; RBH: Royal Brompton & Harefield. Categorical variables are presented as numbers (percentages), while continuous variables are summarized as median (IQR).

7.3.1. Internal testing of the pilot stroke risk prediction signature.

The 10-feature pilot radiomic signature was applied into the sequestered internal testing cohort of 2331 patients to discriminate those at risk of ischaemic stroke following their CCTA (Figure 7.6). The model performed excellently internally in receiver operator characteristic (ROC) curve analysis, with an area under the curve (AUC) of 0.89.

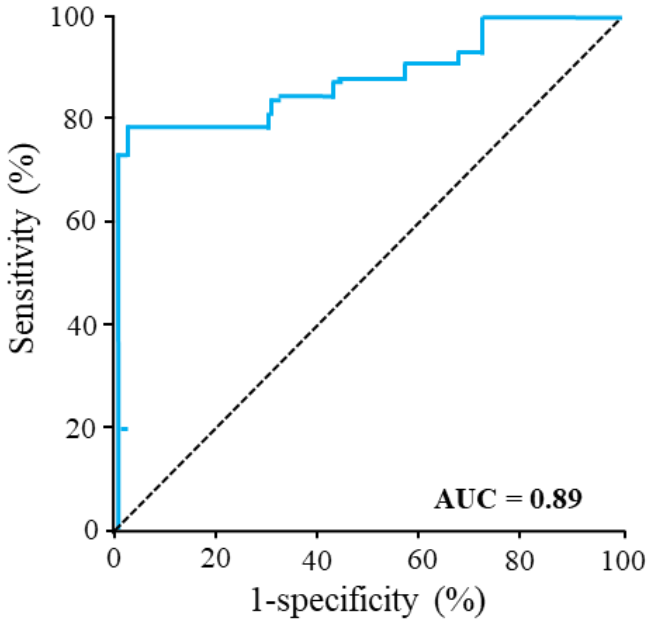


Figure 7.6. Internal testing of the pilot radiomic stroke risk score. In 2331 patients sequestered from the training data, the pilot radiomic risk score for the discrimination of ischaemic stroke achieved an AUC of 0.89 in ROC curve analysis. AUC: area under the curve; ROC: receiver operator characteristic.

7.3.2. External testing of the pilot stroke risk prediction signature.

External validation initially occurred in the geographically distinct sample from the Papworth NHS Foundation Trust. In this sample of 4523 patients, the 10-feature pilot radiomic signature for the discrimination of those at risk of ischaemic stroke following their CCTA did not achieve desirable predictive power for ischaemic stroke. This was likely due to the low event rate within this specific external test cohort, with 87 unique stroke events occurring in the 5 years immediately following the CCTA.

Given this, to test the performance of the radiomic model for data drift within the external test sample the analysis was re-run with a random selection of 50% (n=2261) of the Papworth patients included in the training data, with an external test cohort of the remaining 50% (n=2262) patients

remaining for external testing. Within this reduced external cohort there were 38 ischaemic strokes (43.7% of total in overall Papworth sample). In this cohort ROC curve analysis was performed for the prediction of ischaemic stroke via the pilot Atriomic score (radiomic features only) and the CHA_2DS_2-VASc score (Figure 7.7). The incidence of ischaemic stroke was significantly higher in Papworth patients with a CHA_2DS_2-VASc score ≥ 4 than in those with a CHA_2DS_2-VASc score < 4 (the log-rank test, p <0.001).

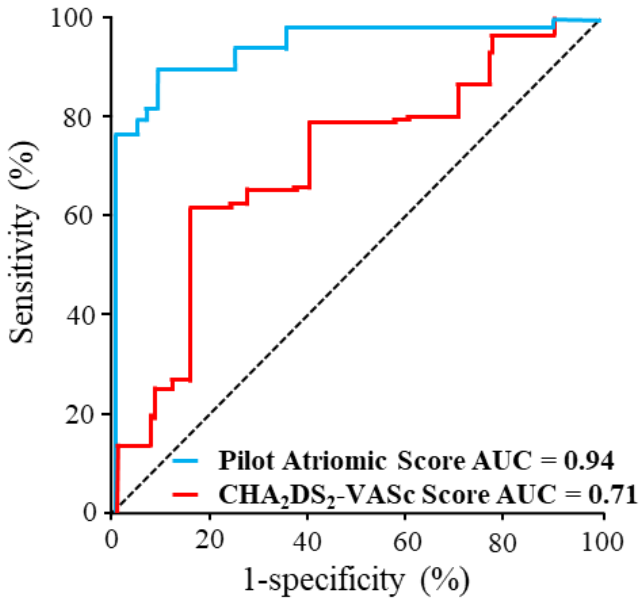


Figure 7.7. External testing of the pilot radiomic stroke risk score in re-training analysis. In analysis of 2262 external patients from Papworth, UK, the pilot radiomic risk score for the discrimination of ischaemic stroke achieved an AUC of 0.94 in ROC curve analysis (blue). 2261 Papworth patients had been included in training data for this analysis. Stroke risk was also assessed via the CHA_2DS_2-VASc score (red). AUC: area under the curve; ROC: receiver operator characteristic.

7.4. Discussion

In this chapter I demonstrate the use of CT radiomics for the detection of ischaemic stroke risk in an early pilot analysis of the ORFAN Arm 4 Study – the results presented here are preliminary and do not represent the finalised radiotranscriptomic Atriomic Risk Algorithm which is undergoing further validation. This analysis demonstrates the promise of radiomic risk signatures developed in large scale clinical cohorts such as the ORFAN Study.

In this preliminary analysis, a radiomic only risk signature consisting of 10 radiomic features was found to be predictive of future ischaemic stroke risk in the 5 years following a clinical CCTA scan being performed. Broad external testing (geographically distinct cohort with no merging of test data with training data) of the pilot radiomic risk signature within the Papworth NHS Trust cohort was not

adequately powered for the accurate prediction of ischaemic stroke events and hence was not possible without the re-training of the model with 50% of the intended external test cohort included within the training pool. The area under the ROC curve in the re-training analysis was 0.94, reflecting the potential of this model for the prediction of stroke events and the need for further external testing with larger cohorts. The preliminary model outperformed the CHA_2DS_2-VASc score for the prediction of ischaemic stroke risk, with a delta area under the ROC curve of 0.23. This analysis included all-comers for CCTA at Papworth regardless of AF status, which is important to note given that the CHA_2DS_2-VASc score is primarily validated for the prediction of ischaemic stroke risk in those with known non-valvular AF.

This preliminary ORFAN Arm 4 Study analysis has limitations that must be considered. As stated, further broad external validation is compulsory to test this model in a cohort with higher numbers of events. Fortunately, this is now available within the ORFAN Arm 4 Study and the scans from ORFAN sites including Milton Keynes, Leicester, St. Bartholomew's, and Milan are undergoing automated radiomic feature extraction for this purpose, with other sites to soon follow. The radiomic model will be retrained using these CCTA datasets as both additional training and for external testing, with a focus on external testing in a larger pooled external cohort with higher events. As stated, this analysis is pilot in nature and the radiomic risk signature will go on to become a possible basis for the radiotranscriptomic ischaemic stroke risk signature that is trained upon the transcriptomic ground truth for atrial myopathy sourced from RNA sequencing data of EAT and atrial tissue samples in the AdipoRedOx Study. The radiotranscriptomic analysis will form the final Atriomic Risk Algorithm that will be tested further in external cohorts from the ORFAN Arm 4 Study. The 10 radiomic features identified here show great promise as being relevant for the risk of ischaemic stroke. Comparison of these radiomic features between patients with elevated expression of genes associated with EAT and myocardial inflammation and fibrosis will be the ultimate test of their relevance. We anticipate the performance of the model to improve significantly with this next critical step.

The high area under the ROC curve when 50% of the external cohort was merged into the training dataset indicates possible differences in the cohorts (between training and testing) that have not been understood. Such differences could be population specific (i.e., a different cohort of patients undergoing CCTA), scanner specific (i.e., non-transferable segmentation and/or radiomic feature extraction due to technical acquisition parameters), or statistical model specific (i.e., limitations in the machine learning model that have not been appreciated that are limiting compatibility of the 10-radiomic feature signature with the external cases). There may also be differential risk of stroke between the Papworth cohort and the training; those who had stroke after the CCTA in the Papworth cohort may have experienced this due to different causes compared to those from Oxford

and Bath. The inter-cohort comparison of risk factors as shown in Table 7.1 did not reveal any fundamental differences between the original training, internal test and external test cohorts, however we have not been able to analyse all potential sources of bias in this preliminary analysis. Such sources of bias are now being explored and include the specific indication for the CCTA, which is extracted from the CCTA request data. We did make efforts to avoid such biases including through the exclusion of heart transplant patients which are more common at the Papworth NHS Trust.

As already stated, due to COVID-19 pandemic related delays in both CCTA transfer from collaborating sites and the delivery of the RNA sequencing data for the atrial and EAT tissue samples from the AdipoRedOx Study a preliminary radiomic only risk signature is presented here as a proof of principle study. The major research priority is the completion of the radiotranscriptomic risk score and, as explained in the subsequent chapter, the data for this have arrived at the OXACCT Core Lab as of late December 2022.

A further limitation of the pilot radiomic stroke risk signature is that the presented analysis is of a single merged VOI consisting of the LA segment and the three peri-LA. Analysis of the radiomic risk signatures, and the differences between them, that can be generated utilising each individual segmentation (i.e. the intra-atrial wall segment alone, the anterior LA wall segment, and the peri-LAA segment) and the permeations of merged segments (i.e. the merged radiomic features of the intra-atrial wall segment and the anterior LA wall segment; and the intra-atrial wall segment and peri-LAA segment) is now needed and is planned – this will be informative of any specific regions of the automated segments that may be causing errors in risk prediction modelling. Moreover, we intend to compare all radiomic and radiotranscriptomic stroke risk assessment tools with clinical risk scores including the QRISK3 algorithm[14] and the FSRS[13]. Only the CHA_2DS_2-VASc score is presented in this analysis. Other relevant risk scores are not presented here as they require clinical data from the time of the scan such as the serum LDL level and data such as treatment status with anti-hypertensive medications. These data are collected within the ORFAN Arm 4 local site databases however these databases were largely undelivered to the OXACCT Core Lab at the time of writing due to the significant delays caused by the global COVID-19 pandemic. The fully developed radiotranscriptomic Atriomic Risk algorithm will be compared to both the QRISK3 and the FSRS.

The number of ORFAN Arm 4 CCTA scans available for inclusion in this analysis was limited due to technical issues that are being actively worked on within the OXACCT Core Lab. The major rate limiting factor for scan inclusion in analysis is the computational power and time required to move tens of thousands of CCTA files from cloud-based servers to physical GPU, and then move the resulting segmentation files to an accessible and analysable off-line environment for the extraction

of the radiomic features from each ROI. The computational power of the GPU for segmentation purposes is not the limiting step, as our OXACCT GPU is designed specifically for purpose and is adequately resourced, it is the bandwidth of data servers and server switches that limit the movement of DICOM data that has resulted in less scans than planned for being available for inclusion.

It is important to consider the flow of CCTA data within the ORFAN Arm 4 Study to understand the technical obstacles that must be overcome. Data is received in a cloud based external Cimar server. The scans must then be downloaded from the cloud and saved onto university research servers. In accordance with best practice, these storage servers are not directly linked to the computational resource available for actioning the automated scan segmentation, and so the scans must be moved to the GPU accessible server when ready for analysis. Results of the segmentation must then be moved back to the storage server once complete so that radiomic extraction can occur.

When moving data between standard University of Oxford servers not custom built for the transmission of medical imaging files there is significant delay. The nature of medical imaging files is such that regular data storage servers struggle with the retrieval and transfer of millions of nested files. Each CCTA scan series has over 600 folders with DICOM data stored within each folder. Each scan can have up to 20 unique series', meaning that the transfer of 10,000 individual patient's CCTA requires a server to search for, retrieve, and push over 120,000,000 individual folders each containing CCTA data[200].

A further technical complication arises from the large number of CCTA series that sites push to the Core Lab when sending their patients' CCTA scans. Only a single series is required to undergo automated analysis, typically the 75% phase series by convention. We found that not all CCTA scans include DICOM tags that include the phase data required to automatically select the correct series for analysis. This significantly impacted upon the inclusion of CCTA in the analysis, as the manual selection of series for automated segmentation is time-consuming to the point of being impossible for tens of thousands of scans.

7.5. Conclusion

In this chapter, I present the generation and partial validation of a pilot radiomic signature for the assessment of ischaemic stroke risk. I show that a radiomic only signature of 10 radiomic features extracted from the LA and peri-LA region of clinical CCTA can provide discrimination of patients at risk of future ischaemic stroke and provide the basis for ongoing testing and validation of a radiotranscriptomic risk signature for the prediction of ischaemic stroke.

This analysis demonstrates the potential of the ORFAN Arm 4 Study as a large-scale clinical CCTA cohort for the discovery and validation of novel disease risk assessment tools. The ORFAN Arm 4 Study will be of major importance to the future generation of clinically relevant CT risk scores to enhance patient care for cardiovascular disease risk. The pilot analysis shown here is generated in two of the eight ORFAN Arm 4 study sites. Once full operational, both radiomic and radiotranscriptomic disease risk scores will be able to be tested in up to 88,223 patients initially, with up to 200,000 patients planned.

8

Chapter 8: Limitations and future directions

This project has limitations. Some limitations of the work relate to technical aspects of CT imaging, which are often acknowledged in similar research work. Other limitations relate to methodological weaknesses and challenging concepts in artificial intelligence science. The project was also impacted by the global COVID-19 pandemic. This Chapter discusses these limitations and the future research objectives that are underway to address them, and fill knowledge gaps raised by the work of the thesis.

8.1. Technical considerations

Computerised tomography imaging was invented and became popular in medicine because of the promise of detailed anatomical insights into human anatomy and disease. This was initially through the qualitative description of the morphological appearance of human organs and bodily structures. The methods utilised in this thesis do not use CT in this traditional manner but instead rely upon the quantitative analysis of the individual voxels that make up each CT slice. The examination of voxel-wise data relies on the absolute and relative attenuation values of voxels to be accurate. This form of CT analysis is inherently more sensitive to technical scan acquisition parameters than traditional qualitative assessment of human subjects with CT.

Scan acquisition parameters can influence the calculated raw values of density assigned to CT voxels, and this adds potential bias into analysis and can limit inter-operability of any findings if not considered at the time of analysis. A classic example of this issue is the technical limitations of the coronary calcium score (CCS). Any change in CT scan slice thickness and tube voltage directly influences the volume of coronary calcium detected in the same patient[201]. CCS is utilised in clinical practice when the initiation of statin therapy remains uncertain, and it is typically utilised as a categorical variable with groups reliant on cut-offs rather than as a continuous variable. This opens the door for patient reclassification due to variation of CT acquisition parameters and is a risk that must be considered by any clinician performing CCS. These issues also apply to the analysis performed in this thesis. The stability of radiomic features has proven to be average, with some studies reporting up to 59% of radiomic features being unstable (ICC<0.9) when variability in CT acquisition settings such as tube voltage, field of view, slice thickness, pitch value and milliamperage is introduced[202].

Given this a diversity of CT scan acquisition parameters need to be incorporated into analysis to ensure any segmentation and risk signatures maintain operability across CT vendors and users. This was performed in the studies that constitute this thesis. In the ORFAN Arm 4 study there is no attempt to standardise CT acquisition protocols, as this is a retrospective collection of CCTA data. However, as is reported in Chapter 2, the acquisition protocols across UK ORFAN sites are very similar in regard to the key acquisition parameters of tube voltage, field of view, slice thickness, pitch value and milliamperage. For the SCOT-HEART Trial, which constituted most of the analysis in Chapter 5, over two-thirds of the cohort were scanned on the same CT scanner[150]. In Chapter 3, for the development of the novel segmentation tools, scans were pooled from across different scanners and sites to ensure that the extraction of EAT from the scans was successful on different machines with slightly different acquisition settings. Pooling such data means that the data is brought closer to the original ground truth of the ROI, and limits possible cofounding by cohort, scanner, and protocol specific variations. The imminent addition of further ORFAN Arm 4 sites into the analysis pool will further improve the external validity of any observations.

8.2. CCTA versus other risk assessment technologies

Within this thesis there was no ability to compare the developed CCTA technologies, such as automated CCTA derived EAT volume and the pilot radiomic stroke risk algorithm, with similar technology derived from other sources. CMR derived EAT volume[203] and even echocardiography derived EAT thickness[204] would be excellent comparators for analysis of patient disease risk conveyed by EAT. This is a limitation that will be difficult to overcome in the immediate future as to compare the CCTA derived technologies a cohort that includes multiple imaging modalities and appropriate follow up for outcomes is required. Such cohorts are very difficult to create. A sub-study within the ORFAN Arm 4 Study that involves the collection of CMR and echocardiography from certain accessible sites is a possible solution to this limitation that is being considered.

Non-imaging technologies must also be considered. There has been significant progress in the AI powered assessment of ECG data for the prediction of new AF which could be compared to any CT derived risk score for atrial myopathy[205]. The ORFAN Arm 4 Study does not collect ECG traces from sites currently, and the addition of such data would be a significant logistical undertaking if ECG traces were to be sourced for patients from around the time of the CCTA. However, this is a critical comparison that needs to be made, as ECG derived detection of atrial myopathy risk would be a major advancement for patient care. In a similar manner to CMR and echocardiography data, the acquisition of ECG traces within the ORFAN Arm 4 Study cohort is being explored at study sites with high quality and comprehensive electronic records for the automated extraction of these data for comparative analysis.

8.3. Artificial intelligence methods

The use of machine learning to create risk signatures that rely upon radiomic features unlocks analysis of CT scans that is impossible for a human interpreter of the same scan. This is the fundamental strength of the methodology, as it is a new and potentially very useful tool to improve patient care through the analysis of complex data. However, there are acceptability issues when such methods function as a 'black box' that makes little sense to the clinician interpreting the data and the patient receiving care that is influenced by such data[206]. This is not a major issue for the analysis presented within the thesis, as the major use of AI presented here is for the automation of a manual process that is easily explained in terms of biological plausibility and relevance – the role of EAT.

The machine learning analysis for the development of radiomic risk signatures is more difficult to explain unless it is informed by biological ground truth that relates to underlying biological mechanisms such as transcriptomics. This will be achieved in the imminent future work to follow on from this thesis, as explained below, to provide biological plausibility to the Atriomic Risk Algorithm. Further methodological enhancement to empower clinicians and patients to understand the tools developed could include the use of deep-learning 'on scan' localisation highlighting[207], which provides visual heatmaps on the CCTA scan demonstrating the area of the LA and peri-LA tissues that the majority of risk is being conveyed from for that individual patient. This would effectively demonstrate the most diseased tissues and allow the clinician to explain how the technology works and what the images mean for the patient. Such methods greatly improve acceptability of the technologies by empowering patients to learn about the origins of the tests and how they can assist the clinician in their provision of healthcare[208].

8.4. Ongoing study plans and imminent research objectives

There is significant research analysis that is ongoing that was delayed due to the COVID-19 pandemic. These immediate research priorities are outlined below.

8.4.1. ORFAN Arm 4 Study priorities

- The completion of data collection from local site databases across all sites. This will allow pathology results from the time of the CCTA scan to be incorporated into analysis and allow complete analysis of cardiovascular risk factors such as smoking status ad family history.
- The completion of CCTA request (clinician requests for CCTA) and report analysis. This will allow analysis of the reasons for CCTA across the UK, and the incorporation of scan indication into analysis for relevant outcomes.

- The onboarding of remaining sites to bring total numbers of patients enrolled in the study up to the maximum of 200,000 in the UK, and 50,000 internationally.
- The second data linkage with NHS Digital and the national registries of SSNAP and NICOR for outcomes data through until the end of 2022.

8.4.2. CCTA analysis priorities

- As of December 2022, the whole heart segmentation algorithm is being applied to all CCTA scans from the first 88223 patients enrolled into the Study. EAT volume data will soon be analysable for the entire cohort. This will enable the analysis of the role of EAT in a far greater number of patients with outcomes compared to what was possible within the SCOT-HEART Trial population. This will be the largest EAT study to date; with patient numbers 20x what has previously been reported. In this thesis we reported that EAT was an independent predictor of all-cause mortality but were unable to use diagnostic data to understand through what diseases this effect is mediated beyond cardiovascular outcomes. We will use the detailed and comprehensive ORFAN Arm 4 outcomes data to understand through what diseases this effect is being mediated. The ORFAN Arm 4 data includes all causes and dates of death from ONS data.
- Radiomic analysis of the whole heart EAT volume is planned. This analysis will involve the extraction of radiomic features for the full EAT segmentations across the 88223 patients. A full EAT radiomic risk analysis will allow the detection of dysfunctional adipose tissue phenotypes that are more relevant to disease outcomes than EAT volume alone. The role of mean attenuation, as has been previously explored as a relevant radiomic feature for EAT[209], will also be investigated individually.
- The LA and peri-LA segmentations are being performed on the full ORFAN Arm 4 cohort. This will allow cohort wide assessment of risk for ischaemic stroke as has already been developed, and also the development of specific risk score for AF. There are enough new AF cases within the ORFAN Arm 4 Study outcomes that we can create risk scores for the development of new AF following the CCTA scan. This will allow a merging of radiomic risk signatures to ensure that our risk prediction models are targeting the atrial myopathy in all of its manifestations.
- The analysis of other CCTA metrics including the perivascular Fat Attenuation Index[163] and the Fat Radiomic Profile[144] will commence in the entire ORFAN Arm 4 cohort. This will allow the largest assessment of the role of coronary artery inflammation to date on relevant cardiovascular outcomes including the full breadth of coronary events (hospitalisations for chest pain, PCI, CABG and death).

8.4.3. Advanced stroke risk prediction signatures – radiotranscriptomic risk

In this thesis, a radiomic risk signature was created for the prediction of the likelihood of future stroke. This signature is a useful tool, however we will be improving this technology through the incorporation of transcriptomic ground truth into the models. The Oxford run AdipoRedOx Study, as utilised already within this thesis for CCTA scans, has collected over 400 right atrial appendage and EAT tissue samples from patients undergoing clinically indicated cardiac surgery. These patients also underwent a research CCTA. All tissue samples have recently completed RNA sequencing and we have a transcriptomic library of atrial appendage and EAT. We will train a radiomic signature to detect the upregulated expression of genes known to be associated to both atrial myocardial and EAT inflammation and fibrosis. This radiotranscriptomic signature will then be applied to the ORFAN Arm 4 cohort at large to test for the detection of AF and stroke risk, and comparison with the radiomic only risk signature already developed.

Figure 8.1 outlines the developmental pipeline in place for the creation of the AF and stroke Atriomic Risk Algorithms in a radiotranscriptomic manner. Our laboratory has previously utilised this approach with the Fat Radiomic Profile[144]. This risk signature of peri-coronary artery adipose tissue detects perivascular structural remodelling associated with CAD, beyond only inflammation. The radiomic signature was trained on the ground truth expression of TNF-A for the detection of inflammation, and COL1A1 and CD31 for the detection of fibrosis and vascularity, respectively. For future analysis, the models developed via the AdipoRedOx cohort can be further fine-tuned within the ORFAN Arm 4 cohort in an outcomes-based analysis to ensure that the risk signature is selecting for patients who are most likely to have disease events.

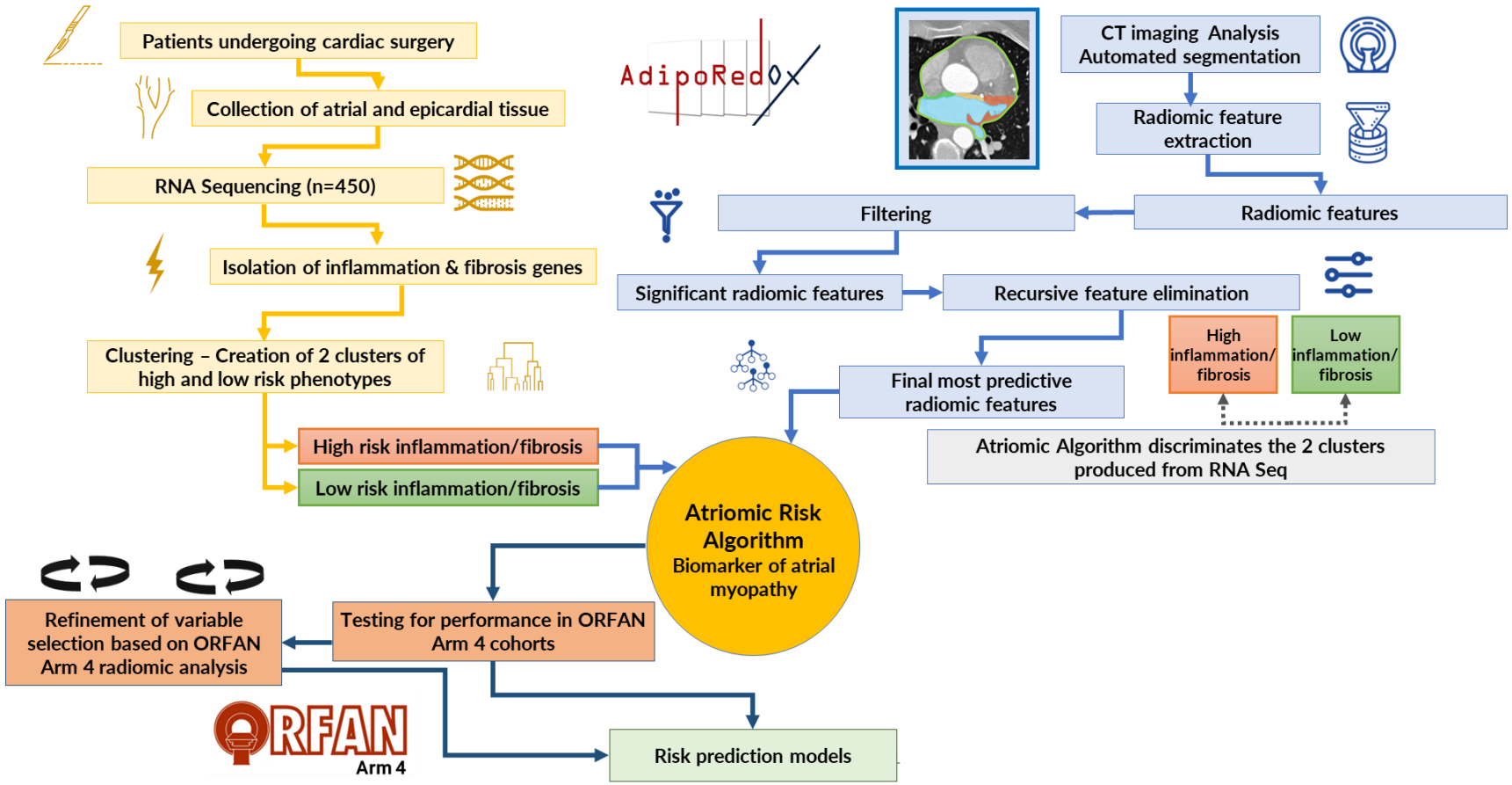


Figure 8.1. The methodological approach to the development of the radiotranscriptomic Atriomic Risk Algorithm. *The Oxford based AdipoRedOx Study allows the generation of biological ground truth in the form of RNA sequencing data. This allows the development of radiomic risk signature trained to detect the expression of genes related to atrial myocardial and EAT inflammation and fibrosis, the two key features of the atrial myopathy. The radiotranscriptomic signature can then be applied into the ORFAN Arm 4 cohort for testing and finetuning on disease outcomes of new AF following the CCTA and new ischaemic stroke following the CCTA. RNA Seq: Ribonucleic acid sequencing*

8.4.4. Data methodology research priorities

The data that the ORFAN Study has collected is of much interest to health economists and epidemiologists interested in the best methods for big data research, NHS data management and data access. In this thesis we report significant non-concordance between NHS Digital HES data and national registry data for stroke events. A more comprehensive analysis of non-concordance between these data sources is underway and will constitute the first major quality analysis of these data sources for nationwide cardiovascular disease outcomes. This is important for researchers and policy makers so that the quality of the datasets available can be understood and limitations can be addressed.

8.5. Long term research plans

The results in this thesis, and the future planned research to develop advanced radiotranscriptomic disease risk biomarkers, have been developed in a major observational study. Such studies are very useful for the initial conceptualisation, development, and validation of a risk models diagnostic and prognostic power, however they do not necessarily allow an assessment of true clinical value. The real clinical value of any risk assessment technology, just like any therapeutic intervention, must be assessed in a randomised controlled trial setting. Such a trial could assess the value of EAT volume assessment as an additional diagnostic to guide the escalation of medical therapy for CAD. The radiomic risk scores, and future radiotranscriptomic risk scores for the atrial myopathy, will require testing in a randomised setting to understand how they can be utilised for the prevention of stroke events. It would be unethical to withhold usual care from patients with known stroke risk, but our method could be tested in patients that do not exhibit any current indication for stroke risk reduction via any means, but who exhibit an abnormally high risk for AF and/or stroke at the time of their clinical CCTA for unrelated purposes (such as stable chest pain). Such patients would normally not come to the attention of physicians for targeted AF and stroke prevention purposes. The anticipated low event rate in any such trial would demand careful power calculations and the cost-effectiveness of any such trial would need major consideration.

Smaller randomised controlled trials may be equally as relevant for the purposes of understanding the role of anti-inflammatory and anti-fibrotic agents on the atrial myopathy. For example, a trial to examine the effect of high dose statin therapy on the atrial myopathy risk as assessed by a radiotranscriptomic risk algorithm would be vital evidence needed to understand possible treatment of stroke risk in those without AF. Novel recombinant protein therapies such as canakinumab, tocilizumab, rituximab and neuregulin may be relevant for such trials[210], along with other established medications such as aspirin, colchicine and statins.

9

Chapter 9: Conclusions

The work presented in this thesis has two major threads: 1) The introduction and creation of a large-scale clinical cohort of CCTA patients – the ORFAN Arm 4 Study, and 2) the development of novel CT risk assessment technologies for cardiovascular diseases, including the automated assessment of EAT volume, and the radiomic assessment of ischaemic stroke risk.

9.1. The ORFAN Arm 4 Study

The development of the cardiovascular risk tools as carried out in this thesis would not have been possible without the ORFAN Study. I expanded the pre-existing ORFAN Study from its original 3 prospective arms to include the hybrid retrospective/prospective fourth arm. This enabled big data and AI approaches to be employed for disease risk assessment. The development of the ORFAN Arm 4 Study as a non-consented clinical cohort required significant collaboration and negotiation with external organisations to ensure the study was appropriately designed to meet the strict requirements of the HRA CAG and the REC, as well as the data protection and safety requirements of each individual site. I created a methodology that allows local clinicians at hospital sites across the UK and the world to enrol their own clinical CT patients into the Study without having to share patient identifying information with the ORFAN Study team. This methodology creates an almost zero chance of a loss in pseudonymisation at any point in the study data flow. The fourth ORFAN Study arm captures patients referred for CCTA since the use of CT for heart disease purposes became an acceptable part of care in the late 2010s. As demonstrated in the number of patients referred for CCTA over the last 12 years, the use of CT for cardiovascular diagnostic purposes has massively increased. This huge increase warranted the creation of a study such as ORFAN Arm 4. With increasing patients undergoing CCTA examination the potential to add significant value to the scan for both clinicians and patients alike is an important objective.

The ORFAN Arm 4 Study required the establishment of wide-reaching data sharing agreements with the major clinical data controllers in the UK: NHS Digital and the Healthcare Quality Improvement Partnership (the data controller for the national cardiac and stroke outcomes registries). These agreements unlock the potential for research projects to target any clinical outcome that is catalogued within NHS Digital records from 2005 through to current, and for highly detailed analysis of cardiovascular disease events through the national registries. The broad inclusion criteria for the

ORFAN Arm 4 Study allows patients who have undergone any CT scan to be included in the study, not just CCTA as is utilised in this thesis. This unlocks research opportunities far broader than is explored in this thesis both within cardiovascular health and beyond.

At the time of writing, the number of patients enrolled in the ORFAN Arm 4 Study with data available is 88,223, all of which had undergone at least one CCTA. The expansion of the Study is underway to achieve 200,000 UK based clinical CCTA patients by the end of 2023. The collection of full baseline data for this large cohort has proven to be a major logistical undertaking that was complicated significantly by the global COVID-19 pandemic.

9.2. Automated EAT volume assessment

Despite significant evidence of the fundamental role of EAT in the development and severity of cardiovascular diseases such as CAD, AF, and heart failure, EAT volume remains on the fringes of cardiovascular risk assessment. When assessing the clinical value of this project, one needs to consider what was known before it started and what was discovered during its course. Until this project there had been no major studies (>2,000 patients) into the possible role of EAT quantification in risk assessment for cardiovascular and other outcomes. This is predominantly due to the challenges of quantification of EAT, with manual segmentation of large number of CT (or other modalities) requiring major commitment of time and resources.

I have developed and validated the most robust automated tool for the quantification of EAT available. The tool reduced the time required for an expert reader to manually segment EAT on a CCTA by 98% (1080 seconds to 22 seconds) when using a standard desktop CPU. The residual U-Net based convolutional neural network architecture utilised for the segmentation model provided an excellent classifier for CT voxels as within or outside of the desired segmentation and allowed dedicated feedback learning loops to fine-tune the automated segmentation tool. The model for the automated segmentation of EAT is validated in a broad external cohort[211] from a geographically distinct sample. Importantly for real-world applicability of this tool, I conducted sub-group analysis of selected cohorts of patients known to have characteristics that render them challenging for CT – obesity, metallic artefact within the heart, calcium blooming artefacts and those with atypical anatomy such as following CABG. Previous deep learning algorithms trained on CT have been shown to have significantly reduced accuracy in real-world data[212]. However, the DLN developed here for EAT quantification performed with almost equal concordance to a human expert reader in all patients regardless of challenging characteristics.

Throughout my project evidence has continued to grow regarding the fundamental relevance of EAT to cardiovascular diseases[213,214]. From 2018 to 2022 new records for the most papers including the

term ‘epicardial adipose tissue’ indexed in PubMed were set each year, with a pronounced focus on the biology underlining the interaction of EAT with myocardial and vascular tissues[92]. This increasing focus in biological mechanisms will lead to a heightened requirement for large scale EAT assessment tools such as the tool developed here. The translational benefits of increased understanding regarding the role of EAT in specific disease mechanisms will not be achieved unless tools such as what we have developed are able to provide rapid and reproducible EAT quantification via means that can be incorporated into both research and clinical environments.

Beyond purely technological development, I have also undertaken the largest study to date into the role of EAT volume in clinical disease outcomes utilising the well phenotyped SCOT-HEART Trial population. This study reveals significant findings related to the independent and powerful role of EAT in CAD and AF[1]. Specific notable findings from this large trial cohort analysis include that EAT volume is an independent risk factor for both AF and CAD at the time of CCTA, and that it can independently predict the risk of long term vascular events. Such events include non-fatal MI, of which EAT was already known to be associated, and stroke, for which there was far less compelling evidence. Most interesting is the independent association discovered between EAT volume and all-cause mortality and non-cardiac mortality. These findings are novel and will serve as the launch pad for targeted enquiry within the larger ORFAN Arm 4 Study cohort as to the exact outcomes that are driving this phenomenon. For the wide-spread clinical application of any measure such as EAT volume, cut-points must be created to allow clinicians to risk stratify patients according to their EAT volume. To this end, I generated two independent cut-points for two different population groups: a cut point for mortality and major vascular events in the intermediate risk patient population trained in the SCOT-HEART Trial population, and a cut point for post-cardiac surgery AF risk in the very high-risk AdipoRedOx Study population. These cut-points were able to provide excellent discriminatory power between high and low risk patients for their respective outcomes and will be further validated in the ORFAN Arm 4 Study population.

9.3. Stroke risk prediction

To demonstrate the utility of the ORFAN Arm 4 Study for disease outcomes research and the use of novel radiomic tools I utilised ischaemic stroke as the outcome of interest given its fundamental relevance to the atrial myopathy. The large event numbers of the ORFAN Arm 4 Study cohort, even at half its planned size, dwarf other observational and randomised CT cohorts, and indeed dwarfs most observational clinical cohorts established for any purpose. There were 8 stroke events recorded in the CCTA arm at the 5-year follow up of the SCOT-HEART Trial[153] and the PROMISE Study has not reported stroke event numbers[215]. We report 2228 unique patients with stroke events, 911 of which were ischaemic stroke events that occurred after the CCTA that rendered the patient

eligible for the study. Once all of the ORFAN Arm 4 Study CCTA scans are available and accessible on OXACCT servers, this resource will be the most powerful study for cardiovascular disease risk assessment utilising CT.

I have demonstrated the potential for the generation of novel CT risk scores using radiomic feature extraction from the LA chamber and segmentations of the surrounding atrial wall and tissues including EAT. The training of these segmentations was undertaken in the same manner as the EAT quantification from the whole heart segment, with excellent external validation results. Sub-group external validation in the challenging CT population cohorts was also excellent for these segmentations. The three peri-atrial segmentations are entirely novel segmentations that have never been attempted in CCTA, or any other form of clinical CT. The ability to study the CT voxels that are likely to be capturing the tissues where the atrial myopathy is occurring is essential for non-invasive detection of this disease process. Until now, only surrogate measures of the atrial myopathy were examinable in a non-invasive manner. The pilot study conducted within this thesis for the radiomic assessment of risk for ischaemic stroke demonstrated the immense potential for such tools and is perhaps the most promising message of this doctoral project. The tool presented in this thesis was trained in 9325 patients, which constituted 10.6% of the entire cohort ORFAN Arm 4 cohort as will soon be accessible on OXACCT servers. The performance of the radiomic only signature was excellent in an external validation group sourced from a distinct hospital site in the ORFAN Arm 4 Study. This technology is an example of the LA and peri-LA specific risk scores that will mature and be further fine-tuned and validated within the full ORFAN cohort. The biological ground truth available from the RNA sequencing data for patients within our own AdipoRedOx Study will unlock even more targeted risk assessment tools that can detect the subtle texture changes associated with subclinical biological changes occurring in patients with atrial myopathy. These approaches will transform the status quo of modern clinical risk assessment – powering the transition to the adoption of personalised models of care.

10

References

1. West HW, Siddique M, Williams MC, Volpe L, Desai R, Lyasheva M, Thomas S, Dangas K, Kotanidis CP, Tomlins P, Mahon C, Kardos A, Adlam D, Graby J, Rodrigues JCL, Shirodaria C, Deanfield J, Mehta NN, Neubauer S, Channon KM, Desai MY, Nicol ED, Newby DE, Antoniades C, ORFAN Investigators. Deep-Learning for Epicardial Adipose Tissue Assessment with Computed Tomography: Implications for Cardiovascular Risk Prediction. *JACC: Cardiovascular Imaging.* 2022; Accepted 17 November 2022.

2. Office of National Statistics. Leading causes of death, UK: 2001 to 2018. 2020; https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/articles/leadingcausesofdeathuk/2001to2018. Accessed 16 November, 2022.

3. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, Boehme AK, Buxton AE, Carson AP, Commodore-Mensah Y, Elkind MSV, Evenson KR, Eze-Nliam C, Ferguson JF, Generoso G, Ho JE, Kalani R, Khan SS, Kissela BM, Knutson KL, Levine DA, Lewis TT, Liu J, Loop MS, Ma J, Mussolino ME, Navaneethan SD, Perak AM, Poudel R, Rezk-Hanna M, Roth GA, Schroeder EB, Shah SH, Thacker EL, VanWagner LB, Virani SS, Voecks JH, Wang N-Y, Yaffe K, Martin SS, null n. Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association. *Circulation.* 2022;145(8):e153-e639.

4. Stroke Association United Kingdom. Atrial Fibrillation: information and resources. 2022; https://www.stroke.org.uk/professionals/atrial-fibrillation-information-and-resources. Accessed December 2, 2022.

5. Stroke Association United Kingdom. Stroke statistics. 2022; https://www.stroke.org.uk/what-is-stroke/stroke-statistics. Accessed December 2, 2022.

6. Patel A, Berdunov V, Quayyum Z, King D, Knapp M, Wittenberg R. Estimated societal costs of stroke in the UK based on a discrete event simulation. *Age and Ageing.* 2020;49(2):270-276.

7. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, Benetos A, Biffi A, Boavida J-M, Capodanno D, Cosyns B, Crawford C, Davos CH, Desormais I, Di Angelantonio E, Franco OH, Halvorsen S, Hobbs FDR, Hollander M, Jankowska EA, Michal M, Sacco S, Sattar N, Tokgozoglu L, Tonstad S, Tsioufis KP, van Dis I, van Gelder IC, Wanner C, Williams B, Group ESD. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice: Developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies With the special contribution of the European Association of Preventive Cardiology (EAPC). *European Heart Journal.* 2021;42(34):3227-3337.

8. Rana JS, Tabada GH, Solomon MD, Lo JC, Jaffe MG, Sung SH, Ballantyne CM, Go AS. Accuracy of the Atherosclerotic Cardiovascular Risk Equation in a Large Contemporary, Multiethnic Population. *J Am Coll Cardiol.* 2016;67(18):2118-2130.

9. Antoniades C, West HW. ESC CVD Prevention Guidelines 2021: improvements, controversies, and opportunities. *Cardiovascular research.* 2022;118(2):e17-e19.

10. Oikonomou EK, Antonopoulos AS, Schottlander D, Marwan M, Mathers C, Tomlins P, Siddique M, Klüner LV, Shirodaria C, Mavrogiannis MC, Thomas S, Fava A, Deanfield J, Channon KM, Neubauer S, Desai MY, Achenbach S, Antoniades C. Standardized measurement of coronary inflammation using cardiovascular computed tomography: integration in clinical care as a prognostic medical device. *Cardiovascular research.* 2021;117(13):2677-2690.

11. Greenland P, LaBree L, Azen SP, Doherty TM, Detrano RC. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. *Jama.* 2004;291(2):210-215.

12. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ.* 2017;357:j2099.

13. Flueckiger P, Longstreth W, Herrington D, Yeboah J. Revised Framingham Stroke Risk Score, Nontraditional Risk Markers, and Incident Stroke in a Multiethnic Cohort. *Stroke.* 2018;49(2):363-369.

14. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest.* 2010;137(2):263-272.

15. Mitchell LB, Southern DA, Galbraith D, Ghali WA, Knudtson M, Wilton SB. Prediction of stroke or TIA in patients without atrial fibrillation using CHADS2 and CHA2DS2-VASc scores. *Heart.* 2014;100(19):1524-1530.

16. Morillas P, Pallarés V, Fácila L, Llisterri JL, Sebastián ME, Gómez M, Castilla E, Camarasa R, Sandin M, García-Honrubia A. The CHADS2 Score to Predict Stroke Risk in the Absence of Atrial Fibrillation in Hypertensive Patients Aged 65 Years or Older. *Revista espanola de cardiologia (English ed).* 2015;68(6):485-491.

17. Lip GY. Stroke and bleeding risk assessment in atrial fibrillation: when, how, and why? *Eur Heart J.* 2013;34(14):1041-1049.

18. Melgaard L, Gorst-Rasmussen A, Lane DA, Rasmussen LH, Larsen TB, Lip GY. Assessment of the CHA2DS2-VASc Score in Predicting Ischemic Stroke, Thromboembolism, and Death in Patients With Heart Failure With and Without Atrial Fibrillation. *Jama.* 2015;314(10):1030-1038.

19. Wolsk E, Lamberts M, Hansen ML, Blanche P, Køber L, Torp-Pedersen C, Lip GY, Gislason G. Thromboembolic risk stratification of patients hospitalized with heart failure in sinus rhythm: a nationwide cohort study. *European journal of heart failure.* 2015;17(8):828-836.

20. Poçi D, Hartford M, Karlsson T, Herlitz J, Edvardsson N, Caidahl K. Role of the CHADS2 score in acute coronary syndromes: risk of subsequent death or stroke in patients with and without atrial fibrillation. *Chest.* 2012;141(6):1431-1440.

21. Welles CC, Whooley MA, Na B, Ganz P, Schiller NB, Turakhia MP. The CHADS2 score predicts ischemic stroke in the absence of atrial fibrillation among subjects with coronary heart disease: data from the Heart and Soul Study. *American heart journal.* 2011;162(3):555-561.

22. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, Prescott E, Storey RF, Deaton C, Cuisset T, Agewall S, Dickstein K, Edvardsen T, Escaned J, Gersh BJ, Svitil P, Gilard M, Hasdai D, Hatala R, Mahfoud F, Masip J, Muneretto C, Valgimigli M, Achenbach S, Bax JJ, Group ESD. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes: The Task Force for the diagnosis and management of chronic coronary syndromes of the European Society of Cardiology (ESC). *European Heart Journal.* 2019.

23. Lawton JS, Tamis-Holland JE, Bangalore S, Bates ER, Beckie TM, Bischoff JM, Bittl JA, Cohen MG, DiMaio JM, Don CW, Fremes SE, Gaudino MF, Goldberger ZD, Grant MC, Jaswal JB, Kurlansky PA, Mehran R, Metkus TS, Nnacheta LC, Rao SV, Sellke FW, Sharma G, Yong CM, Zwischenberger BA. 2021 ACC/AHA/SCAI Guideline for Coronary Artery Revascularization: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation.* 2022;145(3):e18-e114.

24. Patel KV, Pandey A, de Lemos JA. Conceptual Framework for Addressing Residual Atherosclerotic Cardiovascular Disease Risk in the Era of Precision Medicine. *Circulation.* 2018;137(24):2551-2553.

25. Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, Fonseca F, Nicolau J, Koenig W, Anker SD, Kastelein JJP, Cornel JH, Pais P, Pella D, Genest J, Cifkova R, Lorenzatti A, Forster T, Kobalava Z, Vida-Simiti L, Flather M, Shimokawa H, Ogawa H, Dellborg M, Rossi PRF, Troquay RPT, Libby P, Glynn RJ. Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *New England Journal of Medicine.* 2017;377(12):1119-1131.

26. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, Kuder JF, Wang H, Liu T, Wasserman SM, Sever PS, Pedersen TR. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *N Engl J Med.* 2017;376(18):1713-1722.

27. Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, Nissen SE, Pocock S, Poulter NR, Ravn LS, Steinberg WM, Stockner M, Zinman B, Bergenstal RM, Buse JB. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine.* 2016;375(4):311-322.

28. Eikelboom JW, Connolly SJ, Bosch J, Dagenais GR, Hart RG, Shestakovska O, Diaz R, Alings M, Lonn EM, Anand SS, Widimsky P, Hori M, Avezum A, Piegas LS, Branch KRH, Probstfield J, Bhatt DL, Zhu J, Liang Y, Maggioni AP, Lopez-Jaramillo P, O'Donnell M, Kakkar AK, Fox KAA, Parkhomenko AN, Ertl G, Störk S, Keltai M, Ryden L, Pogosova N, Dans AL, Lanas F, Commerford PJ, Torp-Pedersen C, Guzik TJ, Verhamme PB, Vinereanu D, Kim J-H, Tonkin AM, Lewis BS, Felix C, Yusoff K, Steg PG, Metsarinne KP, Cook Bruns N, Misselwitz F, Chen E, Leong D, Yusuf S. Rivaroxaban with or without Aspirin in Stable Cardiovascular Disease. *New England Journal of Medicine.* 2017;377(14):1319-1330.

29. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, Shaw W, Law G, Desai M, Matthews DR. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *New England Journal of Medicine.* 2017;377(7):644-657.

30. Sehested TSG, Bjerre J, Ku S, Chang A, Jahansouz A, Owens DK, Hlatky MA, Goldhaber-Fiebert JD. Cost-effectiveness of Canakinumab for Prevention of Recurrent Cardiovascular Events. *JAMA Cardiol.* 2019;4(2):128-135.

31. Maeda T, Nishi T, Funakoshi S, Tada K, Tsuji M, Satoh A, Kawazoe M, Yoshimura C, Arima H. Residual risks of ischaemic stroke and systemic embolism among atrial fibrillation patients with anticoagulation: large-scale real-world data (F-CREATE project). *Heart.* 2021;107(3):217.

32. Ding WY. Residual Stroke Risk in Atrial Fibrillation. *Arrhythmia & electrophysiology review.* 2021;10(3):147-153.

33. Steensig K, Olesen KKW, Thim T, Nielsen JC, Madsen M, Jensen SE, Jensen LO, Kristensen SD, Lip GYH, Maeng M. Predicting stroke in patients without atrial fibrillation. *European journal of clinical investigation.* 2019:e13103.

34. Packer M. Characterization, Pathogenesis, and Clinical Implications of Inflammation-Related Atrial Myopathy as an Important Cause of Atrial Fibrillation. *Journal of the American Heart Association.* 2020;9(7):e015343.

35. Van Gelder IC, Hagens VE, Bosker HA, Kingma JH, Kamp O, Kingma T, Said SA, Darmanata JI, Timmermans AJ, Tijssen JG, Crijns HJ. A comparison of rate control and rhythm control in patients with recurrent persistent atrial fibrillation. *N Engl J Med.* 2002;347(23):1834-1840.

36. Wyse DG, Waldo AL, DiMarco JP, Domanski MJ, Rosenberg Y, Schron EB, Kellen JC, Greene HL, Mickel MC, Dalquist JE, Corley SD. A comparison of rate control and rhythm control in patients with atrial fibrillation. *N Engl J Med.* 2002;347(23):1825-1833.

37. Glotzer TV, Daoud EG, Wyse DG, Singer DE, Ezekowitz MD, Hilker C, Miller C, Qi D, Ziegler PD. The relationship between daily atrial tachyarrhythmia burden from implantable device diagnostics and stroke risk: the TRENDS study. *Circulation Arrhythmia and electrophysiology.* 2009;2(5):474-480.

38. Ip J, Waldo AL, Lip GY, Rothwell PM, Martin DT, Bersohn MM, Choucair WK, Akar JG, Wathen MS, Rohani P, Halperin JL. Multicenter randomized study of anticoagulation guided by remote rhythm monitoring in patients with implantable cardioverter-defibrillator and CRT-D devices: Rationale, design, and clinical characteristics of the initially enrolled cohort The IMPACT study. *American heart journal.* 2009;158(3):364-370.e361.

39. Brambatti M, Connolly SJ, Gold MR, Morillo CA, Capucci A, Muto C, Lau CP, Van Gelder IC, Hohnloser SH, Carlson M, Fain E, Nakamya J, Mairesse GH, Halytska M, Deng WQ, Israel CW, Healey JS. Temporal relationship between subclinical atrial fibrillation and embolic events. *Circulation.* 2014;129(21):2094-2099.

40. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, Jr., Ellinor PT, Ezekowitz MD, Field ME, Furie KL, Heidenreich PA, Murray KT, Shea JB, Tracy CM, Yancy CW. 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society in Collaboration With the Society of Thoracic Surgeons. *Circulation.* 2019;140(2):e125-e151.

41. Shen MJ, Arora R, Jalife J. Atrial Myopathy. *JACC Basic Transl Sci.* 2019;4(5):640-654.

42. Ronsoni RM, Saffi MAL, Gonçalves MVM, Nakayama IH, Luz Leiria TL. A New Vision at the Interface of Atrial Fibrillation and Stroke. 2021;8.

43. Goldberger JJ, Arora R, Green D, Greenland P, Lee DC, Lloyd-Jones DM, Markl M, Ng J, Shah SJ. Evaluating the Atrial Myopathy Underlying Atrial Fibrillation: Identifying the Arrhythmogenic and Thrombogenic Substrate. *Circulation.* 2015;132(4):278-291.

44. Zipes DP. Atrial fibrillation. A tachycardia-induced atrial cardiomyopathy. *Circulation.* 1997;95(3):562-564.

45. Lindhardsen J, Ahlehoff O, Gislason GH, Madsen OR, Olesen JB, Svendsen JH, Torp-Pedersen C, Hansen PR. Risk of atrial fibrillation and stroke in rheumatoid arthritis: Danish nationwide cohort study. *Bmj.* 2012;344:e1257.

46. Rhee TM, Lee JH, Choi EK, Han KD, Lee H, Park CS, Hwang D, Lee SR, Lim WH, Kang SH, Cha MJ, Cho Y, Oh IY, Oh S. Increased Risk of Atrial Fibrillation and Thromboembolism in Patients with Severe Psoriasis: a Nationwide Population-based Study. *Sci Rep.* 2017;7(1):9973.

47. Lim SY, Bae EH, Han KD, Jung JH, Choi HS, Kim CS, Ma SK, Kim SW. Systemic lupus erythematosus is a risk factor for atrial fibrillation: a nationwide, population-based study. *Clinical and experimental rheumatology.* 2019;37(6):1019-1025.

48. Moon I, Choi EK, Jung JH, Han KD, Choi YJ, Park J, Cho JH, Lee E, Choe W, Lee SR, Cha MJ, Lim WH, Oh S. Ankylosing spondylitis: A novel risk factor for atrial fibrillation - A nationwide population-based study. *International journal of cardiology.* 2019;275:77-82.

49. Choi YJ, Choi EK, Han KD, Park J, Moon I, Lee E, Choe WS, Lee SR, Cha MJ, Lim WH, Oh S. Increased risk of atrial fibrillation in patients with inflammatory bowel disease: A nationwide population-based study. *World journal of gastroenterology.* 2019;25(22):2788-2798.

50. Aksoy H, Okutucu S, Sayin BY, Oto A. Non-invasive electrocardiographic methods for assessment of atrial conduction heterogeneity in ankylosing spondylitis. *European review for medical and pharmacological sciences.* 2016;20(11):2185-2186.

51. Efe TH, Cimen T, Ertem AG, Coskun Y, Bilgin M, Sahan HF, Pamukcu HE, Yayla C, Sunman H, Yuksel I, Yeter E. Atrial Electromechanical Properties in Inflammatory Bowel Disease. *Echocardiography (Mount Kisco, NY).* 2016;33(9):1309-1316.

52. Wanahita N, Messerli FH, Bangalore S, Gami AS, Somers VK, Steinberg JS. Atrial fibrillation and obesity--results of a meta-analysis. *American heart journal.* 2008;155(2):310-315.

53. Huxley RR, Alonso A, Lopez FL, Filion KB, Agarwal SK, Loehr LR, Soliman EZ, Pankow JS, Selvin E. Type 2 diabetes, glucose homeostasis and incident atrial fibrillation: the Atherosclerosis Risk in Communities study. *Heart.* 2012;98(2):133-138.

54. Kwon CH, Kim H, Kim SH, Kim BS, Kim HJ, Kim DK, Han SW, Ryu KH, Dong Sung J. The Impact of Metabolic Syndrome on the Incidence of Atrial Fibrillation: A Nationwide Longitudinal Cohort Study in South Korea. *Journal of clinical medicine.* 2019;8(8).

55. Käräjämäki AJ, Pätsi OP, Savolainen M, Kesäniemi YA, Huikuri H, Ukkola O. Non-Alcoholic Fatty Liver Disease as a Predictor of Atrial Fibrillation in Middle-Aged Population (OPERA Study). *PloS one.* 2015;10(11):e0142937.

56. Baumgartner C, da Costa BR, Collet T-H, Feller M, Floriani C, Bauer DC, Cappola AR, Heckbert SR, Ceresini G, Gussekloo J, den Elzen WPJ, Peeters RP, Luben R, Völzke H, Dörr M, Walsh JP, Bremner A, Iacoviello M, Macfarlane P, Heeringa J, Stott DJ, Westendorp RGJ, Khaw K-T, Magnani JW, Aujesky D, Rodondi N. Thyroid Function Within the Normal Range, Subclinical Hypothyroidism, and the Risk of Atrial Fibrillation. *Circulation.* 2017;136(22):2100-2116.

57. Mourtzinis G, Adamsson Eryd S, Rosengren A, Björck L, Adiels M, Johannsson G, Manhem K. Primary aldosteronism and thyroid disorders in atrial fibrillation: A Swedish nationwide case-control study. *European journal of preventive cardiology.* 2018;25(7):694-701.

58. Packer M. Epicardial Adipose Tissue May Mediate Deleterious Effects of Obesity and Inflammation on the Myocardium. *J Am Coll Cardiol.* 2018;71(20):2360-2372.

59. Rafaqat S, Rafaqat S, Rafaqat S. Pathophysiological role of major adipokines in Atrial Fibrillation. *International Journal of Arrhythmia.* 2021;22(1):18.

60. Scridon A, Dobreanu D, Chevalier P, Şerban RC. Inflammation, a link between obesity and atrial fibrillation. *Inflammation research : official journal of the European Histamine Research Society [et al].* 2015;64(6):383-393.

61. Antonopoulos AS, Antoniades C. The role of epicardial adipose tissue in cardiac biology: classic concepts and emerging roles. *The Journal of physiology.* 2017;595(12):3907-3917.

62. Lipson A, Alexopoulos N, Hartlage GR, Arepalli C, Oeser A, Bian A, Gebretsadik T, Shintani A, Stillman AE, Stein CM, Raggi P. Epicardial adipose tissue is increased in patients with systemic lupus erythematosus. *Atherosclerosis.* 2012;223(2):389-393.

63. Wang X, Guo Z, Zhu Z, Bao Y, Yang B. Epicardial fat tissue in patients with psoriasis:a systematic review and meta-analysis. *Lipids in health and disease.* 2016;15:103.

64. Uysal F, Akbal E, Akbal A, Cevizci S, Arık K, Gazi E. Epicardial Adipose Tissue Is Increased in Patients With Inflammatory Bowel Disease. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine.* 2016;35(9):1859-1864.

65. Rabkin SW. The relationship between epicardial fat and indices of obesity and the metabolic syndrome: a systematic review and meta-analysis. *Metabolic syndrome and related disorders.* 2014;12(1):31-42.

66. Chen WJ, Danad I, Raijmakers PG, Halbmeijer R, Harms HJ, Lammertsma AA, van Rossum AC, Diamant M, Knaapen P. Effect of type 2 diabetes mellitus on epicardial adipose tissue volume and coronary vasomotor function. *The American journal of cardiology.* 2014;113(1):90-97.

67. Fracanzani AL, Pisano G, Consonni D, Tiraboschi S, Baragetti A, Bertelli C, Norata GD, Dongiovanni P, Valenti L, Grigore L, Tonella T, Catapano A, Fargion S. Epicardial Adipose Tissue (EAT) Thickness Is Associated with Cardiovascular and Liver Damage in Nonalcoholic Fatty Liver Disease. *PloS one.* 2016;11(9):e0162473.

68. Korkmaz L, Sahin S, Akyuz AR, Ziyrek M, Anaforoglu I, Kose M, Erkan H, Ağaç MT, Acar Z. Epicardial adipose tissue increased in patients with newly diagnosed subclinical hypothyroidism. *Medical principles and practice : international journal of the Kuwait University, Health Science Centre.* 2013;22(1):42-46.

69. Iacobellis G, Petramala L, Marinelli C, Calvieri C, Zinnamosca L, Concistrè A, Iannucci G, De Toma G, Letizia C. Epicardial Fat Thickness and Primary Aldosteronism. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme.* 2016;48(4):238-241.

70. Nagashima K, Okumura Y, Watanabe I, Nakai T, Ohkubo K, Kofune M, Mano H, Sonoda K, Hiro T, Nikaido M, Hirayama A. Does Location of Epicardial Adipose Tissue Correspond to Endocardial High Dominant Frequency or Complex Fractionated Atrial Electrogram Sites During Atrial Fibrillation? *Circulation: Arrhythmia and Electrophysiology.* 2012;5(4):676-683.

71. Abe I, Teshima Y, Kondo H, Kaku H, Kira S, Ikebe Y, Saito S, Fukui A, Shinohara T, Yufu K, Nakagawa M, Hijiya N, Moriyama M, Shimada T, Miyamoto S, Takahashi N. Association of fibrotic remodeling and cytokines/chemokines content in epicardial adipose tissue with atrial myocardial fibrosis in patients with atrial fibrillation. *Heart rhythm.* 2018;15(11):1717-1727.

72. Bos D, Vernooij MW, Shahzad R, Kavousi M, Hofman A, van Walsum T, Deckers JW, Ikram MA, Heeringa J, Franco OH, van der Lugt A, Leening MJG. Epicardial Fat Volume and the Risk of Atrial Fibrillation in the General Population Free of Cardiovascular Disease. *JACC Cardiovascular imaging.* 2017;10(11):1405-1407.

73. Baragetti A, Pisano G, Bertelli C, Garlaschelli K, Grigore L, Fracanzani AL, Fargion S, Norata GD, Catapano AL. Subclinical atherosclerosis is associated with Epicardial Fat Thickness and hepatic steatosis in the general population. *Nutrition, metabolism, and cardiovascular diseases : NMCD.* 2016;26(2):141-153.

74. Oba K, Maeda M, Maimaituxun G, Yamaguchi S, Arasaki O, Fukuda D, Yagi S, Hirata Y, Nishio S, Iwase T, Takao S, Kusunose K, Yamada H, Soeki T, Wakatsuki T, Harada M, Masuzaki H, Sata M, Shimabukuro M. Effect of the Epicardial Adipose Tissue Volume on the Prevalence of Paroxysmal and Persistent Atrial Fibrillation. *Circulation journal : official journal of the Japanese Circulation Society.* 2018;82(7):1778-1787.

75. Packer M. Leptin-Aldosterone-Neprilysin Axis. *Circulation.* 2018;137(15):1614-1631.

76. Nguyen Dinh Cat A, Antunes TT, Callera GE, Sanchez A, Tsiropoulou S, Dulak-Lis MG, Anagnostopoulou A, He Y, Montezano AC, Jaisser F, Touyz RM. Adipocyte-Specific Mineralocorticoid Receptor Overexpression in Mice Is Associated With Metabolic Syndrome and Vascular Dysfunction: Role of Redox-Sensitive PKG-1 and Rho Kinase. *Diabetes.* 2016;65(8):2392-2403.

77. Garg R, Rao AD, Baimas-George M, Hurwitz S, Foster C, Shah RV, Jerosch-Herold M, Kwong RY, Di Carli MF, Adler GK. Mineralocorticoid receptor blockade improves coronary microvascular function in individuals with type 2 diabetes. *Diabetes.* 2015;64(1):236-242.

78. Tamura N, Ogawa Y, Chusho H, Nakamura K, Nakao K, Suda M, Kasahara M, Hashimoto R, Katsuura G, Mukoyama M, Itoh H, Saito Y, Tanaka I, Otani H, Katsuki M. Cardiac fibrosis in mice lacking brain natriuretic peptide. *Proceedings of the National Academy of Sciences of the United States of America.* 2000;97(8):4239-4244.

79. Bordicchia M, Ceresiani M, Pavani M, Minardi D, Polito M, Wabitsch M, Cannone V, Burnett JC, Jr., Dessì-Fulgheri P, Sarzani R. Insulin/glucose induces natriuretic peptide clearance receptor in human adipocytes: a metabolic link with the cardiac natriuretic pathway. *American journal of physiology Regulatory, integrative and comparative physiology.* 2016;311(1):R104-114.

80. Cao H, Lin J, Chen W, Xu G, Sun C. Baseline adiponectin and leptin levels in predicting an increased risk of disease activity in rheumatoid arthritis: A meta-analysis and systematic review. *Autoimmunity.* 2016;49(8):547-553.

81. Huby A-C, Antonova G, Groenendyk J, Gomez-Sanchez CE, Bollag WB, Filosa JA, de Chantemèle EJB. Adipocyte-Derived Hormone Leptin Is a Direct Regulator of Aldosterone Secretion, Which Promotes Endothelial Dysfunction and Cardiac Fibrosis. *Circulation.* 2015;132(22):2134-2145.

82. Lubitz SA, Faranesh AZ, Selvaggi C, Atlas SJ, McManus DD, Singer DE, Pagoto S, McConnell MV, Pantelopoulos A, Foulkes AS. Detection of Atrial Fibrillation in a Large Population Using Wearable Devices: The Fitbit Heart Study. *Circulation.* 2022;146(19):1415-1424.

83. Kuppahally SS, Akoum N, Burgon NS, Badger TJ, Kholmovski EG, Vijayakumar S, Rao SN, Blauer J, Fish EN, Dibella EV, Macleod RS, McGann C, Litwin SE, Marrouche NF. Left atrial strain and strain rate in patients with paroxysmal and persistent atrial fibrillation: relationship to left atrial structural remodeling detected by delayed-enhancement MRI. *Circulation Cardiovascular imaging.* 2010;3(3):231-239.

84. Binici Z, Intzilakis T, Nielsen OW, Køber L, Sajadieh A. Excessive supraventricular ectopic activity and increased risk of atrial fibrillation and stroke. *Circulation.* 2010;121(17):1904-1911.

85. Larsen BS, Kumarathurai P, Falkenberg J, Nielsen OW, Sajadieh A. Excessive Atrial Ectopy and Short Atrial Runs Increase the Risk of Stroke Beyond Incident Atrial Fibrillation. *J Am Coll Cardiol.* 2015;66(3):232-241.

86. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, Carter RE, Yao X, Rabinstein AA, Erickson BJ, Kapa S, Friedman PA. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet.* 2019;394(10201):861-867.

87. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB, Sinner MF, Sotoodehnia N, Fontes JD, Janssens ACJW, Kronmal RA, Magnani JW, Witteman JC, Chamberlain AM, Lubitz SA, Schnabel RB, Agarwal SK, McManus DD, Ellinor PT, Larson MG, Burke GL, Launer LJ, Hofman A, Levy D, Gottdiener JS, Kääb S, Couper D, Harris TB, Soliman EZ, Stricker BHC, Gudnason V, Heckbert SR, Benjamin EJ. Simple Risk Model Predicts Incidence of Atrial Fibrillation in a Racially and Geographically Diverse Population: the CHARGE-AF Consortium. *Journal of the American Heart Association.*2(2):e000102.

88. Khurshid S, Friedman S, Reeder C, Di Achille P, Diamant N, Singh P, Harrington LX, Wang X, Al-Alusi MA, Sarma G, Foulkes AS, Ellinor PT, Anderson CD, Ho JE, Philippakis AA, Batra P, Lubitz SA. ECG-Based Deep Learning and Clinical Risk Factors to Predict Atrial Fibrillation. *Circulation.* 2022;145(2):122-133.

89. Rohner A, Brinkert M, Kawel N, Buechel RR, Leibundgut G, Grize L, Kühne M, Bremerich J, Kaufmann BA, Zellweger MJ, Buser P, Osswald S, Handke M. Functional assessment of the left atrium by real-time three-dimensional echocardiography using a novel dedicated analysis tool: initial validation studies in comparison with computed tomography. *European journal of echocardiography : the journal of the Working Group on Echocardiography of the European Society of Cardiology.* 2011;12(7):497-505.

90. Daccarett M, Badger TJ, Akoum N, Burgon NS, Mahnkopf C, Vergara G, Kholmovski E, McGann CJ, Parker D, Brachmann J, Macleod RS, Marrouche NF. Association of left atrial fibrosis detected by delayed-enhancement magnetic resonance imaging and the risk of stroke in patients with atrial fibrillation. *J Am Coll Cardiol.* 2011;57(7):831-838.

91. Lee DC, Markl M, Ng J, Carr M, Benefield B, Carr JC, Goldberger JJ. Three-dimensional left atrial blood flow characteristics in patients with atrial fibrillation assessed by 4D flow CMR. *European heart journal cardiovascular Imaging.* 2016;17(11):1259-1268.

92. Iacobellis G. Epicardial adipose tissue in contemporary cardiology. *Nature reviews Cardiology.* 2022;19(9):593-606.

93. Zheng YY, Ma YT, Zhang JY, Xie X. COVID-19 and the cardiovascular system. *Nature reviews Cardiology.* 2020;17(5):259-260.

94. McAninch EA, Fonseca TL, Poggioli R, Panos AL, Salerno TA, Deng Y, Li Y, Bianco AC, Iacobellis G. Epicardial adipose tissue has a unique transcriptome modified in severe coronary artery disease. *Obesity (Silver Spring, Md).* 2015;23(6):1267-1278.

95. Sacks HS, Fain JN, Bahouth SW, Ojha S, Frontini A, Budge H, Cinti S, Symonds ME. Adult epicardial fat exhibits beige features. *The Journal of clinical endocrinology and metabolism.* 2013;98(9):E1448-1455.

96. Mazurek T, Zhang L, Zalewski A, Mannion JD, Diehl JT, Arafat H, Sarov-Blat L, O'Brien S, Keiper EA, Johnson AG, Martin J, Goldstein BJ, Shi Y. Human epicardial adipose tissue is a source of inflammatory mediators. *Circulation.* 2003;108(20):2460-2466.

97. Hirata Y, Tabata M, Kurobe H, Motoki T, Akaike M, Nishio C, Higashida M, Mikasa H, Nakaya Y, Takanashi S, Igarashi T, Kitagawa T, Sata M. Coronary atherosclerosis is associated with macrophage polarization in epicardial adipose tissue. *J Am Coll Cardiol.* 2011;58(3):248-255.

98. Gao X, Mi S, Zhang F, Gong F, Lai Y, Gao F, Zhang X, Wang L, Tao H. Association of chemerin mRNA expression in human epicardial adipose tissue with coronary atherosclerosis. *Cardiovascular diabetology.* 2011;10:87.

99. Langheim S, Dreas L, Veschini L, Maisano F, Foglieni C, Ferrarello S, Sinagra G, Zingone B, Alfieri O, Ferrero E, Maseri A, Ruotolo G. Increased expression and secretion of resistin in epicardial adipose tissue of patients with acute coronary syndrome. *American journal of physiology Heart and circulatory physiology.* 2010;298(3):H746-753.

100. Baker AR, Harte AL, Howell N, Pritlove DC, Ranasinghe AM, da Silva NF, Youssef EM, Khunti K, Davies MJ, Bonser RS, Kumar S, Pagano D, McTernan PG. Epicardial adipose tissue as a source of nuclear factor-kappaB and c-Jun N-terminal kinase mediated inflammation in patients with coronary artery disease. *The Journal of clinical endocrinology and metabolism.* 2009;94(1):261-267.

101. Antonopoulos AS, Margaritis M, Verheule S, Recalde A, Sanna F, Herdman L, Psarros C, Nasrallah H, Coutinho P, Akoumianakis I, Brewer AC, Sayeed R, Krasopoulos G, Petrou M, Tarun A, Tousoulis D, Shah AM, Casadei B, Channon KM, Antoniades C. Mutual Regulation of Epicardial Adipose Tissue and Myocardial Redox State by PPAR-γ/Adiponectin Signalling. *Circulation research.* 2016;118(5):842-855.

102. Antoniades C, Demosthenous M, Reilly S, Margaritis M, Zhang MH, Antonopoulos A, Marinou K, Nahar K, Jayaram R, Tousoulis D, Bakogiannis C, Sayeed R, Triantafyllou C, Koumallos N, Psarros C, Miliou A, Stefanadis C, Channon KM, Casadei B. Myocardial redox state predicts in-hospital clinical outcome after cardiac surgery effects of short-term pre-operative statin treatment. *J Am Coll Cardiol.* 2012;59(1):60-70.

103. Margaritis M, Antonopoulos AS, Digby J, Lee R, Reilly S, Coutinho P, Shirodaria C, Sayeed R, Petrou M, De Silva R, Jalilzadeh S, Demosthenous M, Bakogiannis C, Tousoulis D, Stefanadis C, Choudhury RP, Casadei B, Channon KM, Antoniades C. Interactions between vascular wall and perivascular adipose tissue reveal novel roles for adiponectin in the regulation of endothelial nitric oxide synthase function in human vessels. *Circulation.* 2013;127(22):2209-2221.

104. Antonopoulos AS, Sanna F, Sabharwal N, Thomas S, Oikonomou EK, Herdman L, Margaritis M, Shirodaria C, Kampoli AM, Akoumianakis I, Petrou M, Sayeed R, Krasopoulos G, Psarros C, Ciccone P, Brophy CM, Digby J, Kelion A, Uberoi R, Anthony S, Alexopoulos N, Tousoulis D, Achenbach S, Neubauer S, Channon KM, Antoniades C. Detecting human coronary inflammation by imaging perivascular fat. *Sci Transl Med.* 2017;9(398).

105. Liu D, Ceddia RP, Collins S. Cardiac natriuretic peptides promote adipose 'browning' through mTOR complex-1. *Molecular metabolism.* 2018;9:192-198.

106. Dutour A, Achard V, Sell H, Naour N, Collart F, Gaborit B, Silaghi A, Eckel J, Alessi MC, Henegar C, Clément K. Secretory type II phospholipase A2 is produced and secreted by epicardial adipose tissue and overexpressed in patients with coronary artery disease. *The Journal of clinical endocrinology and metabolism.* 2010;95(2):963-967.

107. Furuhashi M, Fuseya T, Murata M, Hoshina K, Ishimura S, Mita T, Watanabe Y, Omori A, Matsumoto M, Sugaya T, Oikawa T, Nishida J, Kokubu N, Tanaka M, Moniwa N, Yoshida H, Sawada N, Shimamoto K, Miura T. Local Production of Fatty Acid-Binding Protein 4 in Epicardial/Perivascular Fat and Macrophages Is Linked to Coronary Atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2016;36(5):825-834.

108. Pezeshkian M, Mahtabipour MR. Epicardial and subcutaneous adipose tissue Fatty acids profiles in diabetic and non-diabetic patients candidate for coronary artery bypass graft. *BioImpacts : BI.* 2013;3(2):83-89.

109. Camarena V, Sant D, Mohseni M, Salerno T, Zaleski ML, Wang G, Iacobellis G. Novel atherogenic pathways from the differential transcriptome analysis of diabetic epicardial adipose tissue. *Nutrition, metabolism, and cardiovascular diseases : NMCD.* 2017;27(8):739-750.

110. Hatem SN, Redheuil A, Gandjbakhch E. Cardiac adipose tissue and atrial fibrillation: the perils of adiposity. *Cardiovascular research.* 2016;109(4):502-509.

111. Yamaguchi Y, Cavallero S, Patterson M, Shen H, Xu J, Kumar SR, Sucov HM. Adipogenesis and epicardial adipose tissue: a novel fate of the epicardium induced by mesenchymal transformation and PPARγ activation. *Proceedings of the National Academy of Sciences of the United States of America.* 2015;112(7):2070-2075.

112. Suffee N, Moore-Morris T, Farahmand P, Rücker-Martin C, Dilanian G, Fradet M, Sawaki D, Derumeaux G, LePrince P, Clément K, Dugail I, Puceat M, Hatem SN. Atrial natriuretic peptide regulates adipose tissue accumulation in adult atria. *Proceedings of the National Academy of Sciences of the United States of America.* 2017;114(5):E771-e780.

113. Gaborit B, Venteclef N, Ancel P, Pelloux V, Gariboldi V, Leprince P, Amour J, Hatem SN, Jouve E, Dutour A, Clément K. Human epicardial adipose tissue has a specific transcriptomic signature depending on its anatomical peri-atrial, peri-ventricular, or peri-coronary location. *Cardiovascular research.* 2015;108(1):62-73.

114. Shaihov-Teper O, Ram E, Ballan N, Brzezinski RY, Naftali-Shani N, Masoud R, Ziv T, Lewis N, Schary Y, Levin-Kotler LP, Volvovitch D, Zuroff EM, Amunts S, Regev-Rudzki N, Sternik L, Raanani E, Gepstein L, Leor J. Extracellular Vesicles From Epicardial Fat Facilitate Atrial Fibrillation. *Circulation.* 2021;143(25):2475-2493.

115. Wong CX, Stiles MK, John B, Brooks AG, Lau DH, Dimitri H, Kuklik P, Shipp NJ, Sullivan T, Sanders P. Direction-dependent conduction in lone atrial fibrillation. *Heart rhythm.* 2010;7(9):1192-1199.

116. Nalliah CJ, Bell JR, Raaijmakers AJA, Waddell HM, Wells SP, Bernasochi GB, Montgomery MK, Binny S, Watts T, Joshi SB, Lui E, Sim CB, Larobina M, O'Keefe M, Goldblatt J, Royse A, Lee G, Porrello ER, Watt MJ, Kistler PM, Sanders P, Delbridge LMD, Kalman JM. Epicardial Adipose Tissue Accumulation Confers Atrial Conduction Abnormality. *J Am Coll Cardiol.* 2020;76(10):1197-1211.

117. Iacobellis G, Corradi D, Sharma AM. Epicardial adipose tissue: anatomic, biomolecular and clinical relationships with the heart. *Nature clinical practice Cardiovascular medicine.* 2005;2(10):536-543.

118. Balcioğlu AS, Çiçek D, Akinci S, Eldem HO, Bal UA, Okyay K, Müderrisoğlu H. Arrhythmogenic evidence for epicardial adipose tissue: heart rate variability and turbulence are influenced by epicardial fat thickness. *Pacing and clinical electrophysiology : PACE.* 2015;38(1):99-106.

119. Pokushalov E, Kozlov B, Romanov A, Strelnikov A, Bayramova S, Sergeevichev D, Bogachev-Prokophiev A, Zheleznev S, Shipulin V, Lomivorotov VV, Karaskov A, Po SS, Steinberg JS. Long-Term Suppression of Atrial Fibrillation by Botulinum Toxin Injection Into Epicardial Fat Pads in Patients Undergoing Cardiac Surgery: One-Year Follow-Up of a Randomized Pilot Study. *Circulation Arrhythmia and electrophysiology.* 2015;8(6):1334-1341.

120. Oikonomou EK, Antoniades C. The role of adipose tissue in cardiovascular health and disease. *Nature Reviews Cardiology.* 2019;16(2):83-99.

121. Ntambi JM, Young-Cheul K. Adipocyte differentiation and gene expression. *The Journal of nutrition.* 2000;130(12):3122s-3126s.

122. Antonopoulos AS, Margaritis M, Coutinho P, Digby J, Patel R, Psarros C, Ntusi N, Karamitsos TD, Lee R, De Silva R, Petrou M, Sayeed R, Demosthenous M, Bakogiannis C, Wordsworth PB, Tousoulis D, Neubauer S, Channon KM, Antoniades C. Reciprocal effects of systemic inflammation and brain natriuretic peptide on adiponectin biosynthesis in adipose tissue of patients with ischemic heart disease. *Arterioscler Thromb Vasc Biol.* 2014;34(9):2151-2159.

123. Kotanidis CP, Antoniades C. Perivascular fat imaging by computed tomography (CT): a virtual guide. *British journal of pharmacology.* 2021;178(21):4270-4290.

124. Antoniades C, Antonopoulos A, West HW, Inventors. Radiomic Signature of an Epicardial Region. US patent PCT/GB2019/0530582018.

125. Maurovich-Horvat P, Ferencik M, Voros S, Merkely B, Hoffmann U. Comprehensive plaque assessment by coronary CT angiography. *Nature reviews Cardiology.* 2014;11(7):390-402.

126. Clayton B, Roobottom C, Morgan-Hughes G. Assessment of the myocardium with cardiac computed tomography. *European heart journal cardiovascular Imaging.* 2014;15(6):603-609.

127. Eroğlu S. How do we measure epicardial adipose tissue thickness by transthoracic echocardiography? *Anatolian journal of cardiology.* 2015;15(5):416-419.

128. Davidovich D, Gastaldelli A, Sicari R. Imaging cardiac fat. *European Heart Journal - Cardiovascular Imaging.* 2013;14(7):625-630.

129. Bouazizi K, Rahhal A, Kusmia S, Evin M, Defrance C, Cluzel P, Berthet M, Atassi F, Leprince P, Lebreton G, Kachenoura N, Hatem SN, Redheuil A. Differentiation and quantification of fibrosis, fat and fatty fibrosis in human left atrial myocardium using ex vivo MRI. *PloS one.* 2018;13(10):e0205104.

130. Huber AT, Lamy J, Rahhal A, Evin M, Atassi F, Defrance C, Lebreton G, Clément K, Berthet M, Isnard R, Leprince P, Cluzel P, Hatem SN, Kachenoura N, Redheuil A. Cardiac MR Strain: A Noninvasive Biomarker of Fibrofatty Remodeling of the Left Atrial Myocardium. *Radiology.* 2018;286(1):83-92.

131. Franssens BT, Nathoe HM, Leiner T, van der Graaf Y, Visseren FL. Relation between cardiovascular disease risk factors and epicardial adipose tissue density on cardiac computed tomography in patients at high risk of cardiovascular events. *European journal of preventive cardiology.* 2017;24(6):660-670.

132. Ohyama K, Matsumoto Y, Amamizu H, Uzuka H, Nishimiya K, Morosawa S, Hirano M, Watabe H, Funaki Y, Miyata S, Takahashi J, Ito K, Shimokawa H. Association of Coronary Perivascular Adipose Tissue Inflammation and Drug-Eluting Stent-Induced Coronary Hyperconstricting Responses in Pigs: (18)F-Fluorodeoxyglucose Positron Emission Tomography Imaging Study. *Arterioscler Thromb Vasc Biol.* 2017;37(9):1757-1764.

133. Whitaker J, Karády J, Karim R, Tobon-Gomez C, Fastl T, Razeghi O, O'Neill L, Decroocq M, Williams S, Corrado C, Mukherjee RK, Sim I, O'Hare D, Kotadia I, Kolossváry M, Merkely B, Littvay L, Tarnoki AD, Tarnoki DL, Voros S, Razavi R, O'Neill M, Rajani R, Maurovich Horvat P, Niederer S. Standardised computed tomographic assessment of left atrial morphology and tissue thickness in humans. *International journal of cardiology Heart & vasculature.* 2021;32:100694.

134. Barbosa JG, Figueiredo B, Bettencourt N, Tavares JM. Towards automatic quantification of the epicardial fat in non-contrasted CT images. *Computer methods in biomechanics and biomedical engineering.* 2011;14(10):905-914.

135. Shahzad R, Bos D, Metz C, Rossi A, Kirisli H, van der Lugt A, Klein S, Witteman J, de Feyter P, Niessen W, van Vliet L, van Walsum T. Automatic quantification of epicardial fat volume on non-enhanced cardiac CT scans using a multi-atlas segmentation approach. *Medical physics.* 2013;40(9):091910.

136. Commandeur F, Goeller M, Razipour A, Cadet S, Hell MM, Kwiecinski J, Chen X, Chang H-J, Marwan M, Achenbach S, Berman DS, Slomka PJ, Tamarappoo BK, Dey D. Fully Automated CT Quantification of Epicardial Adipose Tissue by Deep Learning: A Multicenter Study. *Radiology: Artificial Intelligence.* 2019;1(6):e190045.

137. Bartoli A, Fournel J, Ait-Yahia L, Cadour F, Tradi F, Ghattas B, Cortaredona S, Million M, Lasbleiz A, Dutour A, Gaborit B, Jacquier A. Automatic Deep-Learning Segmentation of Epicardial Adipose Tissue from Low-Dose Chest CT and Prognosis Impact on COVID-19. *Cells.* 2022;11(6).

138. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2016;278(2):563-577.

139. Kolossváry M, Kellermayer M, Merkely B, Maurovich-Horvat P. Cardiac Computed Tomography Radiomics: A Comprehensive Review on Radiomic Techniques. *Journal of thoracic imaging.* 2018;33(1):26-34.

140. Oikonomou EK, Siddique M, Antoniades C. Artificial intelligence in medical imaging: A radiomic guide to precision phenotyping of cardiovascular disease. *Cardiovascular research.* 2020;116(13):2040-2054.

141. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts H. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer research.* 2017;77(21):e104-e107.

142. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging.* 2012;30(9):1323-1341.

143. Guo X, Liu X, Wang H, Liang Z, Wu W, He Q, Li K, Wang W. Enhanced CT images by the wavelet transform improving diagnostic accuracy of chest nodules. *Journal of digital imaging.* 2011;24(1):44-49.

144. Oikonomou EK, Williams MC, Kotanidis CP, Desai MY, Marwan M, Antonopoulos AS, Thomas KE, Thomas S, Akoumianakis I, Fan LM, Kesavan S, Herdman L, Alashi A, Centeno EH, Lyasheva M, Griffin BP, Flamm SD, Shirodaria C, Sabharwal N, Kelion A, Dweck MR, Van Beek EJR, Deanfield J, Hopewell JC, Neubauer S, Channon KM, Achenbach S, Newby DE, Antoniades C. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. *Eur Heart J.* 2019;40(43):3529-3543.

145. Kotanidis CP, Xie C, Alexander D, Rodrigues JCL, Burnham K, Mentzer A, O'Connor D, Knight J, Siddique M, Lockstone H, Thomas S, Kotronias R, Oikonomou EK, Badi I, Lyasheva M, Shirodaria C, Lumley SF, Constantinides B, Sanderson N, Rodger G, Chau KK, Lodge A, Tsakok M, Gleeson F, Adlam D, Rao P, Indrajeet D, Deshpande A, Bajaj A, Hudson BJ, Srivastava V, Farid S, Krasopoulos G, Sayeed R, Ho LP, Neubauer S, Newby DE, Channon KM, Deanfield J, Antoniades C. Constructing custom-made radiotranscriptomic signatures of vascular inflammation from routine CT angiograms: a prospective outcomes validation study in COVID-19. *The Lancet Digital health.* 2022;4(10):e705-e716.

146. Min JK, Dunning A, Lin FY, Achenbach S, Al-Mallah MH, Berman DS, Budoff MJ, Cademartiri F, Callister TQ, Chang HJ, Cheng V, Chinnaiyan KM, Chow B, Delago A, Hadamitzky M, Hausleiter J, Karlsberg RP, Kaufmann P, Maffei E, Nasir K, Pencina MJ, Raff GL, Shaw LJ, Villines TC. Rationale and design of the CONFIRM (COronary CT Angiography EvaluatioN For Clinical Outcomes: An InteRnational Multicenter) Registry. *Journal of cardiovascular computed tomography.* 2011;5(2):84-92.

147. Douglas PS, Hoffmann U, Lee KL, Mark DB, Al-Khalidi HR, Anstrom K, Dolor RJ, Kosinski A, Krucoff MW, Mudrick DW, Patel MR, Picard MH, Udelson JE, Velazquez EJ, Cooper L.

PROspective Multicenter Imaging Study for Evaluation of chest pain: rationale and design of the PROMISE trial. *American heart journal.* 2014;167(6):796-803.e791.

148. Newby DE, Williams MC, Flapan AD, Forbes JF, Hargreaves AD, Leslie SJ, Lewis SC, McKillop G, McLean S, Reid JH, Sprat JC, Uren NG, van Beek EJ, Boon NA, Clark L, Craig P, Flather MD, McCormack C, Roditi G, Timmis AD, Krishan A, Donaldson G, Fotheringham M, Hall FJ, Neary P, Cram L, Perkins S, Taylor F, Eteiba H, Rae AP, Robb K, Barrie D, Bissett K, Dawson A, Dundas S, Fogarty Y, Ramkumar PG, Houston GJ, Letham D, O'Neill L, Pringle SD, Ritchie V, Sudarshan T, Weir-McCall J, Cormack A, Findlay IN, Hood S, Murphy C, Peat E, Allen B, Baird A, Bertram D, Brian D, Cowan A, Cruden NL, Dweck MR, Flint L, Fyfe S, Keanie C, MacGillivray TJ, Maclachlan DS, MacLeod M, Mirsadraee S, Morrison A, Mills NL, Minns FC, Phillips A, Queripel LJ, Weir NW, Bett F, Divers F, Fairley K, Jacob AJ, Keegan E, White T, Gemmill J, Henry M, McGowan J, Dinnel L, Francis CM, Sandeman D, Yerramasu A, Berry C, Boylan H, Brown A, Duffy K, Frood A, Johnstone J, Lanaghan K, MacDuff R, MacLeod M, McGlynn D, McMillan N, Murdoch L, Noble C, Paterson V, Steedman T, Tzemos N. Role of multidetector computed tomography in the diagnosis and management of patients attending the rapid access chest pain clinic, The Scottish computed tomography of the heart (SCOT-HEART) trial: study protocol for randomized controlled trial. *Trials.* 2012;13:184.

149. van Smeden M, Heinze G, Van Calster B, Asselbergs FW, Vardas PE, Bruining N, de Jaegere P, Moore JH, Denaxas S, Boulesteix AL, Moons KGM. Critical appraisal of artificial intelligence-based prediction models for cardiovascular disease *European Heart Journal.* 2022;43(31):2921-2930.

150. Investigators S-H. CT coronary angiography in patients with suspected angina due to coronary heart disease (SCOT-HEART): an open-label, parallel-group, multicentre trial. *Lancet.* 2015;385(9985):2383-2391.

151. SCOTHEART Investigators. Coronary CT Angiography and 5-Year Risk of Myocardial Infarction. *New England Journal of Medicine.* 2018;379(10):924-933.

152. Antoniades C, Demosthenous M, Reilly S, Margaritis M, Zhang M-H, Antonopoulos A, Marinou K, Nahar K, Jayaram R, Tousoulis D, Bakogiannis C, Sayeed R, Triantafyllou C, Koumallos N, Psarros C, Miliou A, Stefanadis C, Channon KM, Casadei B. Myocardial Redox State Predicts In-Hospital Clinical Outcome After Cardiac Surgery: Effects of Short-Term Pre-Operative Statin Treatment. *Journal of the American College of Cardiology.* 2012;59(1):60-70.

153. Newby DE, Adamson PD, Berry C, Boon NA, Dweck MR, Flather M, Forbes J, Hunter A, Lewis S, MacLean S, Mills NL, Norrie J, Roditi G, Shah ASV, Timmis AD, van Beek EJR, Williams MC. Coronary CT Angiography and 5-Year Risk of Myocardial Infarction. *N Engl J Med.* 2018;379(10):924-933.

154. Thygesen K, Alpert JS, White HD. Universal Definition of Myocardial Infarction. *Circulation.* 2007;116(22):2634-2653.

155. Hicks KA, Tcheng JE, Bozkurt B, Chaitman BR, Cutlip DE, Farb A, Fonarow GC, Jacobs JP, Jaff MR, Lichtman JH, Limacher MC, Mahaffey KW, Mehran R, Nissen SE, Smith EE, Targum SL. 2014 ACC/AHA Key Data Elements and Definitions for Cardiovascular Endpoint Events in Clinical Trials: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Cardiovascular Endpoints Data Standards). *J Am Coll Cardiol.* 2015;66(4):403-469.

156. Cutlip DE, Windecker S, Mehran R, Boam A, Cohen DJ, van Es GA, Steg PG, Morel MA, Mauri L, Vranckx P, McFadden E, Lansky A, Hamon M, Krucoff MW, Serruys PW. Clinical end points in coronary stent trials: a case for standardized definitions. *Circulation.* 2007;115(17):2344-2351.

157. Leipsic J, Abbara S, Achenbach S, Cury R, Earls JP, Mancini GJ, Nieman K, Pontone G, Raff GL. SCCT guidelines for the interpretation and reporting of coronary CT angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *Journal of cardiovascular computed tomography.* 2014;8(5):342-358.

158. Parekh V, Jacobs MA. Radiomics: a new application from established techniques. *Expert review of precision medicine and drug development.* 2016;1(2):207-226.

159. Haralick R, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Trans Syst Man Cybern.* 1973;SMC-3:610-621.

160. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ, Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications.* 2014;5(1):4006.

161. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Jr., Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol.* 1990;15(4):827-832.

162. Kim WH, Kim CG, Kim D-W. Optimal CT Number Range for Adipose Tissue When Determining Lean Body Mass in Whole-Body F-18 FDG PET/CT Studies. *Nucl Med Mol Imaging.* 2012;46(4):294-299.

163. Oikonomou EK, Marwan M, Desai MY, Mancio J, Alashi A, Hutt Centeno E, Thomas S, Herdman L, Kotanidis CP, Thomas KE, Griffin BP, Flamm SD, Antonopoulos AS, Shirodaria C, Sabharwal N, Deanfield J, Neubauer S, Hopewell JC, Channon KM, Achenbach S, Antoniades C. Non-invasive detection of coronary inflammation using computed tomography and prediction of residual cardiovascular risk (the CRISP CT study): a post-hoc analysis of prospective outcome data. *Lancet.* 2018;392(10151):929-939.

164. Cury RC, Abbara S, Achenbach S, Agatston A, Berman DS, Budoff MJ, Dill KE, Jacobs JE, Maroules CD, Rubin GD, Rybicki FJ, Schoepf UJ, Shaw LJ, Stillman AE, White CS, Woodard PK, Leipsic JA. CAD-RADS™: Coronary Artery Disease - Reporting and Data System: An Expert

Consensus Document of the Society of Cardiovascular Computed Tomography (SCCT), the American College of Radiology (ACR) and the North American Society for Cardiovascular Imaging (NASCI). Endorsed by the American College of Cardiology. *Journal of the American College of Radiology : JACR.* 2016;13(12 Pt A):1458-1466.e1459.

165. Cury RC, Abbara S, Achenbach S, Agatston A, Berman DS, Budoff MJ, Dill KE, Jacobs JE, Maroules CD, Rubin GD, Rybicki FJ, Schoepf UJ, Shaw LJ, Stillman AE, White CS, Woodard PK, Leipsic JA. Coronary Artery Disease - Reporting and Data System (CAD-RADS): An Expert Consensus Document of SCCT, ACR and NASCI: Endorsed by the ACC. *JACC Cardiovascular imaging.* 2016;9(9):1099-1113.

166. James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, Lackland DT, LeFevre ML, MacKenzie TD, Ogedegbe O, Smith SC, Jr, Svetkey LP, Taler SJ, Townsend RR, Wright JT, Jr, Narva AS, Ortiz E. 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults: Report From the Panel Members Appointed to the Eighth Joint National Committee (JNC 8). *JAMA.* 2014;311(5):507-520.

167. Diagnosis and classification of diabetes mellitus. *Diabetes care.* 2014;37 Suppl 1:S81-90.

168. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero ST, Smith SC, Jr., Watson K, Wilson PW. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* 2014;63(25 Pt B):2889-2934.

169. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP, Bonny A, Brauer M, Brodmann M, Cahill TJ, Carapetis J, Catapano AL, Chugh SS, Cooper LT, Coresh J, Criqui M, DeCleene N, Eagle KA, Emmons-Bell S, Feigin VL, Fernández-Solà J, Fowkes G, Gakidou E, Grundy SM, He FJ, Howard G, Hu F, Inker L, Karthikeyan G, Kassebaum N, Koroshetz W, Lavie C, Lloyd-Jones D, Lu HS, Mirijello A, Temesgen AM, Mokdad A, Moran AE, Muntner P, Narula J, Neal B, Ntsekhe M, Moraes de Oliveira G, Otto C, Owolabi M, Pratt M, Rajagopalan S, Reitsma M, Ribeiro ALP, Rigotti N, Rodgers A, Sable C, Shakil S, Sliwa-Hahnle K, Stark B, Sundström J, Timpel P, Tleyjeh IM, Valgimigli M, Vos T, Whelton PK, Yacoub M, Zuhlke L, Murray C, Fuster V, Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton A, Benjamin EJ, Benziger CP, Bonny A, Brauer M, Brodmann M, Cahill TJ, Carapetis JR, Catapano AL, Chugh S, Cooper LT, Coresh J, Criqui MH, DeCleene NK, Eagle KA, Emmons-Bell S, Feigin VL, Fernández-Sola J, Fowkes FGR, Gakidou E, Grundy SM, He FJ, Howard G, Hu F, Inker L, Karthikeyan G, Kassebaum NJ, Koroshetz WJ, Lavie C, Lloyd-Jones D, Lu HS, Mirijello A, Misganaw AT, Mokdad AH, Moran AE, Muntner P, Narula J, Neal B, Ntsekhe M, Oliveira GMM, Otto CM, Owolabi MO, Pratt M, Rajagopalan S, Reitsma MB, Ribeiro ALP, Rigotti NA, Rodgers A, Sable CA, Shakil SS, Sliwa K, Stark BA, Sundström J, Timpel P, Tleyjeh II, Valgimigli M, Vos T, Whelton PK, Yacoub M, Zuhlke LJ, Abbasi-Kangevari M, Abdi A, Abedi A, Aboyans V, Abrha WA, Abu-Gharbieh E, Abushouk AI, Acharya D, Adair T, Adebayo OM, Ademi Z, Advani SM, Afshari K, Afshin A, Agarwal G, Agasthi P, Ahmad S, Ahmadi S, Ahmed MB, Aji B, Akalu Y, Akande-Sholabi W, Aklilu A, Akunna CJ, Alahdab F, Al-Eyadhy A, Alhabib KF, Alif SM, Alipour V, Aljunid SM, Alla F, Almasi-Hashiani A, Almustanyir S, Al-Raddadi RM, Amegah AK,

Amini S, Aminorroaya A, Amu H, Amugsi DA, Ancuceanu R, Anderlini D, Andrei T, Andrei CL, Ansari-Moghaddam A, Anteneh ZA, Antonazzo IC, Antony B, Anwer R, Appiah LT, Arabloo J, Ärnlöv J, Artanti KD, Ataro Z, Ausloos M, Avila-Burgos L, Awan AT, Awoke MA, Ayele HT, Ayza MA, Azari S, B DB, Baheiraei N, Baig AA, Bakhtiari A, Banach M, Banik PC, Baptista EA, Barboza MA, Barua L, Basu S, Bedi N, Béjot Y, Bennett DA, Bensenor IM, Berman AE, Bezabih YM, Bhagavathula AS, Bhaskar S, Bhattacharyya K, Bijani A, Bikbov B, Birhanu MM, Boloor A, Brant LC, Brenner H, Briko NI, Butt ZA, Caetano dos Santos FL, Cahill LE, Cahuana-Hurtado L, Cámera LA, Campos-Nonato IR, Cantu-Brito C, Car J, Carrero JJ, Carvalho F, Castañeda-Orjuela CA, Catalá-López F, Cerin E, Charan J, Chattu VK, Chen S, Chin KL, Choi J-YJ, Chu D-T, Chung S-C, Cirillo M, Coffey S, Conti S, Costa VM, Cundiff DK, Dadras O, Dagnew B, Dai X, Damasceno AAM, Dandona L, Dandona R, Davletov K, De la Cruz-Góngora V, De la Hoz FP, De Neve J-W, Denova-Gutiérrez E, Derbew Molla M, Derseh BT, Desai R, Deuschl G, Dharmaratne SD, Dhimal M, Dhungana RR, Dianatinasab M, Diaz D, Djalalinia S, Dokova K, Douiri A, Duncan BB, Duraes AR, Eagan AW, Ebtehaj S, Eftekhari A, Eftekharzadeh S, Ekholuenetale M, El Nahas N, Elgendy IY, Elhadi M, El-Jaafary SI, Esteghamati S, Etisso AE, Eyawo O, Fadhil I, Faraon EJA, Faris PS, Farwati M, Farzadfar F, Fernandes E, Fernandez Prendes C, Ferrara P, Filip I, Fischer F, Flood D, Fukumoto T, Gad MM, Gaidhane S, Ganji M, Garg J, Gebre AK, Gebregiorgis BG, Gebregzabiher KZ, Gebremeskel GG, Getacher L, Obsa AG, Ghajar A, Ghashghaee A, Ghith N, Giampaoli S, Gilani SA, Gill PS, Gillum RF, Glushkova EV, Gnedovskaya EV, Golechha M, Gonfa KB, Goudarzian AH, Goulart AC, Guadamuz JS, Guha A, Guo Y, Gupta R, Hachinski V, Hafezi-Nejad N, Haile TG, Hamadeh RR, Hamidi S, Hankey GJ, Hargono A, Hartono RK, Hashemian M, Hashi A, Hassan S, Hassen HY, Havmoeller RJ, Hay SI, Hayat K, Heidari G, Herteliu C, Holla R, Hosseini M, Hosseinzadeh M, Hostiuc M, Hostiuc S, Househ M, Huang J, Humayun A, Iavicoli I, Ibeneme CU, Ibitoye SE, Ilesanmi OS, Ilic IM, Ilic MD, Iqbal U, Irvani SSN, Islam SMS, Islam RM, Iso H, Iwagami M, Jain V, Javaheri T, Jayapal SK, Jayaram S, Jayawardena R, Jeemon P, Jha RP, Jonas JB, Jonnagaddala J, Joukar F, Jozwiak JJ, Jürisson M, Kabir A, Kahlon T, Kalani R, Kalhor R, Kamath A, Kamel I, Kandel H, Kandel A, Karch A, Kasa AS, Katoto PDMC, Kayode GA, Khader YS, Khammarnia M, Khan MS, Khan MN, Khan M, Khan EA, Khatab K, Kibria GMA, Kim YJ, Kim GR, Kimokoti RW, Kisa S, Kisa A, Kivimäki M, Kolte D, Koolivand A, Korshunov VA, Koulmane Laxminarayana SL, Koyanagi A, Krishan K, Krishnamoorthy V, Kuate Defo B, Kucuk Bicer B, Kulkarni V, Kumar GA, Kumar N, Kurmi OP, Kusuma D, Kwan GF, La Vecchia C, Lacey B, Lallukka T, Lan Q, Lasrado S, Lassi ZS, Lauriola P, Lawrence WR, Laxmaiah A, LeGrand KE, Li M-C, Li B, Li S, Lim SS, Lim L-L, Lin H, Lin Z, Lin R-T, Liu X, Lopez AD, Lorkowski S, Lotufo PA, Lugo A, M NK, Madotto F, Mahmoudi M, Majeed A, Malekzadeh R, Malik AA, Mamun AA, Manafi N, Mansournia MA, Mantovani LG, Martini S, Mathur MR, Mazzaglia G, Mehata S, Mehndiratta MM, Meier T, Menezes RG, Meretoja A, Mestrovic T, Miazgowski B, Miazgowski T, Michalek IM, Miller TR, Mirrakhimov EM, Mirzaei H, Moazen B, Moghadaszadeh M, Mohammad Y, Mohammad DK, Mohammed S, Mohammed MA, Mokhayeri Y, Molokhia M, Montasir AA, Moradi G, Moradzadeh R, Moraga P, Morawska L, Moreno Velásquez I, Morze J, Mubarik S, Muruet W, Musa KI, Nagarajan AJ, Nalini M, Nangia V, Naqvi AA, Narasimha Swamy S, Nascimento BR, Nayak VC, Nazari J, Nazarzadeh M, Negoi RI, Neupane Kandel S, Nguyen HLT, Nixon MR, Norrving B, Noubiap JJ, Nouthe BE, Nowak C, Odukoya OO, Ogbo FA, Olagunju AT, Orru H, Ortiz A, Ostroff SM, Padubidri JR, Palladino R, Pana A, Panda-Jonas S, Parekh U, Park E-C, Parvizi M, Pashazadeh Kan F, Patel UK, Pathak M, Paudel R, Pepito VCF, Perianayagam A, Perico N, Pham HQ, Pilgrim T, Piradov MA, Pishgar F, Podder V, Polibin RV, Pourshams A, Pribadi DRA, Rabiee N, Rabiee M, Radfar A, Rafiei A, Rahim F, Rahimi-Movaghar V, Ur Rahman MH, Rahman MA, Rahmani AM, Rakovac I, Ram P, Ramalingam S, Rana J, Ranasinghe P, Rao SJ, Rathi P, Rawal L, Rawasia WF, Rawassizadeh R, Remuzzi G, Renzaho AMN, Rezapour A, Riahi SM, Roberts-Thomson RL, Roever L, Rohloff P, Romoli M, Roshandel G, Rwegerera GM, Saadatagah S, Saber-Ayad MM, Sabour S, Sacco S, Sadeghi M, Saeedi Moghaddam S, Safari S,

Sahebkar A, Salehi S, Salimzadeh H, Samaei M, Samy AM, Santos IS, Santric-Milicevic MM, Sarrafzadegan N, Sarveazad A, Sathish T, Sawhney M, Saylan M, Schmidt MI, Schutte AE, Senthilkumaran S, Sepanlou SG, Sha F, Shahabi S, Shahid I, Shaikh MA, Shamali M, Shamsizadeh M, Shawon MSR, Sheikh A, Shigematsu M, Shin M-J, Shin JI, Shiri R, Shiue I, Shuval K, Siabani S, Siddiqi TJ, Silva DAS, Singh JA, Mtech AS, Skryabin VY, Skryabina AA, Soheili A, Spurlock EE, Stockfelt L, Stortecky S, Stranges S, Suliankatchi Abdulkader R, Tadbiri H, Tadesse EG, Tadesse DB, Tajdini M, Tariqujjaman M, Teklehaimanot BF, Temsah M-H, Tesema AK, Thakur B, Thankappan KR, Thapar R, Thrift AG, Timalsina B, Tonelli M, Touvier M, Tovani-Palone MR, Tripathi A, Tripathy JP, Truelsen TC, Tsegay GM, Tsegaye GW, Tsilimparis N, Tusa BS, Tyrovolas S, Umapathi KK, Unim B, Unnikrishnan B, Usman MS, Vaduganathan M, Valdez PR, Vasankari TJ, Velazquez DZ, Venketasubramanian N, Vu GT, Vujcic IS, Waheed Y, Wang Y, Wang F, Wei J, Weintraub RG, Weldemariam AH, Westerman R, Winkler AS, Wiysonge CS, Wolfe CDA, Wubishet BL, Xu G, Yadollahpour A, Yamagishi K, Yan LL, Yandrapalli S, Yano Y, Yatsuya H, Yeheyis TY, Yeshaw Y, Yilgwan CS, Yonemoto N, Yu C, Yusefzadeh H, Zachariah G, Zaman SB, Zaman MS, Zamanian M, Zand R, Zandifar A, Zarghi A, Zastrozhin MS, Zastrozhina A, Zhang Z-J, Zhang Y, Zhang W, Zhong C, Zou Z, Zuniga YMH, Murray CJL, Fuster V. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology.* 2020;76(25):2982-3021.

170. Gulati M, Levy PD, Mukherjee D, Amsterdam E, Bhatt DL, Birtcher KK, Blankstein R, Boyd J, Bullock-Palmer RP, Conejo T, Diercks DB, Gentile F, Greenwood JP, Hess EP, Hollenberg SM, Jaber WA, Jneid H, Joglar JA, Morrow DA, O'Connor RE, Ross MA, Shaw LJ. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain. 2021;78(22):e187-e285.

171. Oikonomou Evangelos K, West Henry W, Antoniades C. Cardiac Computed Tomography. *Arteriosclerosis, Thrombosis, and Vascular Biology.* 2019;39(11):2207-2219.

172. Kvist H, Sjöström L, Tylén U. Adipose tissue volume determinations in women by computed tomography: technical considerations. *International journal of obesity.* 1986;10(1):53-67.

173. Mancio J, Azevedo D, Saraiva F, Azevedo AI, Pires-Morais G, Leite-Moreira A, Falcao-Pires I, Lunet N, Bettencourt N. Epicardial adipose tissue volume assessed by computed tomography and coronary artery disease: a systematic review and meta-analysis. *European Heart Journal - Cardiovascular Imaging.* 2017;19(5):490-497.

174. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. Paper presented at: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016; 2016//, 2016; Cham.

175. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-310.

176. Molnar D, Enqvist O, Ulén J, Larsson M, Brandberg J, Johnsson ÅA, Björnson E, Bergström G, Hjelmgren O. Artificial intelligence based automatic quantification of epicardial adipose tissue suitable for large scale population studies. *Scientific Reports.* 2021;11(1):23905.

177. Spearman JV, Meinel FG, Schoepf UJ, Apfaltrer P, Silverman JR, Krazinski AW, Canstein C, De Cecco CN, Costello P, Geyer LL. Automated quantification of epicardial adipose tissue using CT angiography: evaluation of a prototype software. *Eur Radiol.* 2014;24(2):519-526.

178. Yogananda CGB, Wagner BC, Murugesan GK, Madhuranthakam A, Maldjian JA. A Deep Learning Pipeline for Automatic Skull Stripping and Brain Segmentation. Paper presented at: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019); 8-11 April 2019, 2019.

179. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Paper presented at: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; 2015//, 2015; Cham.

180. Antoniades C, West HW. Coronary CT angiography as an 'one-stop shop' to detect the high-risk plaque and the vulnerable patient. *European Heart Journal.* 2021;42(37):3853-3855.

181. Dreisbach JG, Nicol ED, Roobottom CA, Padley S, Roditi G. Challenges in delivering computed tomography coronary angiography as the first-line test for stable chest pain. *Heart (British Cardiac Society).* 2018;104(11):921.

182. Vohl MC, Sladek R, Robitaille J, Gurd S, Marceau P, Richard D, Hudson TJ, Tchernof A. A survey of genes differentially expressed in subcutaneous and visceral adipose tissue in men. *Obesity research.* 2004;12(8):1217-1222.

183. Oikonomou EK, Siddique M, Antoniades C. Artificial intelligence in medical imaging: A radiomic guide to precision phenotyping of cardiovascular disease. *Cardiovascular research.* 2020;116(13):2040-2054.

184. Mancio J, Azevedo D, Saraiva F, Azevedo AI, Pires-Morais G, Leite-Moreira A, Falcao-Pires I, Lunet N, Bettencourt N. Epicardial adipose tissue volume assessed by computed tomography and coronary artery disease: a systematic review and meta-analysis. *European heart journal cardiovascular Imaging.* 2018;19(5):490-497.

185. Venteclef N, Guglielmi V, Balse E, Gaborit B, Cotillard A, Atassi F, Amour J, Leprince P, Dutour A, Clément K, Hatem SN. Human epicardial adipose tissue induces fibrosis of the atrial myocardium through the secretion of adipo-fibrokines. *Eur Heart J.* 2015;36(13):795-805a.

186. Haemers P, Hamdi H, Guedj K, Suffee N, Farahmand P, Popovic N, Claus P, LePrince P, Nicoletti A, Jalife J, Wolke C, Lendeckel U, Jaïs P, Willems R, Hatem SN. Atrial fibrillation is associated with the fibrotic remodelling of adipose tissue in the subepicardium of human and sheep atria. *Eur Heart J.* 2017;38(1):53-61.

187. Pérez-Belmonte LM, Moreno-Santos I, Gómez-Doblas JJ, García-Pinilla JM, Morcillo-Hidalgo L, Garrido-Sánchez L, Santiago-Fernández C, Crespo-Leiro MG, Carrasco-Chinchilla F, Sánchez-Fernández PL, de Teresa-Galván E, Jiménez-Navarro M. Expression of epicardial

adipose tissue thermogenic genes in patients with reduced and preserved ejection fraction heart failure. *Int J Med Sci.* 2017;14(9):891-895.

188. Salatzki J, Foryst-Ludwig A, Bentele K, Blumrich A, Smeir E, Ban Z, Brix S, Grune J, Beyhoff N, Klopfleisch R, Dunst S, Surma MA, Klose C, Rothe M, Heinzel FR, Krannich A, Kershaw EE, Beule D, Schulze PC, Marx N, Kintscher U. Adipose tissue ATGL modifies the cardiac lipidome in pressure-overload-induced left ventricular failure. *PLoS Genet.* 2018;14(1):e1007171-e1007171.

189. SCOTHEART Investigators. CT coronary angiography in patients with suspected angina due to coronary heart disease (SCOT-HEART): an open-label, parallel-group, multicentre trial. *Lancet.* 2015;385(9985):2383-2391.

190. Du Bois D, Du Bois EF. A formula to estimate the approximate surface area if height and weight be known. 1916. *Nutrition (Burbank, Los Angeles County, Calif).* 1989;5(5):303-311; discussion 312-303.

191. Goeller M, Achenbach S, Marwan M, Doris MK, Cadet S, Commandeur F, Chen X, Slomka PJ, Gransar H, Cao JJ, Wong ND, Albrecht MH, Rozanski A, Tamarappoo BK, Berman DS, Dey D. Epicardial adipose tissue density and volume are related to subclinical atherosclerosis, inflammation and major adverse cardiac events in asymptomatic subjects. *Journal of cardiovascular computed tomography.* 2018;12(1):67-73.

192. Nakanishi R, Rajani R, Cheng VY, Gransar H, Nakazato R, Shmilovich H, Otaki Y, Hayes SW, Thomson LE, Friedman JD, Slomka PJ, Berman DS, Dey D. Increase in epicardial fat volume is associated with greater coronary artery calcification progression in subjects at intermediate risk by coronary calcium score: a serial study using non-contrast cardiac CT. *Atherosclerosis.* 2011;218(2):363-368.

193. Antoniades C. 'Dysfunctional' adipose tissue in cardiovascular disease: a reprogrammable target or an innocent bystander? *Cardiovascular research.* 2017;113(9):997-998.

194. Shen MJ, Arora R, Jalife J. Atrial Myopathy. *JACC Basic Transl Sci.* 2019;4(5):640-654.

195. Lin CK, McDonough RJ, Prentice RL, Thomas DM, Steel KE, Rubal BJ, Shry EA, Villines TC, Hulten EA, Slim AM. Assessment of major adverse cardiovascular events and ischemic stroke with coronary computed tomography angiography based upon angiographic diagnosis in a high-volume single center. *SAGE open medicine.* 2014;2:2050312114533535.

196. Office of National Statistics. National identity, England and Wales: Census 2021. 2022. Accessed 13 November, 2022.

197. Cuzick J. A Wilcoxon-type test for trend. *Statistics in medicine.* 1985;4(1):87-90.

198. Office of National Statistics. Ethnic groups: UK Census 2021. 2022; https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021. Accessed 12 November, 2022.

199. Strengthening the UK's Capability in Health Informatics Research Conference Report. 2013.

200. Mildenberger P, Eichelberg M, Martin E. Introduction to the DICOM standard. *Eur Radiol.* 2002;12(4):920-927.

201. van der Bijl N, de Bruin PW, Geleijns J, Bax JJ, Schuijf JD, de Roos A, Kroft LJ. Assessment of coronary artery calcium by using volumetric 320-row multi-detector computed tomography: comparison of 0.5 mm with 3.0 mm slice reconstructions. *Int J Cardiovasc Imaging.* 2010;26(4):473-482.

202. Berenguer R, Pastor-Juan MDR, Canales-Vázquez J, Castro-García M, Villas MV, Mansilla Legorburo F, Sabater S. Radiomics of CT Features May Be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. *Radiology.* 2018;288(2):407-415.

203. Iacobellis G. Magnetic Resonance Imaging of the Epicardial Adipose Tissue. In: *Epicardial Adipose Tissue.* Cham: Springer International Publishing; 2020:71-76.

204. Iacobellis G, Assael F, Ribaudo MC, Zappaterreno A, Alessi G, Di Mario U, Leonetti F. Epicardial fat from echocardiography: a new method for visceral adipose tissue prediction. *Obesity research.* 2003;11(2):304-310.

205. Fabritz L, Obergassel J. Artificial intelligence for early atrial fibrillation detection. *The Lancet.* 2022;400(10359):1173-1175.

206. Durán JM, Jongsma KR. Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *Journal of medical ethics.* 2021.

207. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Paper presented at: 2017 IEEE International Conference on Computer Vision (ICCV); 22-29 Oct. 2017, 2017.

208. Johnson KB, Wei WQ, Weeraratne D, Frisse ME, Misulis K, Rhee K, Zhao J, Snowdon JL. Precision Medicine, AI, and the Future of Personalized Health Care. *Clinical and translational science.* 2021;14(1):86-93.

209. Liu Z, Wang S, Wang Y, Zhou N, Shu J, Stamm C, Jiang M, Luo F. Association of epicardial adipose tissue attenuation with coronary atherosclerosis in patients with a high risk of coronary artery disease. *Atherosclerosis.* 2019;284:230-236.

210. Ciucci G, Colliva A, Vuerich R, Pompilio G, Zacchigna S. Biologics and cardiac disease: challenges and opportunities. *Trends in Pharmacological Sciences.* 2022;43(11):894-905.

211. van Smeden M, Heinze G, Van Calster B, Asselbergs FW, Vardas PE, Bruining N, de Jaegere P, Moore JH, Denaxas S, Boulesteix AL, Moons KGM. Critical appraisal of artificial intelligence-based prediction models for cardiovascular disease. *Eur Heart J.* 2022;43(31):2921-2930.

212. Rao B, Zohrabian V, Cedeno P, Saha A, Pahade J, Davis MA. Utility of Artificial Intelligence Tool as a Prospective Radiology Peer Reviewer — Detection of Unreported Intracranial Hemorrhage. *Academic Radiology.* 2021;28(1):85-93.

213. Krishnan A, Chilton E, Raman J, Saxena P, McFarlane C, Trollope AF, Kinobe R, Chilton L. Are Interactions between Epicardial Adipose Tissue, Cardiac Fibroblasts and Cardiac Myocytes Instrumental in Atrial Fibrosis and Atrial Fibrillation? *Cells.* 2021;10(9).

214. Pugliese NR, Paneni F, Mazzola M, De Biase N, Del Punta L, Gargani L, Mengozzi A, Virdis A, Nesti L, Taddei S, Flammer A, Borlaug BA, Ruschitzka F, Masi S. Impact of epicardial adipose tissue on cardiovascular haemodynamics, metabolic profile, and prognosis in heart failure. *European journal of heart failure.* 2021;23(11):1858-1871.

215. Bittner Daniel O, Mayrhofer T, Budoff M, Szilveszter B, Foldyna B, Hallett Travis R, Ivanov A, Janjua S, Meyersohn Nandini M, Staziaki Pedro V, Achenbach S, Ferencik M, Douglas Pamela S, Hoffmann U, Lu Michael T. Prognostic Value of Coronary CTA in Stable Chest Pain. *JACC: Cardiovascular Imaging.* 2020;13(7):1534-1545.

11

Appendix

Appendix 1. Radiomic definitions

(Adapted from http://pyradiomics.readthedocs.io/en/latest/features.html – downloaded April 22, 2022)

The definitions of the radiomic features have been derived from the pyradiomics library and the pyradiomics community maintains the copyright for the definitions mentioned below (© Copyright 2016, pyradiomics community, http://github.com/radiomics/pyradiomics, last accessed April 22, 2022).

The pyradiomics package is covered by the open source 3-clause BSD License (© 2022 Harvard Medical School: https://github.com/Radiomics/pyradiomics/blob/master/LICENSE.txt)

In my analysis, I measured a total of 872 radiomic features in CT segmentations.

a. First Order Statistics

These statistics describe the central tendency, variability, uniformity, asymmetry, skewness and magnitude of the attenuation values in a given region of interest (ROI), disregarding the spatial relationship of the individual voxels. As such, they describe quantitative and qualitative features of the whole ROI. A total of 19 features were calculated for each one of the eight wavelet transformations and the original CT image, as follows:

Let:

- $\boldsymbol{X}$ be a set of $\boldsymbol{N_p}$ voxels included in the region of interest (ROI)
- $\boldsymbol{P(i)}$ be the first order histogram with $\boldsymbol{N_g}$ discrete intensity levels, where $\boldsymbol{N_g}$ is the number of non-zero bins, equally spaced from 0 with a width.
- $\boldsymbol{p(i)}$ be the normalised first order histogram and equal to $\frac{\mathbf{P}(i)}{N_p}$
- **c** is a value that shifts the intensities to prevent negative values in X. This ensures that voxels with the lowest gray values contribute the least to Energy, instead of voxels with gray level intensity closest to 0. Since the HU range of AT (-190 to -30 HU) does not include zero, c was set at c=0. Therefore, higher energy corresponds to less radiodense AT, and therefore a higher lipophilic content.
- $\boldsymbol{\epsilon}$ is an arbitrarily small positive number (≈2.2×10^{-16})

Appendix Table A2. First-order radiomic features for peri-atrial segmentation

Radiomic feature	Interpretation
$\mathbf{Energy} = \sum_{i=1}^{N_p}(\mathrm{X}(i)+c)^2$	**Energy** is a measure of the magnitude of voxel values in an image. A larger value implies a greater sum of the squares of these values.

Total Energy $= V_{voxel}\sum_{i=1}^{N_p}(\mathrm{X}(i)+c)^2$	**Total Energy** is the value of Energy feature scaled by the volume of the voxel in cubic mm.
Entropy $= -\sum_{i=1}^{N_g} p(i)\log_2(p(i)+\epsilon)$	**Entropy** specifies the uncertainty/randomness in the image values. It measures the average amount of information required to encode the image values
Minimum $= min(\mathrm{X})$	The **minimum** gray level intensity within the ROI.
The **10th percentile** of X	The **10th percentile** of X
The **90th percentile** of X	The **90th percentile** of X
Maximum $= max(\mathrm{X})$	The **maximum** gray level intensity within the ROI.
Mean $= \frac{1}{N_p}\sum_{i=1}^{N_p}\mathrm{X}(i)$	The average (**mean**) gray level intensity within the ROI.
Median	The **median** gray level intensity within the ROI.
Interquartile range $= \mathrm{P}_{75}-\mathrm{P}_{25}$	Here P_{25} and P_{75} are the 25th and 75th percentile of the image array, respectively.
Range $= max(\mathrm{X})-min(\mathrm{X})$	The range of gray values in the ROI.
MAD $= \frac{1}{N_p}\sum_{i=1}^{N_p}\lvert\mathrm{X}(i)-\overline{X}\rvert$	**Mean Absolute Deviation (MAD)** is the mean distance of all intensity values from the Mean Value of the image array.
rMAD $= \frac{1}{N_{10-90}}\sum_{i=1}^{N_{10-90}}\lvert\mathrm{X}_{10-90}(i)-\overline{X}_{10-90}\rvert$	**Robust Mean Absolute Deviation (rMAD)** is the mean distance of all intensity values from the Mean Value calculated on the subset of image array with gray levels in between, or equal to the 10th and 90th percentile.
RMS $= \sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathrm{X}(i)+c)^2}$	**Root Mean Squared (RMS)** is the square-root of the mean of all the squared intensity values. It is another measure of the magnitude of the image values. This feature is volume-confounded, a larger value of c increases the effect of volume-confounding.
Skewness $= \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathrm{X}(i)-\overline{X})^3}{(\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathrm{X}(i)-\overline{X})^2})^3}$	**Skewness** measures the *asymmetry* of the distribution of values about the Mean value. Depending on where the tail is elongated and the mass of the distribution is concentrated, this value can be positive or negative. (Where μ^3 is the 3rd central moment).
Kurtosis $= \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathrm{X}(i)-\overline{X})^4}{(\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathrm{X}(i)-\overline{X})^2)^2}$	**Kurtosis** is a measure of the '*peakedness*' of the distribution of values in the image ROI. A higher kurtosis implies that the mass of the distribution is concentrated towards the tail(s) rather than towards the mean. A lower kurtosis implies the reverse: that the mass of the distribution is concentrated towards a spike near the Mean

	value. (Where μ4 is the 4th central moment).
Variance $= \frac{1}{N_p}\sum_{i=1}^{N_p}(X(i)-\overline{X})^2$	**Variance** is the the mean of the squared distances of each intensity value from the Mean value. This is a measure of the spread of the distribution about the mean.
Uniformity $= \sum_{i=1}^{N_g} p(i)^2$	**Uniformity** is a measure of the sum of the squares of each intensity value. This is a measure of the heterogeneity of the image array, ***where a greater uniformity implies a greater heterogeneity or a greater range of discrete intensity values***.

b. Shape-related Statistics

Shape-related statistics describe the size and shape of a given ROI, without taking into account the attenuation values of its voxels. Since they are independent of the gray level intensities, shape-related statistics were consistent across all wavelet transformation and the original CT image, and therefore were only calculated once. These were defined as follows:

Let:

V the volume of the ROI in mm^3

A the surface area of the ROI in mm^2

Appendix Table A3. First-shape related radiomic features for peri-atrial segmentation

Radiomic feature	Interpretation
$\textbf{Volume} = \sum_{i=1}^{N} V_i$	The **volume** of the ROI V is approximated by multiplying the number of voxels in the ROI by the volume of a single voxel $\mathbf{V_i}$.
$\textbf{Surface Area} = \sum_{i=1}^{N} \frac{1}{2}\lvert a_i b_i \times a_i c_i \rvert$	**Surface Area** is an approximation of the surface of the ROI in mm^2, calculated using a marching cubes algorithm, where ***N*** is the number of triangles forming the surface mesh of the volume (ROI), $\boldsymbol{a_i b_i}$ and $\boldsymbol{a_i c_i}$ are the edges of the $\boldsymbol{i^{th}}$ triangle formed by points $\boldsymbol{a_i}$, $\boldsymbol{b_i}$ and $\boldsymbol{c_i}$.
Surface to volume ratio $= \frac{A}{V}$	Here, a lower value indicates a more compact (sphere-like) shape. This feature is not dimensionless, and is therefore (partly) dependent on the volume of the ROI.
Sphericity $= \frac{\sqrt[3]{36\pi V^2}}{A}$	**Sphericity** is a measure of the roundness of the shape of the tumor region relative to a sphere. It is a dimensionless measure, independent of scale and orientation. The value range is 0<sphericity≤1, where a value of 1 indicates a perfect sphere (a sphere has the smallest possible surface area for a given volume, compared to other solids).
Volume Number	Total number of discrete volumes in the ROI.
Voxel Number	Total number of discrete voxels in the ROI.

Maximum 3D diameter	Maximum 3D diameter is defined as the largest pairwise Euclidean distance between surface voxels in the ROI (Feret Diameter).
Maximum 2D diameter (Slice)	Maximum 2D diameter (Slice) is defined as the largest pairwise Euclidean distance between ROI surface voxels in the row-column (generally the axial) plane.
Maximum 2D diameter (Column)	Maximum 2D diameter (Column) is defined as the largest pairwise Euclidean distance between ROI surface voxels in the row-slice (usually the coronal) plane.
Maximum 2D diameter (Row)	Maximum 2D diameter (Row) is defined as the largest pairwise Euclidean distance between tumor surface voxels in the column-slice (usually the sagittal) plane.
Major axis $= 4\sqrt{\lambda_{major}}$	
Minor axis $= 4\sqrt{\lambda_{minor}}$	
Least axis $= 4\sqrt{\lambda_{least}}$	
Elongation $= \sqrt{\frac{\lambda_{minor}}{\lambda_{major}}}$	Here, $\boldsymbol{\lambda_{major}}$ and $\boldsymbol{\lambda_{minor}}$ are the lengths of the largest and second largest principal component axes. The values range between 1 (circle-like (non-elongated)) and 0 (single point or 1 dimensional line).
Flatness $= \sqrt{\frac{\lambda_{least}}{\lambda_{major}}}$	Here, $\boldsymbol{\lambda_{major}}$ and $\boldsymbol{\lambda_{minor}}$ are the lengths of the largest and smallest principal component axes. The values range between 1 (non-flat, sphere-like) and 0 (a flat object).

Adapted from http://pyradiomics.readthedocs.io/en/latest/features.html – last accessed March 22, 2018. © Copyright 2016, pyradiomics community, http://github.com/radiomics/pyradiomics Revision eae15eff.

c. Gray Level Co-occurrence Matrix (GLCM)

In simple words, a GLCM describes the number of times a voxel of a given attenuation value ***i*** is located next to a voxel of ***j***. A GLCM of size $\boldsymbol{N_g \times N_g}$ describes the second-order joint probability function of an image region constrained by the mask and is defined as $P(\boldsymbol{i,j}|\boldsymbol{\delta,\vartheta})$. The $(\boldsymbol{i,j})^{th}$ element of this matrix represents the number of times the combination of levels ***i*** and ***j*** occur in two pixels in the image, that are separated by a distance of $\boldsymbol{\delta}$ pixels along angle $\boldsymbol{\vartheta}$. The distance $\boldsymbol{\delta}$ from the center voxel is defined as the distance according to the infinity norm. For $\boldsymbol{\delta=1}$, this results in 2 neighbors for each of 13 angles in 3D (26-connectivity) and for $\boldsymbol{\delta=2}$ a 98-connectivity (49 unique angles). In order to get rotationally invariant results, statistics are calculated in all directions and then averaged, to ensure a symmetrical GLCM.

Let:

$\boldsymbol{\epsilon}$ be an arbitrarily small positive number **(≈2.2×10−16)**

P(*i,j*) be the co-occurence matrix for an arbitrary $\boldsymbol{\delta}$ and $\boldsymbol{\vartheta}$

p(*i,j*) be the normalised co-occurence matrix and equal to $\frac{\mathrm{P}(i,j)}{\sum \mathrm{P}(i,j)}$

Ng be the number of discrete intensity levels in the image

$\boldsymbol{p_x(i)} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{N_g}} \boldsymbol{P(i,j)}$ be the marginal row probabilities

$\boldsymbol{p_y(j)} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N_g}} \boldsymbol{P(i,j)}$ be the marginal column probabilities

$\boldsymbol{\mu_x}$ be the mean gray level intensity of $\boldsymbol{p_x}$ and defined as $\boldsymbol{\mu_x} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N_g}} \boldsymbol{p_x(i)i}$

$\boldsymbol{\mu_y}$ be the mean gray level intensity of $\boldsymbol{p_y}$ and defined as $\boldsymbol{\mu_y} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{N_g}} \boldsymbol{p_y(j)j}$

$\boldsymbol{\sigma_x}$ be the standard deviation of $\boldsymbol{p_x}$

$\boldsymbol{\sigma_y}$ be the standard deviation of $\boldsymbol{p_y}$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \text{ , where } i+j=k, \text{ and } k = 2,3,\dots,2N_g$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \text{ , where } |i-j| = k, \text{ and } k = 0,1,\dots,N_g - 1$$

$HX = -\sum_{i=1}^{N_g} p_x(i)\log_2\left(p_x(i) + \epsilon\right)$ be the entropy of $\boldsymbol{p_x}$

$HY = -\sum_{j=1}^{N_g} p_y(j)\log_2\left(p_y(j) + \epsilon\right)$ be the entropy of $\boldsymbol{p_y}$

$$HXY1 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)\log_2\left(p_x(i)p_y(j) + \epsilon\right)$$

$$HXY2 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j)\log_2\left(p_x(i)p_y(j) + \epsilon\right)$$

For distance weighting, GLCM matrices are weighted by weighting factor W and then summed and normalised.. Weighting factor W is calculated for the distance between neighbouring voxels by $\boldsymbol{W = e^{-\|d\|^2}}$, where d is the distance for the associated angle.

Appendix Table A4. Gray Level Co-occurrence Matrix (GLCM) for peri-atrial segmentation

Radiomic feature	Interpretation
Autocorrelation $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)ij$	**Autocorrelation** is a measure of the magnitude of the fineness and coarseness of texture.

Joint average $= \mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)i$	Returns the mean gray level intensity of the i distribution.
Cluster prominence $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^4 p(i,j)$	**Cluster Prominence** is a measure of the skewness and asymmetry of the GLCM. A higher value implies more asymmetry around the mean while a lower value indicates a peak near the mean value and less variation around the mean.
Cluster tendency $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^2 p(i,j)$	**Cluster Tendency** is a measure of groupings of voxels with similar gray-level values.
Cluster shade $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^3 p(i,j)$	**Cluster Shade** is a measure of the skewness and uniformity of the GLCM. A higher cluster shade implies greater asymmetry about the mean.
Contrast $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i,j)$	**Contrast** is a measure of the local intensity variation, favoring values away from the diagonal (i=j). A larger value correlates with a greater disparity in intensity values among neighbouring voxels.
Correlation $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)ij - \mu_x \mu_y}{\sigma_x(i)\sigma_y(j)}$	**Correlation** is a value between 0 (uncorrelated) and 1 (perfectly correlated) showing the linear dependency of gray level values to their respective voxels in the GLCM.
Difference average $= \sum_{k=0}^{N_g - 1} k p_{x-y}(k)$	**Difference Average** measures the relationship between occurrences of pairs with similar intensity values and occurrences of pairs with differing intensity values.
Difference entropy $= \sum_{k=0}^{N_g - 1} p_{x-y}(k) \log_2 (p_{x-y}(k) + \epsilon)$	**Difference Entropy** is a measure of the randomness/variability in neighborhood intensity value differences.
Difference variance $= \sum_{k=0}^{N_g - 1} (k - DA)^2 p_{x-y}(k)$	**Difference Variance** is a measure of heterogeneity that places higher weights on differing intensity level pairs that deviate more from the mean.
Joint energy $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i,j))^2$	**Joint energy** is a measure of homogeneous patterns in the image. A greater **joint energy** implies that there are more instances of intensity value pairs in the image that neighbor each other at higher frequencies. (also known as Angular Second

	Moment).
Joint entropy $= -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log_2\left(p(i,j)+\epsilon\right)$	**Joint entropy** is a measure of the randomness/variability in neighborhood intensity values.
IMC 1 $= \frac{HXY - HXY1}{\max\{HX, HY\}}$	**Informational measure of correlation 1**
IMC 2 $= \sqrt{1 - e^{-2(HXY2 - HXY)}}$	**Informational measure of correlation 2**
IDM $= \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{p(i,j)}{1+\lvert i-j\rvert^2}$	IDM (**inverse difference moment** a.k.a Homogeneity 2) is a measure of the local homogeneity of an image. IDM weights are the inverse of the Contrast weights (decreasing exponentially from the diagonal i=j in the GLCM).
IDMN $= \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{p(i,j)}{1+(\frac{\lvert i-j\rvert^2}{N_g^2})}$	IDMN (**inverse difference moment normalised**) is a measure of the local homogeneity of an image. IDMN weights are the inverse of the Contrast weights (decreasing exponentially from the diagonal i=j in the GLCM). Unlike Homogeneity2, IDMN normalises the square of the difference between neighbouring intensity values by dividing over the square of the total number of discrete intensity values.
ID $= \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{p(i,j)}{1+\lvert i-j\rvert}$	ID (**inverse difference** a.k.a. Homogeneity 1) is another measure of the local homogeneity of an image. With more uniform gray levels, the denominator will remain low, resulting in a higher overall value.
IDN $= \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{p(i,j)}{1+(\frac{\lvert i-j\rvert}{N_g})}$	IDN (**inverse difference normalised**) is another measure of the local homogeneity of an image. Unlike Homogeneity1, IDN normalises the difference between the neighbouring intensity values by dividing over the total number of discrete intensity values.
Inverse variance $= \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{p(i,j)}{\lvert i-j\rvert^2}, i \neq j$	
Maximum probability $= max(p(i,j))$	**Maximum Probability** is occurrences of the most predominant pair of neighbouring intensity values (also known as Joint maximum).

Sum average $= \sum_{k=2}^{2N_g} p_{x+y}(k)k$	**Sum Average** measures the relationship between occurrences of pairs with lower intensity values and occurrences of pairs with higher intensity values.
Sum entropy $= \sum_{k=2}^{2N_g} p_{x+y}(k)\log_2\left(p_{x+y}(k)+\epsilon\right)$	**Sum Entropy** is a sum of neighborhood intensity value differences.
Sum squares $= \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu_x)^2 p(i,j)$	**Sum of Squares** or Variance is a measure in the distribution of neigboring intensity level pairs about the mean intensity level in the GLCM. (Defined by IBSI as Joint Variance).

Adapted from http://pyradiomics.readthedocs.io/en/latest/features.html – last accessed March 22, 2018. © Copyright 2016, pyradiomics community, http://github.com/radiomics/pyradiomics Revision eae15eff.

d. Gray Level Size Zone Matrix (GLSZM)

A Gray Level Size Zone (GLSZM) describes gray level zones in a ROI, which are defined as the number of connected voxels that share the same gray level intensity. A voxel is considered connected if the distance is 1 according to the infinity norm (26-connected region in a 3D, 8-connected region in 2D). In a gray level size zone matrix ***P(i,j)*** the **(i,j)th** element equals the number of zones with gray level ***i*** and size ***j*** appear in image. Contrary to GLCM and GLRLM, the GLSZM is rotation independent, with only one matrix calculated for all directions in the ROI.

Let:

$\mathbf{N_g}$ be the number of discreet intensity values in the image

$\mathbf{N_s}$ be the number of discreet zone sizes in the image

$\mathbf{N_p}$ be the number of voxels in the image

$\mathbf{N_z}$ be the number of zones in the ROI, which is equal to $\sum_{i=1}^{N_g} \quad \sum_{j=1}^{N_s} \mathrm{P}(i,j)$ and $\boldsymbol{1 \leq N_z \leq N_p}$

P(i,j) be the size zone matrix

p(i,j) be the normalised size zone matrix, defined as $p(i,j) = \frac{\mathrm{P}(i,j)}{N_z}$

$\boldsymbol{\epsilon}$ is an arbitrarily small positive number (≈2.2×10−16).

Appendix Table A5. Gray Level Size Zone Matrix (GLSZM) statistics for peri-atrial segmentation

Radiomic feature	Interpretation
SAE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathrm{P}(i,j)}{j^2}}{N_z}$	**SAE (small area emphasis)** is a measure of the distribution of small size zones, with a greater value indicative of smaller size zones and more

	fine textures.
LAE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathrm{P}(i,j) j^2}{N_z}$	**LAE (large area emphasis)** is a measure of the distribution of large area size zones, with a greater value indicative of larger size zones and more coarse textures.
GLN $= \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} \mathrm{P}(i,j))^2}{N_z}$	**GLN (gray level non-uniformity)** measures the variability of gray-level intensity values in the image, with a lower value indicating more homogeneity in intensity values.
GLNN $= \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} \mathrm{P}(i,j))^2}{N_z^2}$	**GLNN (gray level non-uniformity normalised)** measures the variability of gray-level intensity values in the image, with a lower value indicating a greater similarity in intensity values. This is the normalised version of the GLN formula.
SZN $= \frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} \mathrm{P}(i,j))^2}{N_z}$	**SZN (size zone non-uniformity)** measures the variability of size zone volumes in the image, with a lower value indicating more homogeneity in size zone volumes.
SZNN $= \frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} \mathrm{P}(i,j))^2}{N_z^2}$	**SZNN (size zone non-uniformity normalised)** measures the variability of size zone volumes throughout the image, with a lower value indicating more homogeneity among zone size volumes in the image. This is the normalised version of the SZN formula.
Zone Percentage $= \frac{N_z}{N_p}$	**ZP (Zone Percentage)** measures the coarseness of the texture by taking the ratio of number of zones and number of voxels in the ROI. Values are in range $\frac{1}{N_p} \leq ZP \leq 1$, with higher values indicating a larger portion of the ROI consists of small zones (indicates a more fine texture).
GLV $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)(i-\mu)^2,$ where $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) i$	**Gray level variance (GLV)** measures the variance in gray level intensities for the zones.
ZV $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)(j-\mu)^2,$ where $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) j$	**Zone Variance (ZV)** measures the variance in zone size volumes for the zones.
ZE $= -\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) \log_2 (p(i,j) + \epsilon)$	**Zone Entropy (ZE)** measures the uncertainty/randomness in the distribution of zone sizes and gray levels. A higher value indicates more heterogeneneity in the texture patterns.
LGLZE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathrm{P}(i,j)}{i^2}}{N_z}$	**LGLZE (low gray level zone emphasis)** measures the distribution of lower gray-level size zones, with a higher value indicating a greater proportion of

	lower gray-level values and size zones in the image.
HGLZE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathrm{P}(i,j) i^2}{N_z}$	**HGLZE (high gray level zone emphasis)** measures the distribution of the higher gray-level values, with a higher value indicating a greater proportion of higher gray-level values and size zones in the image.
SALGLE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathrm{P}(i,j)}{i^2 j^2}}{N_z}$	**SALGLE (small area low gray level emphasis)** measures the proportion in the image of the joint distribution of smaller size zones with lower gray-level values.
SAHGLE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathrm{P}(i,j) i^2}{j^2}}{N_z}$	**SAHGLE (small area high gray level emphasis)** measures the proportion in the image of the joint distribution of smaller size zones with higher gray-level values.
LALGLE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathrm{P}(i,j) j^2}{i^2}}{N_z}$	**LALGLE (low area low gray level emphasis)** measures the proportion in the image of the joint distribution of larger size zones with lower gray-level values.
LAHGLE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathrm{P}(i,j) i^2 j^2}{N_z}$	**LAHGLE (low area high gray level emphasis)** measures the proportion in the image of the joint distribution of larger size zones with higher gray-level values.

Adapted from http://pyradiomics.readthedocs.io/en/latest/features.html – last accessed March 22, 2018.
© Copyright 2016, pyradiomics community, http://github.com/radiomics/pyradiomics Revision eae15eff.

e. Gray Level Run Length Matrix (GLRLM)

A Gray Level Run Length Matrix (GLRLM) describes gray level runs, which are defined as the length in number of pixels, of consecutive pixels that have the same gray level value. In a gray level run length matrix ***P(i,j|ϑ)***, the ***(i,j)th*** element describes the number of runs with gray level ***i*** and length ***j*** occur in the image (ROI) along angle ***ϑ***.

Let:

N_g be the number of discreet intensity values in the image

N_r be the number of discreet run lengths in the image

N_p be the number of voxels in the image

$N_z(\vartheta)$ be the number of runs in the image along angle θ, which is equal to $\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathrm{P}(i,j|\theta)$and $1 \leq N_z(\theta) \leq N_p$

P(i,j|ϑ) be the run length matrix for an arbitrary direction θ

p(i,j|ϑ) be the normalised run length matrix, defined as $p(i,j|\theta) = \frac{\mathrm{P}(i,j|\theta)}{N_z(\theta)}$

ϵ is an arbitrarily small positive number (≈2.2×10−16).

By default, the value of a feature is calculated on the GLRLM for each angle separately, after which the mean of these values is returned. If distance weighting is enabled, GLRLMs are weighted by the distance between neighbouring voxels and then summed and normalised. Features are then calculated on the resultant matrix. The distance between neighbouring voxels is calculated for each angle using the norm specified in 'weightingNorm'.

Appendix Table A6. Gray Level Run Length Matrix (GLRLM) statistics for for peri-atrial segmentation

Radiomic feature	Interpretation
$\textbf{SRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathrm{P}(i,j\|\theta)}{j^2}}{N_z(\theta)}$	**SRE (Short Run Emphasis)** is a measure of the distribution of short run lengths, with a greater value indicative of shorter run lengths and more fine textural textures.
$\textbf{LRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathrm{P}(i,j\|\theta) j^2}{N_z(\theta)}$	**LRE (Long Run Emphasis)** is a measure of the distribution of long run lengths, with a greater value indicative of longer run lengths and more coarse structural textures.
$\textbf{GLN} = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} \mathrm{P}(i,j\|\theta))^2}{N_z(\theta)}$	**GLN (Gray Level Non-uniformity)** measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values.
$\textbf{GLNN} = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} \mathrm{P}(i,j\|\theta))^2}{N_z(\theta)^2}$	**GLNN (Gray Level Non-uniformity Normalised)** measures the similarity of gray-level intensity values in the image, where a lower GLNN value correlates with a greater similarity in intensity values. This is the normalised version of the GLN formula.
$\textbf{RLN} = \frac{\sum_{j=1}^{N_r} (\sum_{i=1}^{N_g} \mathrm{P}(i,j\|\theta))^2}{N_z(\theta)}$	**RLN (Run Length Non-uniformity)** measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image.
$\textbf{RLNN} = \frac{\sum_{j=1}^{N_r} (\sum_{i=1}^{N_g} \mathrm{P}(,j\|\theta))^2}{N_z(\theta)^2}$	**RLNN (Run Length Non-uniformity)** measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image. This is the normalised version of the RLN formula.
$\textbf{RP} = \frac{N_z(\theta)}{N_p}$	**RP (Run Percentage)** measures the coarseness of the texture by taking the ratio of number of runs and number of

	voxels in the ROI. Values are in range $\frac{1}{N_p} \leq RP \leq 1$, with higher values indicating a larger portion of the ROI consists of short runs (indicates a more fine texture).
GLV $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j\|\theta)(i-\mu)^2,$ where $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j\|\theta) i$	**GLV (Gray Level Variance)** measures the variance in gray level intensity for the runs.
RV $= \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j\|\theta)(j-\mu)^2,$ where $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j\|\theta) j$	**RV (Run Variance)** is a measure of the variance in runs for the run lengths.
RE $= -\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j\|\theta) \log_2 (p(i,j\|\theta) + \epsilon)$	**RE (Run Entropy)** measures the uncertainty/randomness in the distribution of run lengths and gray levels. A higher value indicates more heterogeneity in the texture patterns.
LGLRE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathrm{P}(i,j\|\theta)}{i^2}}{N_z(\theta)}$	**LGLRE (low gray level run emphasis)** measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.
HGLRE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathrm{P}(i,j\|\theta) i^2}{N_z(\theta)}$	**HGLRE (high gray level run emphasis)** measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.
SRLGLE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathrm{P}(i,j\|\theta)}{i^2 j^2}}{N_z(\theta)}$	**SRLGLE (short run low gray level emphasis)** measures the joint distribution of shorter run lengths with lower gray-level values.
SRHGLE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathrm{P}(i,j\|\theta) i^2}{j^2}}{N_z(\theta)}$	**SRHGLE (short run high gray level emphasis)** measures the joint distribution of shorter run lengths with higher gray-level values.
LRLGLRE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathrm{P}(i,j\|\theta) j^2}{i^2}}{N_z(\theta)}$	**LRLGLRE (long run low gray level emphasis)** measures the joint distribution of long run lengths with lower gray-level values.
LRHGLRE $= \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathrm{P}(i,j\|\theta) i^2 j^2}{N_z(\theta)}$	**LRHGLRE (long run high gray level run emphasis)** measures the joint distribution of long run lengths with higher gray-level values.

Adapted from http://pyradiomics.readthedocs.io/en/latest/features.html – last accessed March 22, 2018. © Copyright 2016, pyradiomics community, http://github.com/radiomics/pyradiomics Revision eae15eff.

f. Neigbouring Gray Tone Difference Matrix (NGTDM) Features

A Neighbouring Gray Tone Difference Matrix quantifies the difference between a gray value and the average gray value of its neighbours within distance ***δ***. The sum of absolute differences for gray level ***i*** is stored in the matrix. Let $\mathbf{X_{gl}}$be a set of segmented voxels and $x_{gl}(j_x, j_y, j_z) \in X_{gl}$ be the gray level of a voxel at postion (j_x, j_y, j_z), then the average gray level of the neigbourhood is:

$$\overline{A}_i = \overline{A}(j_x, j_y, j_z)$$
$$= \frac{1}{W} \sum_{k_x=-\delta}^{\delta} \sum_{k_y=-\delta}^{\delta} \sum_{k_z=-\delta}^{\delta} x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z),$$

where $(k_x, k_y, k_z) \neq (0,0,0)$ and $x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z) \in X_{gl}$

Here, ***W*** is the number of voxels in the neighbourhood that are also in $\boldsymbol{X_{gl}}$.

Let:

$\boldsymbol{n_i}$ be the number of voxels in $\boldsymbol{X_{gl}}$ with gray level ***i***

Nv,p be the total number of voxels in $\boldsymbol{X_{gl}}$ and equal to $\sum n_i$ (i.e. the number of voxels with a valid region; at least 1 neighbor). $N_{v,p} \leq N_p$, where N_p is the total number of voxels in the ROI.

$\boldsymbol{p_i}$ be the gray level probability and equal to n_i/N_v

$s_i = \{\sum^{n_i} |i - \overline{A}_i|$ for $n_i \neq 0$; 0 for $n_i = 0$ be the sum of absolute differences for gray level i

$\boldsymbol{N_g}$ be the number of discreet gray levels

$\boldsymbol{N_{g,p}}$ be the number of gray levels where $\boldsymbol{p_i} \neq 0$

Appendix Table A7. Neigbouring Gray Tone Difference Matrix (NGTDM) for peri-atrial segmentation

Radiomic feature	Interpretation
$\boldsymbol{Coarseness} = \frac{1}{\sum_{i=1}^{N_g} p_i s_i}$	**Coarseness** is a measure of average difference between the center voxel and its neighbourhood and is an indication of the spatial rate of change. A higher value indicates a lower spatial change rate and

	a locally more uniform texture.
$$\boldsymbol{Contrast} = \left(\frac{1}{N_{g,p}(N_{g,p}-1)} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_i p_j (i - j)^2 \right) \left(\frac{1}{N_{v,p}} \sum_{i=1}^{N_g} s_i \right),$$ where $p_i \neq 0, p_j \neq 0$	**Contrast** is a measure of the spatial intensity change, but is also dependent on the overall gray level dynamic range. Contrast is high when both the dynamic range and the spatial change rate are high, i.e. an image with a large range of gray levels, with large changes between voxels and their neighbourhood.
$$\boldsymbol{Busyness} = \frac{\sum_{i=1}^{N_g} p_i s_i}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \lvert i p_i - j p_j \rvert},$$ where $p_i \neq 0, p_j \neq 0$	A measure of the change from a pixel to its neighbour. A high value for busyness indicates a 'busy' image, with rapid changes of intensity between pixels and its neighbourhood.
$$\boldsymbol{Complexity} = \frac{1}{N_{v,p}} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \lvert i - j \rvert \frac{p_i s_i + p_j s_j}{p_i + p_j},$$ where $p_i \neq 0, p_j \neq 0$	An image is considered complex when there are many primitive components in the image, i.e. the image is non-uniform and there are many rapid changes in gray level intensity.
$$\boldsymbol{Strength} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p_i + p_j)(i - j)^2}{\sum_{i=1}^{N_g} s_i},$$ where $p_i \neq 0, p_j \neq 0$	**Strength** is a measure of the primitives in an image. Its value is high when the primitives are easily defined and visible, i.e. an image with slow change in intensity but more large coarse differences in gray level intensities.
Adapted from http://pyradiomics.readthedocs.io/en/latest/features.html – last accessed March 22, 2018. © Copyright 2016, pyradiomics community, http://github.com/radiomics/pyradiomics Revision eae15eff.	

g. Gray Level Dependence Matrix (GLDM)

A Gray Level Dependence Matrix (GLDM) quantifies gray level dependencies in an image. A gray level dependency is defined as the number of connected voxels within distance ***δ*** that are dependent on the center voxel. A neighbouring voxel with gray level ***j*** is considered dependent on center voxel with gray level ***i*** if $\boldsymbol{|i-j| \leq \alpha}$. In a gray level dependence matrix ***P(i,j)*** the $\boldsymbol{(i,j)^{th}}$ element describes the number of times a voxel with gray level ***i*** with ***j*** dependent voxels in its neighbourhood appears in image.

$\boldsymbol{N_g}$ be the number of discreet intensity values in the image

N_d be the number of discreet dependency sizes in the image

N_z be the number of dependency zones in the image, which is equal to $\sum_{i=1}^{N_g} \quad \sum_{j=1}^{N_d} \mathrm{P}(i,j)$

P(i,j) be the dependence matrix

p(i,j) be the normalised dependence matrix, defined as $p(i,j) = \frac{\mathrm{P}(i,j)}{N_z}$

Appendix Table A7. Gray Level Dependence Matrix (GLDM) statistics for peri-atrial segmentation

Radiomic feature	Interpretation
$\boldsymbol{SDE} = \frac{\sum_{i=1}^{N_g} \quad \sum_{j=1}^{N_d} \frac{\mathrm{P}(i,j)}{i^2}}{N_z}$	**SDE (Small Dependence Emphasis)**: A measure of the distribution of small dependencies, with a greater value indicative of smaller dependence and less homogeneous textures.
$\boldsymbol{LDE} = \frac{\sum_{i=1}^{N_g} \quad \sum_{j=1}^{N_d} \mathrm{P}(i,j) j^2}{N_z}$	**LDE (Large Dependence Emphasis)**: A measure of the distribution of large dependencies, with a greater value indicative of larger dependence and more homogeneous textures.
$\boldsymbol{GLN} = \frac{\sum_{i=1}^{N_g} \quad (\sum_{j=1}^{N_d} \mathrm{P}(i,j))^2}{\sum_{i=1}^{N_g} \quad \sum_{j=1}^{N_d} \quad \mathrm{P}(i,j)}$	**GLN (Gray Level Non-Uniformity)**: Measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values.
$\boldsymbol{DN} = \frac{\sum_{j=1}^{N_d} \quad (\sum_{i=1}^{N_g} \mathrm{P}(i,j))^2}{N_z}$	**DN (Dependence Non-Uniformity)**: Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image.
$\boldsymbol{DNN} = \frac{\sum_{j=1}^{N_d} \quad (\sum_{i=1}^{N_g} \mathrm{P}(i,j))^2}{N_z^2}$	**DNN (Dependence Non-Uniformity Normalised)**: Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image. This is the normalised version of the DLN formula.
$\boldsymbol{GLV} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j)(i-\mu)^2$, where $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} ip(i,j)$	**GLV (Gray Level Variance)**: Measures the variance in gray level in the image.
$\boldsymbol{DV} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j)(j-\mu)^2$,	**DV (Dependence Variance)**: Measures the variance in dependence size in the image.

where $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} jp(i,j)$	
$\boldsymbol{DE} = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j)\log_2 (p(i,j) + \epsilon)$	**DE (Dependence Entropy)**: Measures the entropy in dependence size in the image.
$\boldsymbol{LGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathrm{P}(i,j)}{i^2}}{N_z}$	**LGLE (Low Gray Level Emphasis)**: Measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.
$\boldsymbol{HGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathrm{P}(i,j)i^2}{N_z}$	**HGLE (High Gray Level Emphasis)**: Measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.
$\boldsymbol{SDLGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathrm{P}(i,j)}{i^2 j^2}}{N_z}$	**SDLGLE (Small Dependence Low Gray Level Emphasis)**: Measures the joint distribution of small dependence with lower gray-level values.C
$\boldsymbol{SDHGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathrm{P}(i,j)i^2}{j^2}}{N_z}$	**SDHGLE (Small Dependence High Gray Level Emphasis)**: Measures the joint distribution of small dependence with higher gray-level values.
$\boldsymbol{LDLGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathrm{P}(i,j)j^2}{i^2}}{N_z}$	**LDLGLE (Large Dependence Low Gray Level Emphasis)**: Measures the joint distribution of large dependence with lower gray-level values.
$\boldsymbol{LDHGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathrm{P}(i,j)i^2 j^2}{N_z}$	**LDHGLE (Large Dependence High Gray Level Emphasis)**: Measures the joint distribution of large dependence with higher gray-level values.

Adapted from http://pyradiomics.readthedocs.io/en/latest/features.html – last accessed March 22, 2018. © Copyright 2016, pyradiomics community, http://github.com/radiomics/pyradiomics Revision eae15eff.

Appendix 2. ORFAN Study approvals

ORFAN Study Research Ethics Committee approvals

NHS
Health Research Authority
South Central - Oxford C Research Ethics Committee
Level 3, Block B
Whitefriars Building
Lewins Mead
Bristol
BS1 2NT

Telephone: 0117 342 1383

04 November 2015

Professor Charalambos Antoniades
Associate Professor of Cardiovascular Medicine
University of Oxford
Division of Cardiovascular Medicine, University of Oxford
Level 6 West Wing, John Radcliffe Hospital
Headley Way, Oxford
Oxford
OX3 9DU

Dear Professor Antoniades

Study title:	**The Oxford Risk Factors And Non-invasive imaging Study**
REC reference:	**15/SC/0545**
Protocol number:	**11578**
IRAS project ID:	**164361**

Thank you for your letter, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Vice-Chair.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact the REC Manager, Maeve Groot Bluemink, nrescommittee.southcentral-oxfordc@nhs.net.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non clinical trials this is not currently mandatory.

If a sponsor wishes to contest the need for registration they should contact Catherine Blewett (catherineblewett@nhs.net), the HRA does not, however, expect exceptions to be made. Guidance on where to register is provided within IRAS.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non-NHS sites

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Covering letter on headed paper [Cover letter]	V1.0	14 August 2015
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Sponsor Letter]	V1.0	14 August 2015
IRAS Checklist XML [Checklist_29102015]		29 October 2015
Letter from funder [Funding letters 1&2]	V1.0	05 August 2015
Letter from sponsor [Sponsor letter]	V1.0	14 August 2015
Letters of invitation to participant [Study Arm 1]	V1.0	05 August 2015
Letters of invitation to participant [Study Arm 2]	V1.0	05 August 2015
Letters of invitation to participant [Study Arm 3]	V1.0	05 August 2015
Non-validated questionnaire [Annual Questionnaire]	1.0	04 August 2015
Participant consent form [Consent 1]	v2.0	22 October 2015
Participant consent form [Consent 2]	v2.0	22 October 2015
Participant information sheet (PIS) [PIL]	v2.0	22 October 2015
REC Application Form [REC_Form_24082015]		24 August 2015
Referee's report or other scientific critique report [Peer Review]	V1.0	05 August 2015
Research protocol or project proposal [Protocol]	V2.0	22 October 2015
Summary CV for Chief Investigator (CI) [Prof Antoniades CV]	V1.0	05 August 2015

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at
http://www.hra.nhs.uk/hra-training/

15/SC/0545	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project.

Yours sincerely



Pp Susan Lousada
Vice Chair

Email:nrescommittee.southcentral-oxfordc@nhs.net

Enclosures: "After ethical review – guidance for researchers"

Copy to: *Ms Heather House, Oxford University Hospitals NHS Trust*

Henry West

From: oxfordc.rec@hra.nhs.uk <noreply@harp.org.uk>
Sent: 18 July 2022 12:59
To: antoniad@well.ox.ac.uk; ouhtma@ouh.nhs.uk; heather.house@ouh.nhs.uk; CTRG Sponsorship Correspondence
Cc: Maria Lyasheva
Subject: IRAS Project ID 164361. HRA and HCRW Approval for the Amendment

Dear Dr. Antoniades,

IRAS Project ID:	164361
Short Study Title:	The Oxford Risk Factors And Non-invasive imaging Study
Amendment No./Sponsor Ref:	Minor Amendment 3
Amendment Date:	07 July 2022
Amendment Type:	Non Substantial Non-CTIMP

I am pleased to confirm **HRA and HCRW Approval** for the above referenced amendment.

You should implement this amendment at NHS organisations in England and Wales, in line with the guidance in the amendment tool.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/.

Please contact amendments@hra.nhs.uk for any queries relating to the assessment of this amendment.

Kind regards

Miss Gemma Oakes
Approvals Specialist
Health Research Authority
Ground Floor | Skipton House | 80 London Road | London | SE1 6LH
E.amendments@hra.nhs.uk
W. www.hra.nhs.uk

Sign up to receive our newsletter HRA Latest.

ORFAN Study Health Research Authority Confidentiality Advisory Group approvals

Health Research Authority
Skipton House
80 London Road
London
SE1 6LH

Tel: 020 7104 8100
Email: cag@hra.nhs.uk

07 January 2021

Professor Charalambos Antoniades
Professor of Cardiovascular Medicine
Honorary Consultant Cardiologist
Oxford University Hospitals NHS Foundation Trust
Level 6, West Wing,
John Radcliffe Hospital,
Oxford,
OX3 9DU

Dear Professor Antoniades,

Application title:	**The Oxford Risk Factors And Non-invasive imaging Study: ORFAN**
CAG reference:	**20/CAG/0157**
IRAS project ID:	**164361**
REC reference:	**15/SC/0545**

Thank you for submitting a **research** application under Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002 ('section 251 support') to process confidential patient information without consent.

Supported applications allow the controller(s) of the relevant data sources, if they wish, to provide specified information to the applicant for the purposes of the relevant activity without being in breach of the common law duty of confidence. Support provides a lawful basis to allow the information to be processed by the relevant parties for the specified purposes without incurring a breach of the common law duty of confidence only. Applicants must ensure the activity remains fully compliant with all other relevant legislation.

The role of the Confidentiality Advisory Group (CAG) is to review applications submitted under these Regulations and to provide advice to the Health Research Authority on whether application activity should be supported, and if so, any relevant conditions. This application was considered at the precedent set CAG meeting held on 04 December 2020.

This outcome should be read in conjunction with the provisional support letter dated 16 December 2020

Health Research Authority decision

The Health Research Authority, having considered the advice from the Confidentiality Advisory Group as set out below, has determined the following:

The application, to allow the disclosure of confidential patient information from participating NHS Trusts to NHS Digital, Barts health NHS Trust (on behalf of the National Institute for Cardiovascular Outcomes Research (NICOR), Kings College London (on behalf of the Sentinel Stroke National Audit Programme (SSNAP), and local NHS Trust NIHR Biomedical Research Centres (BRC's) for the purposes of time limited linkage with clinical datasets as listed in the application, in order for anonymised linked datasets to be disclosed to the applicants at the University of Oxford, and to allow these third parties to hold the linkage file throughout the life of the ORFAN Study to facilitate annual follow up for 10 years, is fully supported, subject to compliance with the standard and specific conditions of support.

Please note that the legal basis to allow access to the specified confidential patient information without consent is now in effect.

Context

Purpose of application

This application from the University of Oxford set out the purpose of medical research which aims to develop and validate novel imaging biomarkers to predict future heart attacks and other cardiovascular complications. This is a multi-arm study, with arm 4 only relevant to this application. Following the COVID-19 pandemic, the ORFAN study will additionally ascertain COVID-19 status (suspected or confirmed). The influence of COVID-19 infection on the outcome measures of the study will be explored.

Obesity is a known risk factor for cardiovascular disease (CVD). Body fat accumulation is associated with increased metabolic risk and it is now recognised that not only body fat quantity but also adipose tissue (AT) quality determines increased CVD risk. CT scans provide a reliable, accurate and non-invasive way to assess body adiposity. Increased volume of AT as measured by CT is independently associated with increased CVD risk in large clinical cohorts; similarly, increased epicardial AT volume is an independent predictor of cardiovascular event in subjects undergoing coronary CTA. Developing imaging biomarkers for assessing body adiposity could potentially provide a useful screening tool for the primary or secondary prevention of CVD.

A substantial proportion of Covid-19 patients have suffered adverse cardiovascular outcomes. If applicants can identify patients at risk for rapid deterioration during acute Covid-19 infection by quantifying the background vascular inflammation and the tendency of micro-thromboses, this will enable deployment of appropriate therapeutic measures (e.g. anticoagulation) and guide allocation of NHS resources.

NHS Trust direct care teams will identify retrospective patients with eligible scans. They are assigned a unique study ID, and the minimum required identifiers extracted alongside this ID are sent to;

- NHS Digital, for the purposes of linkage with HES, ONS, and other datasets as detailed below, including the emergency care dataset, medicines dispensed in primary care, and covid-19 hospitalisations,
- Barts Health NHS Trust, on behalf of the National Institute for Cardiovascular Outcomes Research (NICOR), for the purposes of linkage with clinical outcomes, such as heart attack
- Kings College London, on behalf of the Sentinel Stroke National Audit Programme (SSNAP), for the purposes of linkage with clinical outcomes, such as stroke,

- Participating NIHR Biomedical Research Centres (BRC's) for the purposes of linking with baseline clinical data extracted from the Electronic Patient record (EPR) of associated participating trust.

The above parties will perform the data linkage, will modify date of birth and date of death and will delete identifiable data, with only the unique study ID remaining. The linked datasets will be sent to the ORFAN study team, at University of Oxford. The third parties that performed the data linkage will keep the linkage file for 10 years to update the relevant outcomes on an annual basis, to avoid individual NHS Trusts having to send identifiers to the third parties multiple times.

The scans are pseudonymised with the unique study ID by the direct care team and securely sent to University of Oxford. The direct care team will also screen electronic health records for health outcomes. Participating trusts also send health outcomes alongside the unique study ID to the ORFAN study team, at university of Oxford. The ORFAN team will link the datasets received from NHS Digital, NICOR, SSNAP and the NIHR BRCs, with the clinical data including the scans from the participating Trusts.

75,000 retrospective patients data will be collected over a 2 year period from when CAG support is provided. The data collected will span different time points as each participating Trust, as they commenced CT services at different times. However no patient with a scan prior to 2010 will be included.

A recommendation for class 1, 4 & 6 support was requested to cover the relevant unconsented activities as described in the application.

Confidential patient information requested

The following sets out a summary of the specified cohort, listed data sources and key identifiers. Where applicable, full datasets and data flows are provided in the application form and relevant supporting documentation as this letter represents only a summary of the full detail.

Cohort	75,000 retrospective patients in the UK (England and wales only for the purposes of CAG) who have undergone clinical CTA's or unenhanced CT chest, abdomen and pelvis scans at participating NHS Trusts. The timescale for inclusion in the ORFAN study Arm 4 is different for each collaborating NHS Trust, however no patient with a scan before 2010 will be included in the study. Inclusion criteria detailed in the application
Data sources	1. NHS Digital; (HES, ONS, Medicines dispensed in primary care, Emergency care dataset, Covid-19 hospitalisations – datastes as detailed below) a. Hospital Episode Statistics (HES) – Accident and Emergency dataset b. HES – Admitted patient care dataset c. HES - Outpatients d. HES – Critical care e. Emergency Care Data Set (ECDS) f. Medicines dispensed in Primary Care (NHSBSA data) g. Civil Registration (Deaths) - Secondary Care

	Cut h. COVID-19 Hospitalization in England Surveillance Syste i. HES:Civil Registration (Deaths) bridge HQIP is the data controller for the below datasets, however the legal entity where identifiers are sent are; 2. Barts health NHS Trust (on behalf of the National Institute for Cardiovascular Outcomes Research (NICOR) 3. Kings College London, on behalf of the Sentinel Stroke National Audit Programme (SSNAP) 4. NHS Trusts participating – PACS radiology systems and electronic patient records at; • Oxford University Hospitals • Royal United Hospitals Bath • Milton Keynes University Hospital • University Hospitals of Leicester • Barts Health • Royal Brompton and Harefield • Leeds Teaching Hospitals • Royal Papworth Hospital • Cambridge University Hospitals • Guy's and St Thomas' NHS Foundation Trust • New Cross Hospital • Royal Wolverhampton NHS Trust • Sandwell & West Birmingham Hospitals • Queen Elizabeth Hospital Birmingham • University Hospitals Birmingham NHS Foundation Trust • University Hospital of Manchester Foundation Trust 5. NIHR Biomedical Research Centres (BRCs) that operate at Trusts with collaborating clinical teams – data extracted from electronic patient record (EPR) of the associated local trust • Oxford • Royal Brompton and Harefield • Leicester • Leeds • Cambridge • Barts
Identifiers required for linkage purposes **Minimum required as confirmed with each data processor**	**NHS Digital** 1) NHS Number 2) Date of birth 3) Postcode (if NHS Number is not available) 4) Name (if NHS Number is not available)

	NICOR 1) NHS Number 2) Date of birth 3) Postcode **SSNAP** 1) NHS Number 2) Date of birth 3) Postcode (if NHS Number is not available) 4) Name (if NHS Number is not available) **NIHR BRCs** 1) NHS Number 2) Date of birth 3) Local hospital number (MRN) 4) Date of relevant CT scan
Identifiers required for analysis purposes	1. Ethnicity 2. Unique study ID 3. Covid-19 status 4. Gender 5. date of death in the following format MM/YYYY. 6. Date of birth in the following format MM/YYYY This dataset can be considered anonymous as the applicants will not be able to re-identify any patient.
Additional information	The third party that performed the data linkage will keep linkage files to enable yearly outcome updates for 10 years.

Confidentiality Advisory Group advice

This letter summarises the outstanding elements set out in the provisional support letter, and the applicant response. The applicant response was considered by the Confidentiality Advice Team.

1. **Please provide REC Favourable Opinion regarding substantial amendment 6.**

The applicant provided the REC Favourable Opinion to the CAG on 07 January 2021.

Confidentiality Advisory Group advice conclusion

The CAG agreed that the minimum criteria under the Regulations appeared to have been met, and therefore advised recommending support to the Health Research Authority, subject to compliance with the specific and standard conditions of support as set out below.

Specific conditions of support

1. Favourable opinion from REC. This application is relevant to substantial amendment 6 - **Favourable Opinion Confirmed 05 January 2021**

2. Continual achievement of 'Standards Met' in relation to the relevant DSPT submission (or any future security assurance changes) for the duration of support. Evidence to be

provided (through NHS Digital confirmation they have reviewed and confirmed the DSPT submission as standards met' for the duration of support, and at time of each annual review.
All organisations processing confidential patient information, including the Trusts where patients are identified will be required to have security assurances in place. However, as there are more than five organisations, the CAT team will not check each one individually; it is the responsibility of the applicant to ensure these are in place.

As the above conditions have been accepted, this letter provides confirmation of final support. I will arrange for the register of approved applications on the HRA website to be updated with this information.

Application maintenance

Annual review

Please note that this legal support is subject to submission of an annual review report, for the duration of support, to show that the minimal amount of patient information is being processed and support is still necessary, how you have met the conditions or report plans, any public benefits that have arisen and action towards meeting them. It is also your responsibility to submit this report every 12 months for the entire duration that confidential patient information is being processed without consent.

The next annual review should be provided no later than **07 January 2022** and preferably 4 weeks before this date. Reminders are not issued so please ensure this is provided annually to avoid jeopardising the status of the support. Submission of an annual review in line with this schedule remains necessary even where there has been a delay to the commencement of the supported activity, or a halt in data processing. Please ensure you review the HRA website to ensure you are completing the most up to date 'section 251' annual review form as these may change.

For an annual review to be valid, there must also be evidence that the relevant DSPT submission(s) for organisations processing confidential patient information without consent are in place and have been reviewed by NHS Digital. Please plan to contact NHS Digital in advance of the CAG annual review submission date to check they have reviewed the relevant DSPTs and have confirmed these are satisfactory.

Register of Approved Applications

All supported applications to process confidential patient information without consent are listed in the published 'Register of Approved Applications'. It is a statutory requirement for the Register to be published and it is available on the CAG section of the Health Research Authority website. It contains applicant contact details, a summary of the research and other pertinent points.

This Register is used by controllers to check whether support is in place.

Changes to the application

The application and relevant documents set out the scope of the support which is in place for the application activity and any relevant restrictions around this.

Any amendments which are made to the scope of this support, including but not limited to, purpose, data flows, data sources, items of confidential patient information and processors, require submission of a formal amendment to the application. Changes to processors will require evidence of satisfactory DSPT submission. The amendment form can be found in the Confidentiality Advisory Group pages on the Health Research Authority website.

Support for any submitted amendment would not come into effect until a positive outcome letter has been issued.

Changes to the controller

Amendments which involve a change to the named controller for the application activity require the submission of a new and signed CAG application form and supporting documentation to support the application amendment. This is necessary to ensure that the application held on file appropriately reflects the organisation taking responsibility for the manner and purpose of data processing within the application, and that the legal support in place is related to the correct legal entity.

Applicants are advised to make contact with the Confidentiality Advice Team to discuss a change in controllership for an existing application in sufficient time ahead of the transfer of project responsibility to discuss the submission process timings.

Further information and relevant forms to amend the support is available on the HRA website.

Reviewed documents

The documents reviewed at the meeting were:

Document	*Version*	*Date*
CAG application from (signed/authorised) [ORFAN CAG_Form_V5.17.2 03.12.2020]		03 December 2020
Covering letter on headed paper [ORFAN CAG Cover Letter 23.07.2020 - Copy]		23 July 2020
Other [20CAG0157 CAT advice form with responses 03.12.2020]		03 December 2020
Other [ORFAN Data Flow Diagram Arm 4 only V1.1 02.12.2020]	1.1	02 December 2020
Other [ORFAN NHS Digital Privacy Notice Arm 4 only V1.1 03.12.2020]	1.1	03 December 2020
Patient Information Materials [PPI ORFAN Study Arm 4 Survey V1.3 03.12.2020]	1.3	03 December 2020
Patient Information Materials [ORFAN Study – Patient Notification Poster V1.0 03.12.2020]	1.0	03 December 2020
Patient Information Materials [ORFAN Study Arm 4 – Webpage Patient Notification V1.0 03.12.2020]	1.0	03 December 2020
Research protocol or project proposal [ORFAN Protocol V8.0 Tracked Changes 23_11_2020]	8	23 November 2020
Write recommendation from Caldicott Guardian (or equivalent) of applicant's organisation [ORFAN Written recommendation from Caldicott Guardian 09.11.2020]		09 November 2020
164361 15 SC 0545 AM07_Favourable_opinion_of_a_substantial_amendment 05 01 2021		05 January 2021

Membership of the Committee

The members of the Confidentiality Advisory Group who were present at the consideration of this item are listed below.

There were no declarations of interest in relation to this item.

Please do not hesitate to contact me if you have any queries following this letter. I would be grateful if you could quote the above reference number in all future correspondence.

With the Group's best wishes for the success of this project.

Yours sincerely

Caroline Watchurst
Confidentiality Advisor
On behalf of Health Research Authority

Email: cag@hra.nhs.uk

Enclosures:	*List of members who considered application* *Standard conditions of approval*
Copy to:	oxfordc.rec@hra.nhs.uk

Confidentiality Advisory Group precedent-set meeting attendance
04 December 2020

Members present:

Name	
Ms Clare Sanderson	CAG alternative vice-chair
Professor Barry Evans	CAG member
Mr Tony Kane	CAG member

Also in attendance:

Name	*Position (or reason for attending)*
Ms Caroline Watchurst	HRA Confidentiality Advisor
Dr Paul Mills	HRA Confidentiality Advice Service Manager

NHS
Health Research Authority

Standard conditions of support

Support to process the specified confidential patient information without consent, given by the Health Research Authority, is subject to compliance with the following standard conditions of support.

The applicant and those processing the information under the terms of the support will ensure that:

1. The specified confidential patient information is only used for the purpose(s) set out in the application.

2. Confidentiality is preserved and there are no disclosures of information in aggregate or patient level form that may inferentially identify a person, nor will any attempt be made to identify individuals, households or organisations in the data.

3. Requirements of the Statistics and Registration Services Act 2007 are adhered to regarding publication when relevant, in addition to other national guidance.

4. All staff with access to confidential patient information have contractual obligations of confidentiality, enforceable through disciplinary procedures.

5. All staff with access to confidential patient information have received appropriate ongoing training to ensure they are aware of their responsibilities and are acting in compliance with the application detail.

6. Activities must be compliant with the General Data Protection Regulation and Data Protection Act 2018.

7. Audit of data processing by a designated agent is facilitated and supported.

8. The wishes of patients who have withheld or withdrawn their consent are respected.

9. Any significant changes (for example, people, purpose, data flows, data items, security arrangements) must be approved via formal amendment prior to changes coming into effect.

10. An annual review report is submitted to the CAG every 12 months from the date of the final support letter, for the duration of the support.

11. Any breaches of confidentiality around the supported flows of information should be reported to CAG within 10 working days of the incident, along with remedial actions taken/to be taken. This does not remove the need to follow national/legal requirements for reporting relevant security breaches.

Health Research Authority
2 Redman Place
Stratford
London
E20 1JQ

Tel: 020 7104 8100
Email: cag@hra.nhs.uk

10 August 2022

Professor Charalambos Antoniades
Professor of Cardiovascular Medicine
Honorary Consultant Cardiologist
Oxford University Hospitals NHS Foundation Trust
Level 6, West Wing,
John Radcliffe Hospital,
Oxford,
OX3 9DU

Dear Professor Antoniades,

Application title:	**The Oxford Risk Factors And Non-invasive imaging Study: ORFAN**
CAG reference:	**20/CAG/0157**
IRAS project ID:	**164361**
REC reference:	**15/SC/0545**

Thank you for your amendment request to the above research application, submitted for support under Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002 to process confidential patient information without consent. Supported applications enable the data controller to provide specified information to the applicant for the purposes of the relevant activity, without being in breach of the common law duty of confidentiality, although other relevant legislative provisions will still be applicable.

The role of the Confidentiality Advisory Group (CAG) is to review applications submitted under these Regulations and to provide advice to the Health Research Authority on whether an application should be supported, and if so, any relevant conditions.

Health Research Authority decision

The Health Research Authority, having considered the advice from the Confidentiality Advisory Group as set out below, has determined the following:

The amendment, to include The Liverpool Heart and Chest Hospital NHS Foundation Trust as an additional data processor, as a newly participating site, and to collect post mortem CT scans, is supported, subject to compliance with the standard conditions of support.

Amendment request

This application aims to develop and validate novel imaging biomarkers to predict future heart attacks and other cardiovascular complications. Support is currently in place to allow the disclosure of confidential patient information from participating NHS Trusts to NHS

Digital, Arden & Gem CSU (on behalf of the National Institute for Cardiovascular Outcomes Research (NICOR)), Kings College London (on behalf of the Sentinel Stroke National Audit Programme (SSNAP)), and local NHS Trust NIHR Biomedical Research Centres (BRC's) for the purposes of time limited linkage with clinical datasets, in order for anonymised linked datasets to be disclosed to the applicants at the University of Oxford. Support is currently in place to use confidential patient information regarding 75,000 retrospective patients in the UK (England and Wales only for the purposes of CAG), who have undergone clinical Computed tomography angiography (CTA) scans or unenhanced computed tomography (CT) chest, abdomen and pelvis scans at participating NHS Trusts. The timescale for inclusion in the study is different for each collaborating NHS Trust, however, no patient with a scan before 2010 will be included in the study.

This amendment seeks support to include The Liverpool Heart and Chest Hospital NHS Foundation Trust as a new participating site (new data processor for CAG). The amendment also seeks support to include post mortem CT scans if they occurred, to ensure applicants can confirm scientific findings with scans performed after death associated with heart disease. CT scans are increasingly used for post-mortem investigation of cause of death. This is not an additional data flow, as 's251' support is already in place for this data flow, and it is not increasingly disclosive, however this is an additional data item.

The applicant has also provided updated patient facing documents which are accepted as notifications to CAG.

Confidentiality Advice Team advice

The amendment requested was considered by the Confidentiality Advice Team, who raised no queries regarding this amendment.

Confidentiality Advice Team conclusion

In line with the considerations above, the CAT agreed that the minimum criteria under the Regulations appeared to have been met for this amendment, and therefore advised recommending support to the Health Research Authority.

Specific conditions of support

1. Confirmation provided from the IG Delivery Team at NHS Digital to the CAG that the relevant Data Security and Protection Toolkit (DSPT) submission(s) has achieved the 'Standards Met' threshold: **Confirmed: All organisations processing confidential patient information, including the Trusts where patients are identified will be required to have security assurances in place. However, as there are more than five organisations, the CAT team will not check each one individually; it is the responsibility of the applicant to ensure these are in place.**

2. Confirmation of a favourable opinion from a Research Ethics Committee.
   Confirmed non substantial 07 July 2022

Reviewed documents

Document	*Version*	*Date*
Amendment request form		07 July 2022
IRAS Project ID 164361. HRA and HCRW Approval for the Amendment		07 July 2022

Please do not hesitate to contact me if you have any queries following this letter. I would be grateful if you could quote the above reference number in all future correspondence.

Yours sincerely

Caroline Watchurst
Confidentiality Advisor

On behalf of the Health Research Authority

Email: cag@hra.nhs.uk

Enclosures: Standard conditions of Support

cc. oxfordc.rec@hra.nhs.uk

NHS
Health Research Authority

Standard conditions of support

Support to process confidential patient information without consent, given by the Health Research Authority, is subject to the following standard conditions of support.

The applicant and those processing the information will ensure that:

1. The specified confidential patient information is only used for the purpose(s) set out in the application.

2. Confidentiality is preserved and there are no disclosures of information in aggregate or patient level form that may inferentially identify a person, nor will any attempt be made to identify individuals, households or organisations in the data.

3. Requirements of the Statistics and Registration Services Act 2007 are adhered to regarding publication when relevant, in addition to other national guidance.

4. All staff with access to confidential patient information have contractual obligations of confidentiality, enforceable through disciplinary procedures.

5. All staff with access to confidential patient information have received appropriate ongoing training to ensure they are aware of their responsibilities.

6. Activities remain consistent with the General Data Protection Regulation and Data Protection Act 2018.

7. Audit of data processing by a designated agent is facilitated and supported.

8. The wishes of patients who have withheld or withdrawn their consent are respected.

9. Any significant changes (for example, people, purpose, data flows, data items, security arrangements) must be supported via formal amendment prior to changes coming into effect.

10. An annual review report is submitted to the CAG every 12 months from the date of the final support letter, for the duration of the support.

11. Any breaches of confidentiality around the supported flows of information should be reported to CAG within 10 working days of the incident, along with remedial actions taken / to be taken. This does not remove the need to follow national/legal requirements for reporting relevant security breaches.

Appendix 3. ORFAN Study patient and public involvement documentation

ORFAN PPI survey documentation as used

Survey of radiology patients regarding the use of patient data

The ORFAN Study is a University of Oxford research study investigating the use of CT scans for detecting early signs of heart attacks, strokes and other disease related to the heart. The study has three arms including a study arm that does not approached patients in person but collects deidentified patient information. **Such research is only possible with special approval from the Health Research Authority's Confidentiality Advisory Group.**

This survey is to understand the views of patients about the careful use of their deidentified information for research without asking for their permission.

Why was I asked to do this?

. We wish to hear the views of people like yourself and use your feedback to best run the ORFAN study. This survey will not collect any of your personal information and completing the study does not make you a study participant.

Do I have to complete this survey?

No, you do not have to complete this survey. Completion of this survey is entirely optional and does not influence any clinical care during your visit to the hospital.

I wish to know more about the ORFAN Study, who can I contact?

There is a study team member available where you have been asked to complete this survey, so please ask at reception to speak to them if you have questions. You may also email Clinical Researcher Maria Lyasheva at maria.lyasheva@cardiov.ox.ac.uk with any questions.

Please turn over the page

subject:	ORFAN Study Arm 4 Survey	ethics ref:	15/SC/0545	version/date:	V1.3 3rd December 2020
short title:	The ORFAN Study	CI:	Prof. Charalambos Antoniades	page:	Page 1 of 3

Information relevant to the survey – please read this before completing the survey questions:

The ORFAN Study is run by the University of Oxford and needs to collect medical information from a large numbers of patients to be able to answer the research questions it is asking. To have enough information for our study we need to include anonymous data collection techniques that allow the study team to access relevant information about patients, without any access to identifying information about the patients. This means the researchers will never see information such as a patient's name, NHS/hospital numbers, date of birth, phone number or address. The only people who will have access to this information are those who normally do as part of providing you with clinical care such as the doctors, nurses and allied health staff.

To match the patient with other relevant clinical information held about them, such as if that patient had suffered from a heart attack at another hospital. The clinical care team will be asked to send identifying information such as a patient's NHS number and date of birth to an approved and secure third-party organisation such as NHS Digital (www.digital.nhs.uk). This is only possible with special approval from the Health Research Authority's Confidentiality Advisory Group. The deidentified information would then be shared with the ORFAN Study Researchers in Oxford.

Survey Questions

For all questions, please use the scale of 0 to 10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement

Please indicate your response to questions 1, 2 and 3 with a mark at the appropriate level of agreement/disagreement

Question 1:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10

Please turn over the page

Question 2:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties organisations, such as NHS Digital, for the additional collection of relevant research information'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10

Question 3:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual patient or contact the patient in any way.'

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10

Question 4:

Please provide any comments on the type of research described in this survey, including any concerns you have:

Thank you for taking the time to be involved in this survey.

ORFAN PPI Survey Report as utilised in REC and CAG applications

ORFAN Arm 4 data flow PPI survey report

This document summarises the survey that was carried out to inform the setup and management of the ORFAN Study Arm 4 in regard to the proposed flow of patient information.

This document briefly outlines how the survey was conducted, what the aggregate results of the survey are, and provides some examples of survey responses.

The survey document and example responses are attached at the end of the summary.

1. Why this survey was conducted

The Study investigators of the ORFAN study intend to apply for Confidentiality Advisory Group support to include patients who have not provided direct consent within study arm 4. The researchers intend to link these patient's identifiable information with information already held by other third-party organisations such as NHS Digital.

This is a change in the protocol of the ORFAN Study and a change in how data would flow within the ORFAN Study.

For this purpose, a survey was thought appropriate to hear patients' views on the proposed data flow for the amended ORFAN Arm 4 protocol and data flow.

2. How this survey was conducted

The survey was prepared by ORFAN researcher Dr Henry West and approved by University of Oxford Radcliffe Department of Medicine clinical research coordinator Ms Hanan Lamlum.

The survey was held in an ad-hoc fashion on weekdays when clinical research staff were available in November and December 2020.

The patients approached to participate were a convenience sample of those referred to outpatient radiology services at both the John Radcliffe Hospital (OUH NHS Trust) and The Manor Hospital in Oxford. All patients asked to complete the survey were attending radiology for a CT scan, and were public patients. The patients were not necessarily receiving heart CT scans. The patients we all aged over 18-years-old.

Patients were approached in person and asked if they were interested in completing the survey about a proposed University of Oxford research study. Those who were not interested were thanked for their time and not approached in any further way.

3. Survey results

The survey asked 3 questions that each gave the respondent a scale of 0-10 to respond as to their agreement or disagreement with a certain statement about the proposed data flow within ORFAN.

Question 4 was a free comment box with the respondent asked to provide any comments on the type of research described in this survey, including any concerns that they have.

The instruction for questions 1, 2 and 3 was as follows:

"For all questions, please use the scale of 0 to 10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement

Please indicate your response to questions 1, 2 and 3 with a mark at the appropriate level of agreement/disagreement"

Results of Question 1:

Statement provided: *'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.*

Number of complete responses: 25

Mean response: 9.12

Range of responses: 7 to 10

Mode Response: 10

Results of Question 2:

Statement provided: *'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties organisations, such as NHS Digital, for the additional collection of relevant research information'.*

Number of complete responses: 25

Mean response: 8.88

Range of responses: 5 to 10

Mode Response: 10

Results of Question 3:

Statement provided: *'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual patient or contact the patient in any way.'*

Number of complete responses: 25

Mean response: 8.92

Range of responses: 5 to 10

Mode Response: 10

Question 4 results:

Of 25 respondents, 13 proved comments. All comments are included below in transcribed forma and some are attached in appendix B.:

The majority of comments are very positive, with the concerns raised being in relation to the security of the data and the data flow and the concern of patients being contacted in an unsolicited manner. One respondent also expressed a wish for access to all his data.

Examples responses to the comment box are included here, with a focus on recurring themes:

'*Everything is fine, felt very comfortable*' (respondent 24)

'*I think that volunteers should be given access to their own biometric data, i.e. blood results and images*' (respondent 25)

'*It must be secure for patient' sake*' (respondent 23)

'*Happy to help!*' (respondent 15)

'*It makes sense. Thanks for your work*' (respondent 13)

'*As long as I/patients are not contacted*' (respondent 11)

'*No concerns*' (Respondent 10)

''*Happy to help*' (Respondent 9)

'*As long as it is secure*' (respondent 7)

''*We should support research* initiatives' (Respondent 6)

'*No concerns*' (Respondent 4)

'*Thank you. It must be safe to the patients*' (Respondent 2)

4. Discussion of results

The results were very supportive of the proposed data flow. The average agreement with each question was very high, and one participant was in the 'neutral'.

These results, and the supportive comments are reassuring that the population referred to public radiology services are supportive of the proposed data flow. Of the 13 who commented, the concerns raised were mostly general data protection concerns of relevance to all research, These include unsolicited contact made by the researchers (not possible in Arm 4 due to the research team never collecting contact details), and the security of the data (security procedures are detailed at length in the ORFAN Study Protocol and supporting documents.).

5. Appendix

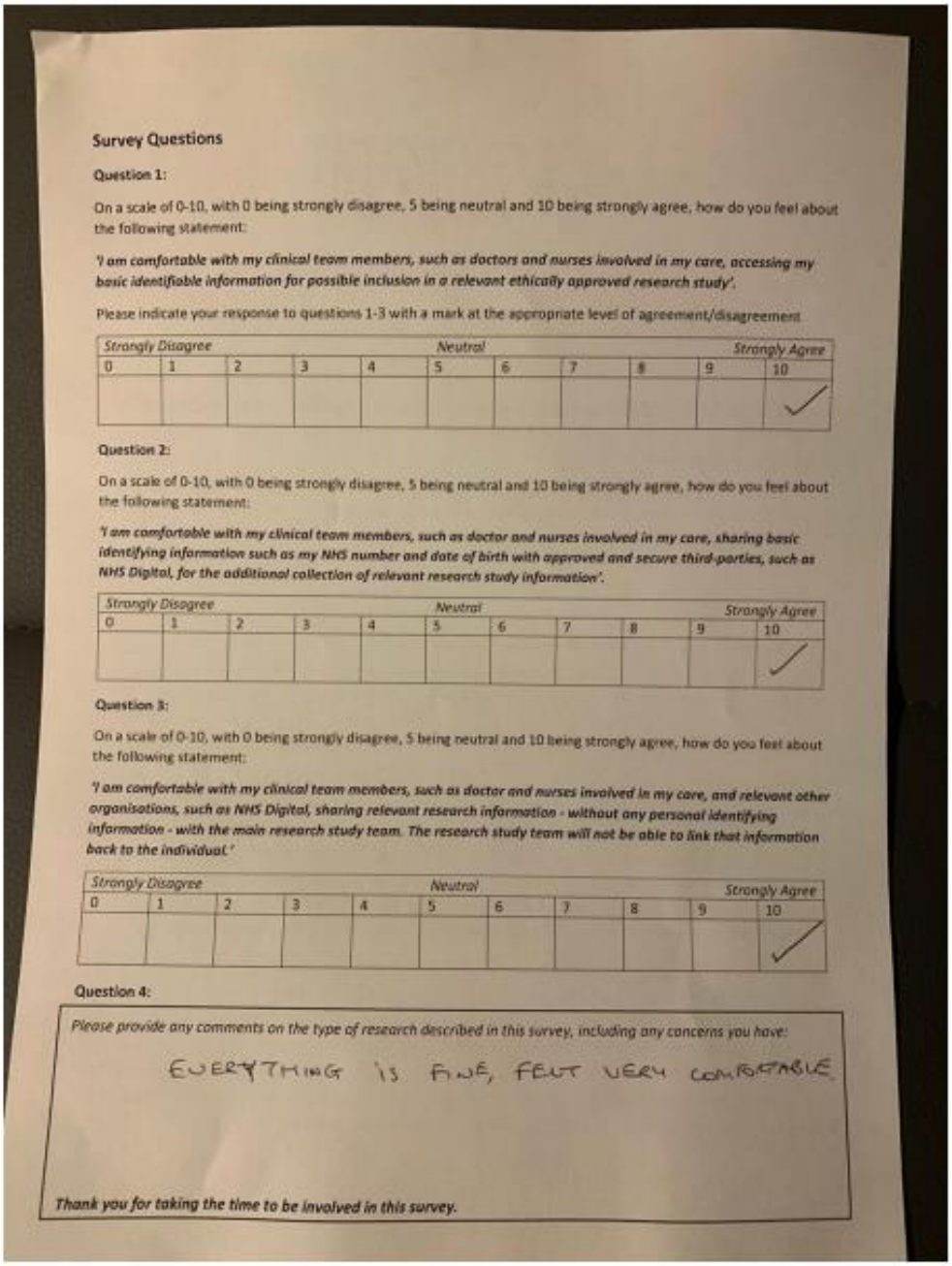
A. Survey form as used in the collection of responses – version 1.3
B. Scanned responses of respondents included in section 3 of this report

Appendix A - Survey form as used in the collection of responses – version 1.3

[REMOVED AS ALREADY INCLUDED AS APPENDIX 3.1]

Appendix B: Scanned responses of respondents included in section 3 of this report

Respondent 24:



Survey Questions

Question 1:

On a scale of 0-10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.

Please indicate your response to questions 1-3 with a mark at the appropriate level of agreement/disagreement

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
										✓

Question 2:

On a scale of 0-10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement:

'I am comfortable with my clinical team members, such as doctor and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties, such as NHS Digital, for the additional collection of relevant research study information'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
										✓

Question 3:

On a scale of 0-10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement:

'I am comfortable with my clinical team members, such as doctor and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual.'

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
										✓

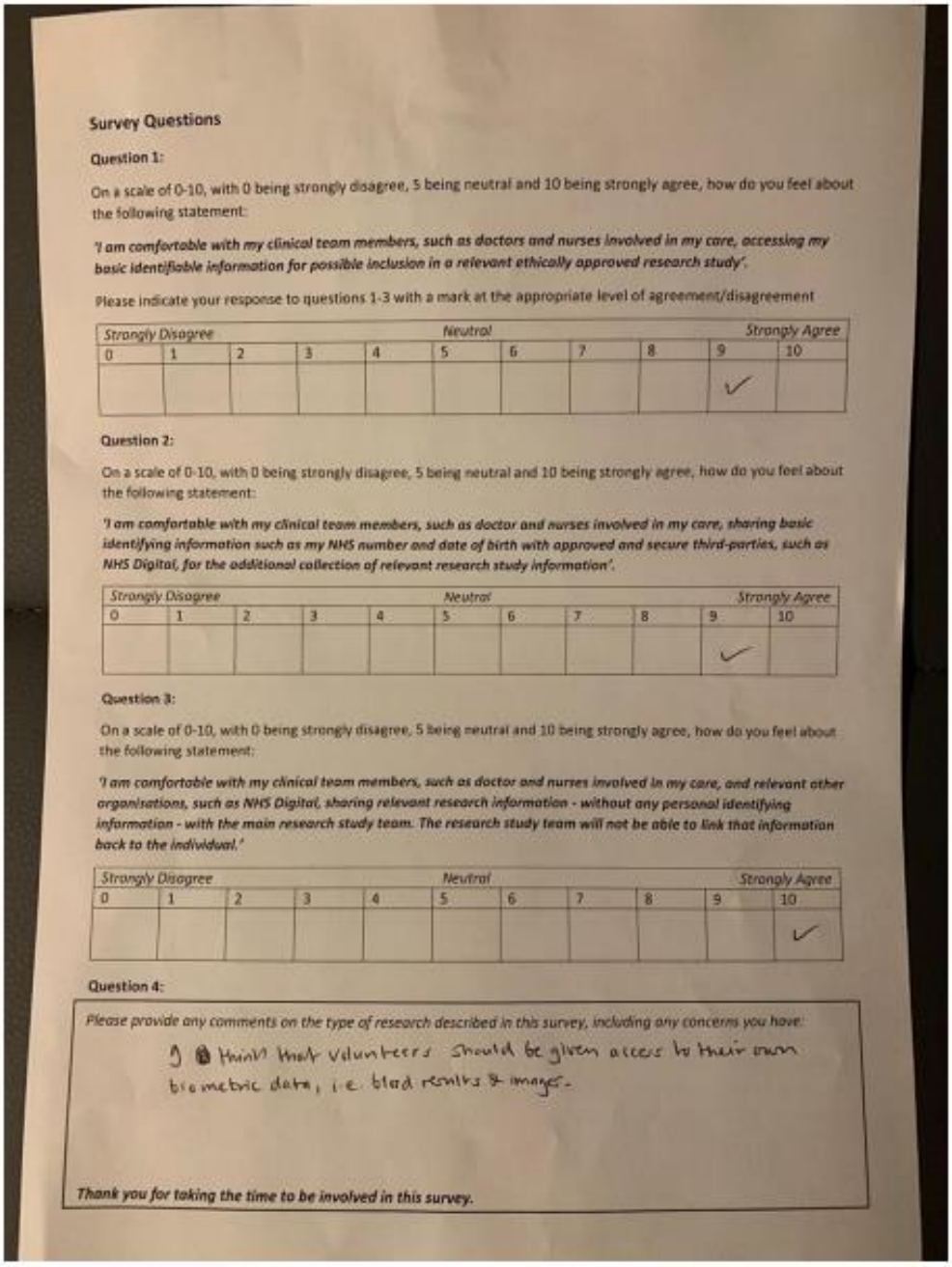
Question 4:

Please provide any comments on the type of research described in this survey, including any concerns you have:

EVERYTHING IS FINE, FELT VERY COMFORTABLE

Thank you for taking the time to be involved in this survey.

Respondent 25:



Survey Questions

Question 1:

On a scale of 0-10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.

Please indicate your response to questions 1-3 with a mark at the appropriate level of agreement/disagreement

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
									✓	

Question 2:

On a scale of 0-10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement:

'I am comfortable with my clinical team members, such as doctor and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties, such as NHS Digital, for the additional collection of relevant research study information'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
									✓	

Question 3:

On a scale of 0-10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement:

'I am comfortable with my clinical team members, such as doctor and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual.'

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
										✓

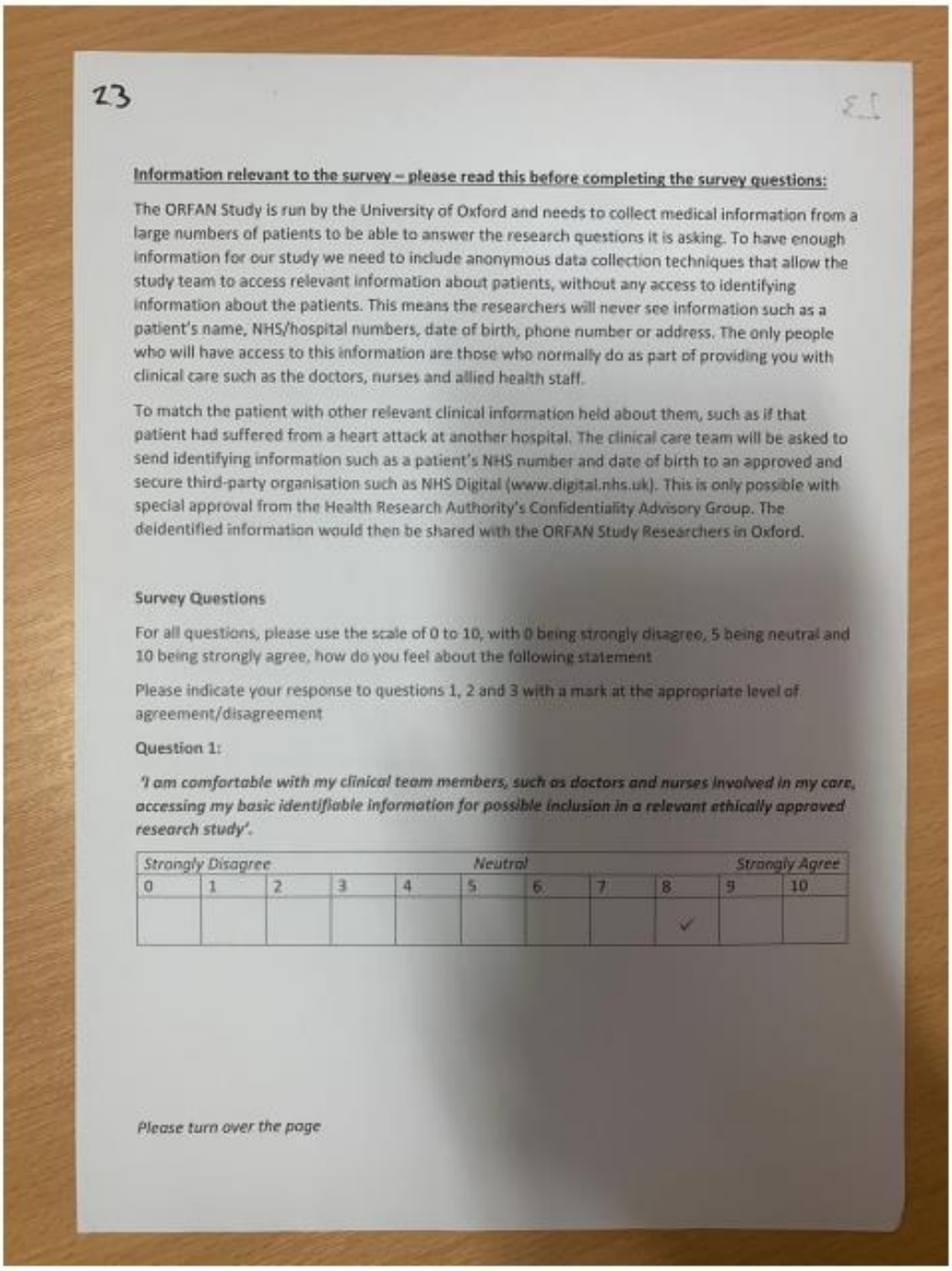
Question 4:

Please provide any comments on the type of research described in this survey, including any concerns you have:

I think that volunteers should be given access to their own biometric data, i.e. blood results & images.

Thank you for taking the time to be involved in this survey.

Respondent 23, page 1



23

Information relevant to the survey – please read this before completing the survey questions:

The ORFAN Study is run by the University of Oxford and needs to collect medical information from a large numbers of patients to be able to answer the research questions it is asking. To have enough information for our study we need to include anonymous data collection techniques that allow the study team to access relevant information about patients, without any access to identifying information about the patients. This means the researchers will never see information such as a patient's name, NHS/hospital numbers, date of birth, phone number or address. The only people who will have access to this information are those who normally do as part of providing you with clinical care such as the doctors, nurses and allied health staff.

To match the patient with other relevant clinical information held about them, such as if that patient had suffered from a heart attack at another hospital. The clinical care team will be asked to send identifying information such as a patient's NHS number and date of birth to an approved and secure third-party organisation such as NHS Digital (www.digital.nhs.uk). This is only possible with special approval from the Health Research Authority's Confidentiality Advisory Group. The deidentified information would then be shared with the ORFAN Study Researchers in Oxford.

Survey Questions

For all questions, please use the scale of 0 to 10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement

Please indicate your response to questions 1, 2 and 3 with a mark at the appropriate level of agreement/disagreement

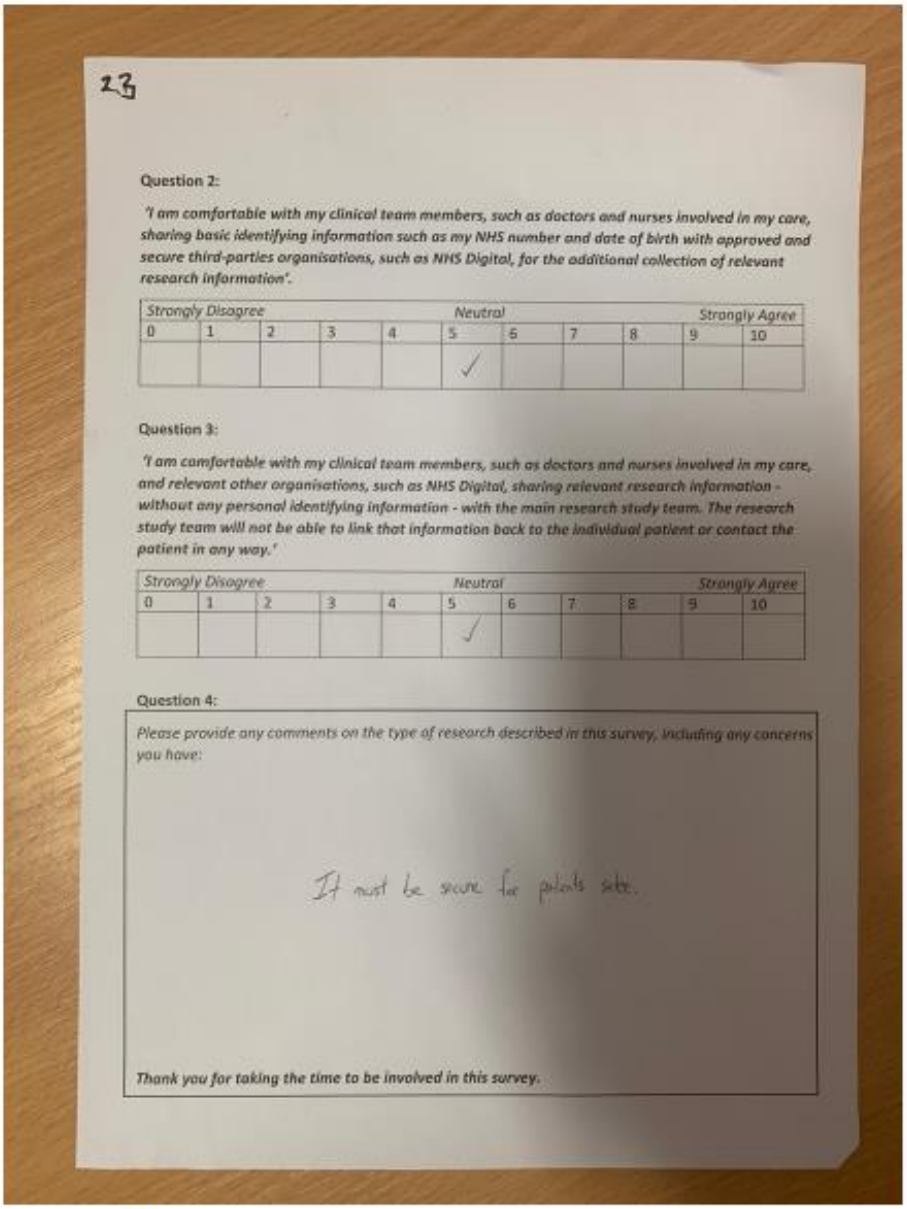
Question 1:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
								✓		

Please turn over the page

Respondent 23, page 2:



23

Question 2:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties organisations, such as NHS Digital, for the additional collection of relevant research information'.

Strongly Disagree					Neutral					Strongly Agree
0	1	2	3	4	5	6	7	8	9	10
					✓					

Question 3:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual patient or contact the patient in any way.'

Strongly Disagree					Neutral					Strongly Agree
0	1	2	3	4	5	6	7	8	9	10
					✓					

Question 4:

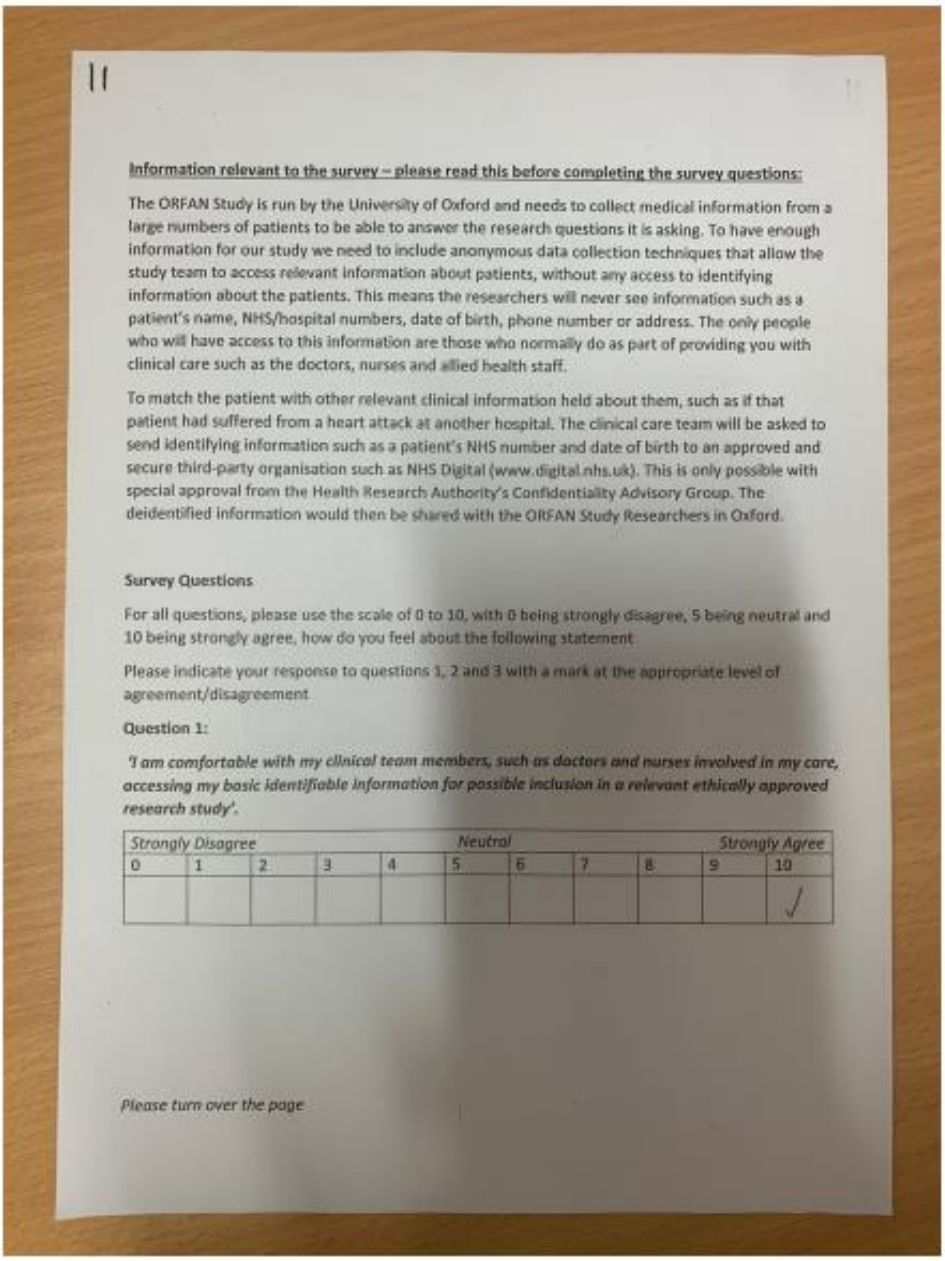
Please provide any comments on the type of research described in this survey, including any concerns you have:

It must be secure for patients sake.

Thank you for taking the time to be involved in this survey.

Respondent 11, page 1:

subject:	ORFAN Survey Report	ethics ref:	15/SC/0545	version/date:	V1.0.3 3rd December 2020
short title:	The ORFAN Study	CI:	Prof. Charalambos Antoniades	page:	Page 8 of 12



11

Information relevant to the survey – please read this before completing the survey questions:

The ORFAN Study is run by the University of Oxford and needs to collect medical information from a large numbers of patients to be able to answer the research questions it is asking. To have enough information for our study we need to include anonymous data collection techniques that allow the study team to access relevant information about patients, without any access to identifying information about the patients. This means the researchers will never see information such as a patient's name, NHS/hospital numbers, date of birth, phone number or address. The only people who will have access to this information are those who normally do as part of providing you with clinical care such as the doctors, nurses and allied health staff.

To match the patient with other relevant clinical information held about them, such as if that patient had suffered from a heart attack at another hospital. The clinical care team will be asked to send identifying information such as a patient's NHS number and date of birth to an approved and secure third-party organisation such as NHS Digital (www.digital.nhs.uk). This is only possible with special approval from the Health Research Authority's Confidentiality Advisory Group. The deidentified information would then be shared with the ORFAN Study Researchers in Oxford.

Survey Questions

For all questions, please use the scale of 0 to 10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement

Please indicate your response to questions 1, 2 and 3 with a mark at the appropriate level of agreement/disagreement

Question 1:

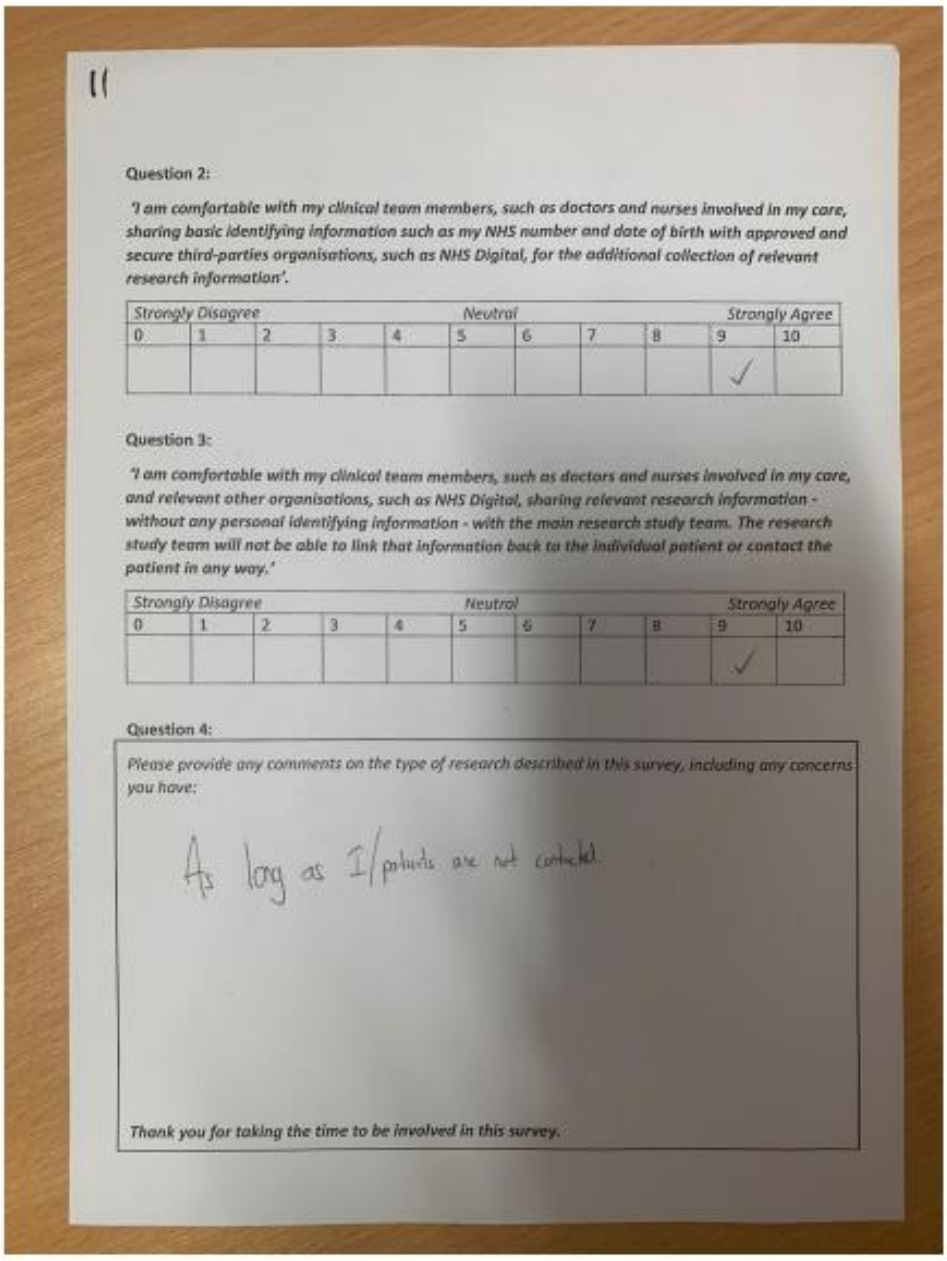
'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
										√

Please turn over the page

Respondent 11, page 2:

subject:	ORFAN Survey Report	ethics ref:	15/SC/0545	version/date:	V1.0.3 3rd December 2020
short title:	The ORFAN Study	CI:	Prof. Charalambos Antoniades	page:	Page 9 of 12



11

Question 2:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties organisations, such as NHS Digital, for the additional collection of relevant research information'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
									✓	

Question 3:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual patient or contact the patient in any way.'

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
									✓	

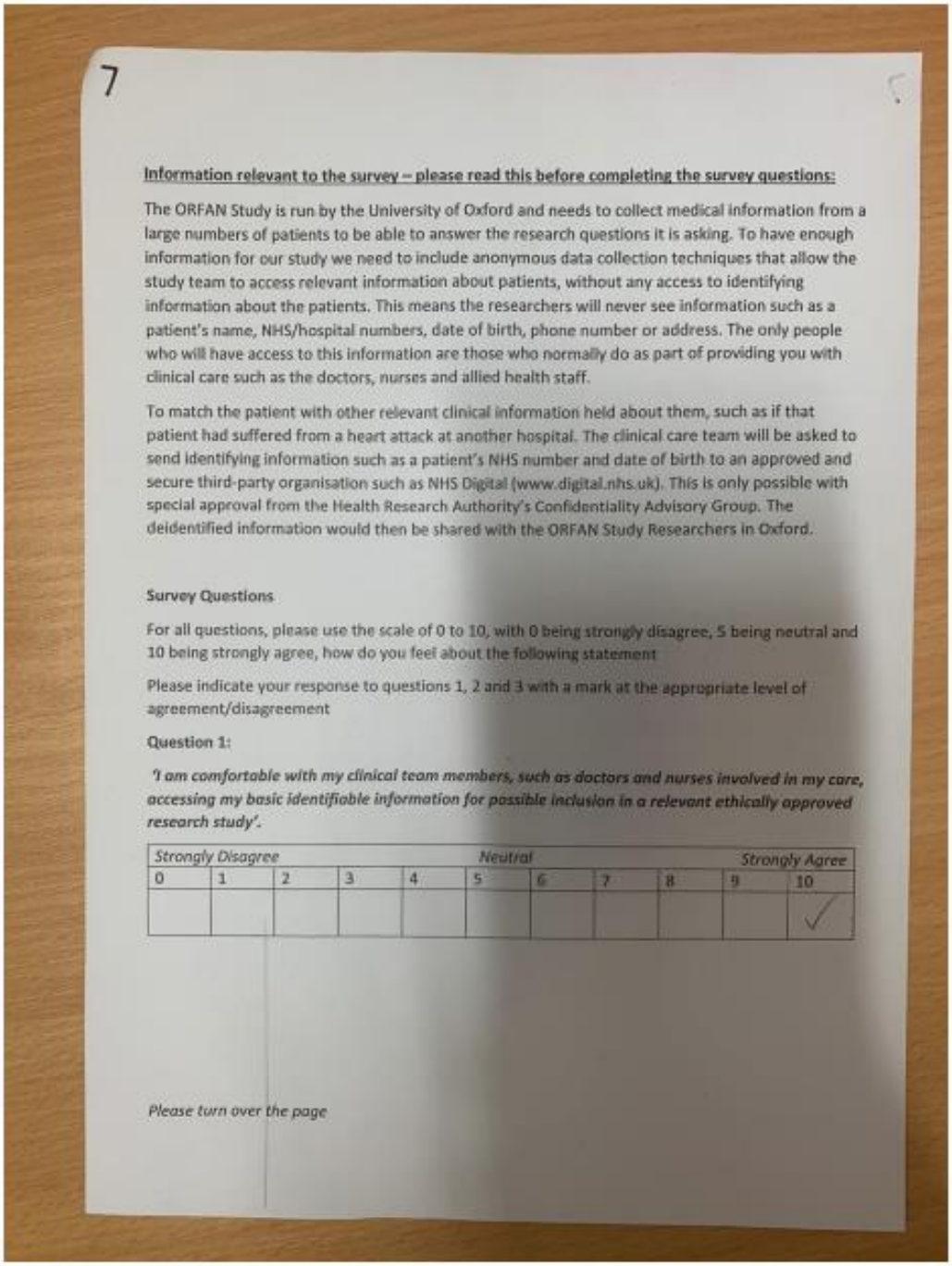
Question 4:

Please provide any comments on the type of research described in this survey, including any concerns you have:

As long as I/patients are not contacted.

Thank you for taking the time to be involved in this survey.

Respondent 7, page 1:



7

Information relevant to the survey – please read this before completing the survey questions:

The ORFAN Study is run by the University of Oxford and needs to collect medical information from a large numbers of patients to be able to answer the research questions it is asking. To have enough information for our study we need to include anonymous data collection techniques that allow the study team to access relevant information about patients, without any access to identifying information about the patients. This means the researchers will never see information such as a patient's name, NHS/hospital numbers, date of birth, phone number or address. The only people who will have access to this information are those who normally do as part of providing you with clinical care such as the doctors, nurses and allied health staff.

To match the patient with other relevant clinical information held about them, such as if that patient had suffered from a heart attack at another hospital. The clinical care team will be asked to send identifying information such as a patient's NHS number and date of birth to an approved and secure third-party organisation such as NHS Digital (www.digital.nhs.uk). This is only possible with special approval from the Health Research Authority's Confidentiality Advisory Group. The deidentified information would then be shared with the ORFAN Study Researchers in Oxford.

Survey Questions

For all questions, please use the scale of 0 to 10, with 0 being strongly disagree, 5 being neutral and 10 being strongly agree, how do you feel about the following statement

Please indicate your response to questions 1, 2 and 3 with a mark at the appropriate level of agreement/disagreement

Question 1:

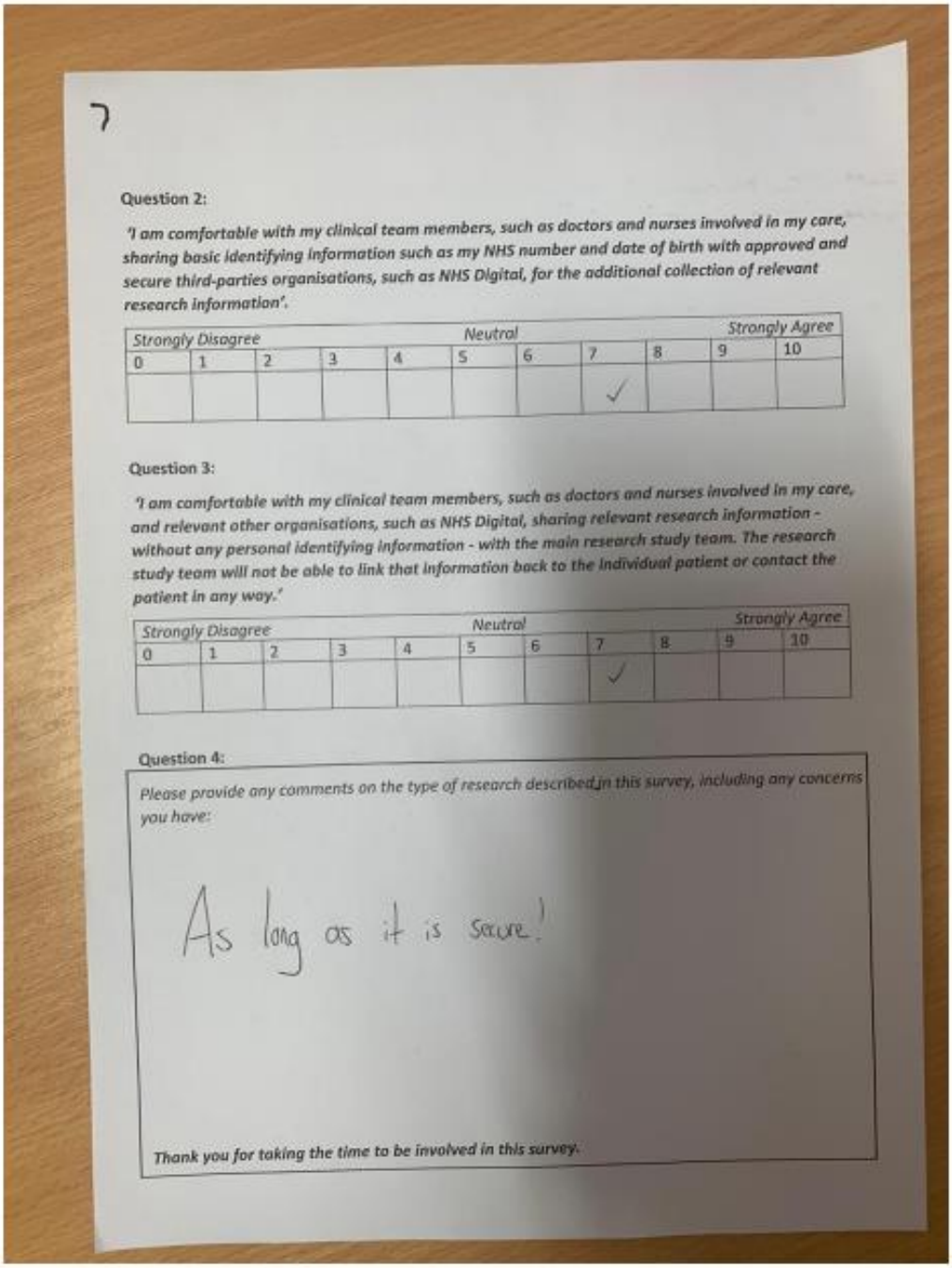
'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, accessing my basic identifiable information for possible inclusion in a relevant ethically approved research study'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
										✓

Please turn over the page

Respondent 7, page 2:

subject:	ORFAN Survey Report	ethics ref:	15/SC/0545	version/date:	V1.0.3 3rd December 2020
short title:	The ORFAN Study	CI:	Prof. Charalambos Antoniades	page:	Page 11 of 12



7

Question 2:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, sharing basic identifying information such as my NHS number and date of birth with approved and secure third-parties organisations, such as NHS Digital, for the additional collection of relevant research information'.

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
							✓			

Question 3:

'I am comfortable with my clinical team members, such as doctors and nurses involved in my care, and relevant other organisations, such as NHS Digital, sharing relevant research information - without any personal identifying information - with the main research study team. The research study team will not be able to link that information back to the individual patient or contact the patient in any way.'

Strongly Disagree					*Neutral*					*Strongly Agree*
0	1	2	3	4	5	6	7	8	9	10
							✓			

Question 4:

Please provide any comments on the type of research described in this survey, including any concerns you have:

As long as it is secure!

Thank you for taking the time to be involved in this survey.

Appendix 4. ORFAN Study protocol - version 10

DIVISION OF CARDIOVASCULAR MEDICINE
RADCLIFFE DEPARTMENT OF MEDICINE
LEVEL 6 WEST WING,
JOHN RADCLIFFE HOSPITAL, OXFORD OX3 9DU

Study Title: **T**he **O**xford **R**isk **F**actors And **N**on-invasive Imaging Study

Internal Reference Number / Short title The ORFAN Study

Ethics Ref: 15/SC/0545

Date and Version No: 30 May 2022 Version 10

Chief Investigator: Professor Charalambos Antoniades[1]

Investigators: Dr Alexios Antonopoulos[1]

Ms Sheena Thomas[1]

Ms Maria Lyasheva[1]

Dr Evangelos Oikonomou[1]

Dr Nikant Sabharwal[2]

Dr Cheerag Shirodaria[2]

Dr Susan Anthony[2]

Dr Andrew Kelion[2]

Prof Adrian Banning[2]

Dr Rafail Angelos Kotronias[1,2]

Dr Henry West[1]

Dr Cheng Xie[1,2]

Dr Rajesh Kumar Kharbanda[1,2]

Dr Attila Kardos[1,4]

Dr David Adlam[3]

Dr Amrita Bajaj[3]

Dr Intrajeet Das[3]

Dr Aparna Deshpande[3]

Dr Praveen Rao[3]

Dr Tarun Mittal[5]

Dr Saeed Mirsadraee[5]

Dr Ed Nicol[5]

Dr Jonathan Rodrigues[6]

Dr Benjamin Hudson[6]

Prof John Greenwood[7]

Prof Colin Berry[8,9]

Prof Stephan Achenbach[10]

Dr Mohamed Marwan[10]

Dr Milind Y Desai[11]

Dr Nicholas Screaton[12]

Dr Pál Maurovich-Horvat[13]

Prof Guo-Wei He[14]

Dr Wen-Hua Lin[14]

Dr Li-Juan Fan[14]

Prof Naohiko Takahashi[15]

Dr Hidekazu Kondo[15]

Dr Neng Dai[16]

Prof. Junbo Ge[16]

Prof Bon-Kwon Koo[17]

Dr Gianluca Pontone[18]

Dr Marco Guglielmo[18]

Prof Ron Blankstein[19,20]

Prof Theodora Benedek[21]

Dr Ronak Rajani[22]

Dr Mak Sze Mun[22]

Dr Giulia Benedetti[22]

Dr Rebecca Louise Preston[22]

Dr Elisa McAlindon[23]

Dr Shahzad Munir[23]

Dr Derek Leslie Connolly[24]

Dr William Bradlow[25,26]

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Dr Matthias Schmitt[27,28]

Prof David Ernest Newby[29]

Prof Steffen E Petersen[30,31]

Dr Francesca Pugliese[32]

Dr Timothy A Fairbairn[33]

Dr Nehal N Mehta[34]

Prof Stefan Neubauer[1]

Prof John Deanfield[35,36]

Prof Keith Channon[1]

1. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford
2. Oxford University Hospitals NHS Foundation Trust
3. University Hospitals of Leicester NHS Trust
4. Milton Keynes University Hospital NHS Foundation Trust
5. Royal Brompton and Harefield NHS Foundation Trust
6. Royal United Hospitals Bath NHS Foundation Trust
7. Leeds Teaching Hospitals NHS Foundation Trust
8. NHS Greater Glasgow and Clyde, NHS Scotland
9. Golden Jubilee National Hospital, NHS Scotland
10. Department of Cardiology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
11. Cleveland Clinic Heart and Vascular Institute, Cleveland, U.S.A
12. Royal Papworth Hospital NHS Trust, Cambridge
13. Department of Radiology, MTA-SE Cardiovascular Imaging Research Group, Budapest, Hungary
14. TEDA International Cardiovascular Hospital, Tianjin, China
15. Oita University, Japan
16. Fudan University, China
17. Seoul National University, Seoul, South Korea
18. Centro Cardiologico Monzino IRCCS, University of Milan, Italy
19. Harvard Medical School, Boston, USA
20. Brigham and Women's Hospital, Boston, USA
21. University of Medicine and Pharmacy of Tirgu Mures, Romania
22. Guy's and St Thomas' NHS Foundation Trust
23. Heart and Lung Centre, New Cross Hospital, Wolverhampton
24. Sandwell & West Birmingham Hospitals NHS Trust
25. University Hospitals Birmingham (UHB) NHS Trust
26. University of Birmingham
27. University Hospital of Manchester Foundation Trust, Manchester
28. University of Manchester, Manchester
29. University of Edinburgh, Royal Infirmary, Edinburgh
30. The William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London
31. Barts Heart Centre, Barts Health NHS Trust
32. NIHR Barts Cardiovascular Biomedical Research Centre, the William Harvey Research Institute, Queen Mary University of London & Barts Health NHS Trust
33. Liverpool Heart and Chest Hospital NHS Foundation Trust
34. National Institutes of Health, National Heart, Lung, and Blood Institute
35. University College London
36. National Institute of Cardiovascular Outcomes Research (NICOR)

Sponsor: University of Oxford

Funder: Oxford Biomedical Research Centre, British Heart Foundation, Innovate UK, National Consortium of Intelligent Medical Imaging and the

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

European Commission, National Institute for Health Research (NIHR), and NHSX

Chief Investigator Signature:

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Disclosures

Professor Antoniades is the founder, Chief Scientific Officer and a shareholder for Caristo Diagnostics, a spin out company from the University of Oxford. Professor Neubauer is the Founder and Chief Translational Officer for Caristo Diagnostics. Professor Channon is the Founder and Chief Medical Officer for Caristo Diagnostics. Caristo Diagnostics may be used for anonymisation of images as well as image analysis.

Confidentiality Statement

This document contains confidential information that must not be disclosed to anyone other than the Sponsor, the Investigator Team, host organisation, and members of the Research Ethics Committee, unless authorised to do so.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

TABLE OF CONTENTS

1. SYNOPSIS 8
2. ABBREVIATIONS 9
3. BACKGROUND AND RATIONALE 11
3.1. Adipose Tissue and Cardiovascular Disease 11
3.2. Assessing Body Adiposity by Imaging 11
3.3. Unresolved Issues 11
3.4. Coronavirus disease 2019 (COVID-19) 12
3.5. Pilot Data 13
3.6. Study Hypothesis 16
4. OBJECTIVES AND OUTCOME MEASURES 17
5. STUDY DESIGN 18
5.1. Summary of Study Design 18
5.2. COVID-19 research 19
5.3. Clinical Protocol 21
5.4. Experimental Design 26
6. PARTICIPANT IDENTIFICATION 27
6.1. Study Participants 27
6.1.1. Inclusion Criteria for study Arms 1, 2 and 3 27
6.1.2. Inclusion Criteria for study Arm 4 27
6.1.3. Exclusion Criteria for Study Arms 1, 2 and 3 27
6.1.4. Exclusion Criteria for Study Arm 4 27
7. STUDY PROCEDURES 28
7.1. Recruitment 28
7.2. Screening and Eligibility Assessment 28
7.3. Baseline Assessments 28
7.4. Subsequent Visits and Long-Term Follow Up 29
7.5. Sample/Image Handling 29
7.6. Discontinuation/Withdrawal of Participants from Study 30
7.7. Definition of End of Study 30
7.8. Definition of Serious Adverse Events 30
7.9. Reporting Procedures for Serious Adverse Events 30
8. STATISTICS AND ANALYSIS 31

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

8.1. Description of Statistical Methods 31
8.2. The Level of Statistical Significance 32
8.3. Procedure for Accounting for Missing, Unused and Spurious Data 32
8.4. Power Calculations 32
8.5 Interim Analysis 34
9. DATA MANAGEMENT 35
9.1. Access to Data 35
9.2. Data Recording and Record Keeping 35
9.3 Image Data Anonymisation 38
10. QUALITY ASSURANCE PROCEDURES 39
11. ETHICAL AND REGULATORY CONSIDERATIONS 40
11.1. Declaration of Helsinki 40
11.2. ICH Guidelines for Good Clinical Practice 40
11.3. Section 251 Support 40
11.4. Approvals 40
11.5. Reporting 40
11.6. Participant Confidentiality 41
11.7. Approval for Sending Samples to External Specialised Centres for Processing 41
11.8. Expenses and Benefits 41
12. FINANCE AND INSURANCE 42
12.1. Funding 42
12.2. Insurance 42
13. PUBLICATION POLICY 43
14. REFERENCES 44
15. APPENDIX A: TELEPHONE QUESTIONNAIRE 46
AMENDMENT HISTORY 48

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

1. SYNOPSIS

Study Title	The Oxford Risk Factors And Non-invasive imaging Study	
Internal ref. no. / short title	The ORFAN Study	
Study Design	Multi- centre, prospective cohort, observational study	
Study Participants	High risk individuals with and without coronary atherosclerosis	
Planned Sample Size	265500; Arm 1: 500, Arm 2: 10000 Arm 3: 5000, Arm 4: up to 250000 (Nationally ~200000, Internationally ~50000)	
Planned Study Period	2015-2030	
	Objectives	**Outcome Measures**
Primary	To investigate whether biomarkers of metabolic risk can predict atherosclerosis disease progression or major adverse cardiovascular events	1. Measurements of plasma/imaging markers of cardiometabolic risk 2. Atherosclerosis progression by Computerised Tomography 3. Major adverse cardiovascular events over 10 years.
Secondary	Secondary objective 1: To identify novel biomarkers able to predict cardiovascular disease pathogenesis and extent of pre-existing vascular disease	1. Measurement of coronary plaque burden by Computerised Tomography. 2. Measurement of coronary calcification by Computerised Tomography 3. Measurement of aortic dimensions by Computerised Tomography 4. Measurement of changes in cardiac structure by Computerised Tomography 5. Measurement of novel indices of adipose tissue expansion and distribution, by using Computerised Tomography and/or ultrasound
	Secondary objective 2: Develop and validate imaging biomarkers to predict clinical outcomes in patients with COVID-19 infection.	1. Measurement of modification to cardiovascular disease pathogenesis and modulation of pre-existing vascular disease that is attributable to infection with coronavirus disease 2019 (COVID-19).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

2. ABBREVIATIONS

AT	Adipose Tissue
CABG	Coronary Artery Bypass Graft
CAG	Confidentiality Advisory Group
CVD	Cardiovascular Disease
CI	Chief Investigator
CO	Cardiology Outpatients
CRF	Case Report Form
CT	Computed Tomography
CTA	Computed Tomography Angiography
CCTA	Coronary Computed Tomography Angiography
CTRG	Clinical Trials & Research Governance, University of Oxford
eCRF	Electronic Case Report Form
GCP	Good Clinical Practice
GP	General Practitioner
HIPAA	Health Insurance Portability and Accountability Act
ICF	Informed Consent Form
ICH	International Conference of Harmonisation
JRH	John Radcliffe Hospital
NCIMI	National Consortium for Intelligent Medical Imaging
NHS	National Health Service
NICOR	National Institute Of Cardiovascular Outcomes Research
NIHR	National Institute for Health Research
NRES	National Research Ethics Service
OCB	Oxford CABG Bioresource
ONS	Office for National Statistics
OUH	Oxford University Hospitals
PACS	Picture Archiving and Communication System
PI	Principal Investigator
PIL	Participant Information Leaflet
PROF	Professor
R&D	NHS Trust Research and Development Department
REC	Research Ethics Committee
SFRP5	Secreted Frizzled-Related Protein-5

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

SSNAP	Sentinel Stroke National Audit Programme
UK	United Kingdom

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

3. BACKGROUND AND RATIONALE

3.1. Adipose Tissue and Cardiovascular Disease

Obesity is a known risk factor for cardiovascular disease (CVD) development and a global burden for national health systems. Body fat accumulation is associated with increased metabolic risk and increased body adiposity has been independently associated with adverse clinical outcomes. It is now becoming increasingly recognised that not only body fat quantity but also adipose tissue (AT) quality determines increased CVD risk (1,2).

Over recent years it has been acknowledged that AT is a key player in vascular disease development, via the release of active adipokines (cell secreting proteins). Adipokines released in circulation via AT depots such as visceral or subcutaneous AT can in an endocrine manner affect vascular biology. Likewise, perivascular AT can affect vascular biology by releasing adipokines that act on vascular wall in a paracrine manner (2-4).

3.2. Assessing Body Adiposity by Imaging

Anthropometric measurements such as body mass index or waist-to-hip ratio are widely used for assessing body adiposity, however these crude indices have inherent limitations, providing little information about body adiposity.

Ultrasound can be used to assess body adiposity, by measurement of the thickness of subcutaneous (5) or epicardial (6) fat pad. Given the non-invasive nature of ultrasound examinations and its widespread availability, ultrasound has been increasingly used for assessing body adiposity.

More recently, the widespread use of computed tomography (CT) in clinical practice has offered an additional way to assess body adiposity. AT can be reliably detected on CT scan images based on its relative radiodensity compared to water (7-9). By applying specific thresholds in the Hounsfield Units of CT scans, body fat can be detected. Thus, CT provides a reliable, accurate and non-invasive way to assess body adiposity. Increased volume of visceral or subcutaneous AT as measured by computed tomography is independently associated with increased CVD risk in large clinical cohorts (10,11); similarly, increased epicardial AT volume is an independent predictor of cardiovascular event in subjects undergoing coronary CT angiography (CCTA), over and above coronary calcification (8,9,12).

In this context, developing imaging biomarkers for assessing body adiposity seems to be promising as it can potentially provide a useful screening tool for the primary or secondary prevention of CVD.

3.3. Unresolved Issues

Even though various circulating levels of adipokines have been associated with CVD risk, the use of a single adipokine as a plasma biomarker that could be used in clinical practice for predicting CVD risk has proved particularly challenging. Therefore, the identification of a plasma biomarker that could relate metabolic status to CVD risk is still much needed.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Evidence suggests that recently discovered adipokines, with a still unclarified biological role, e.g. the secreted frizzled-related protein-5 (*SFRP5*) (13), may be important players in obesity-related cardiovascular disease. Measurement of circulating levels of *SFRP5* or other adipokines in a clinical cohort with prospective follow up, would help the identification of plasma biomarkers that predict cardiovascular events. Moreover, retrospective studies of selected subpopulations with known clinical outcomes would allow for the design of nested case-control studies to explore the association of newly discovered adipokines with cardiovascular risk.

Observed associations in clinical studies are often the result of confounding factors and do not necessarily imply causality. To address the issue of causality, we will search for novel genetic indices (e.g. single nucleotide polymorphisms related to the expression of *SFRP5* and others). This will help the building of genetic models that can be used in Mendelian randomisation studies which are a powerful clinical tool to assess the causal relationship of identified biomarkers (e.g. *SFRP5* and others) with cardiovascular disease risk.

In addition, whilst visceral or subcutaneous fat depot volumes quantified by CT have been well-validated as biomarkers of metabolic risk, the role of perivascular AT characteristics as an imaging biomarker has been largely overlooked. Given its close anatomical proximity with the vascular wall, perivascular AT could be important for coronary artery disease pathogenesis. We have recently shown that perivascular AT affects vascular biology by its secretory products (4). Perivascular AT could also be associated with development/progression of aortic aneurysms and peripheral arterial disease, even though this has never been tested to date.

CT imaging has been widely used to assess the volume of the various body fat depots as an index of body adiposity, however little is known about the value of assessing AT quality by CT imaging. Qualitative assessment of adipose tissue by CT imaging could be equally important to quantitative AT measurements for predicting CVD risk. Qualitative assessment of fat by CT has been recently suggested to provide information about AT biology.

A recent clinical study demonstrated that CT density of visceral and subcutaneous AT, as assessed by the average Hounsfield Units of AT, is independently associated with increased CVD risk in a community-based cohort of elderly subjects (14). Nevertheless, these data come only from a single study, and cannot be safely extrapolated to other cohorts of patients, such as to those with coronary artery disease, aortic aneurysms or peripheral arterial disease. Moreover, it is unknown whether qualitative assessment of perivascular AT could provide prognostic information over and above abdominal or subcutaneous adiposity measurements. Testing such hypotheses in large clinical cohorts is vital as it could help develop new imaging biomarkers for vascular disease prediction and help in risk stratification purposes.

3.4. Coronavirus disease 2019 (COVID-19)

In December 2019 the first cases of coronavirus disease 2019 (COVID-19) occurred through zoonotic transmission in China. The causative virus, namely severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is capable of human-to-human transmission and spread rapidly to other parts of China and subsequently to nearly every country in the globe, causing a pandemic as declared by the World Health Organization (15). By the end of September, SARS-CoV-2 has infected more than 35 million people worldwide and killed more than 1 million people (16).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

A substantial proportion of COVID-19 patients suffer adverse cardiovascular outcomes (17-19). Global data indicates the at-risk population have cardiovascular risk factors (e.g. age, coronary artery disease, obesity, hypertension, diabetes). What is common between these risk factors is a degree of vascular injury driven by background vascular inflammation and most likely a degree of pre-existing atherosclerosis.

COVID-19 infection triggers a "cytokine storm" which is responsible for the multi-organ damage that drives mortality. Evidence suggests that inflammation-induced vascular injury leads to multiple micro-thromboses, which propagate multi-organ failure. It is unclear whether the baseline vascular inflammatory burden or indeed pre-existing subclinical atherosclerosis, make the cardiovascular system vulnerable to this cytokine storm. If we can identify patients at risk for rapid deterioration during acute COVID-19 infection by quantifying the background vascular inflammation and the tendency of microthrombosis,this will enable deployment of appropriate therapeutic measures (e.g. anticoagulation) and guide allocation of NHS resources. We are developing and testing biomarkers that can non-invasively track inflammation in the vascular wall by analysing the radiotranscriptomic changes of perivascular fat from clinical CT angiograms (20,21) (Figure A4).

In previous viral pandemics (e.g. influenza) (22), the acute inflammatory assault of the vascular wall resets residual inflammatory risk, leading to a surge of heart attacks and strokes for years after the infection. A technology that quantifies the degree of vascular assault in COVID-19 can transform the patients' post-discharge management. We are seeking to test a radiotranscriptomic approach that covers this unmet need by analyzing clinical CT scans obtained during the acute phase of the infection. Such an approach could in future guide deployment of preventive strategies for the management of COVID-19 patients post-discharge (see pilot data below).

The influence of COVID-19 infection on the outcome measures of the study will be explored. The exploration of the relationship between outcome measures and COVID-19 will be strengthened through the identification of suitable ORFAN controls for a nested case control study.

3.5. Pilot Data

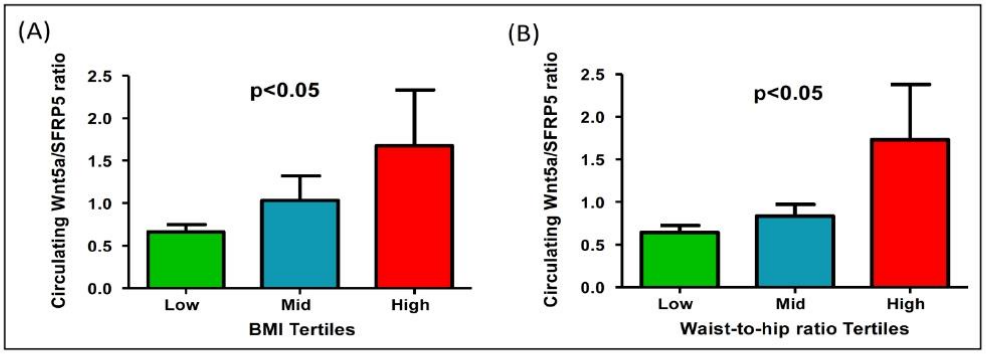


Fig. A2. In pilot data with 80 patients recruited under another study (AdipoRedOx, ref: REC C, 11/SC/0140) there was a significant correlation between wnt5a/*SFRP5* ratio and body mass index (BMI, A) and waist to hip ratio (B). P values from one-way ANOVA

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

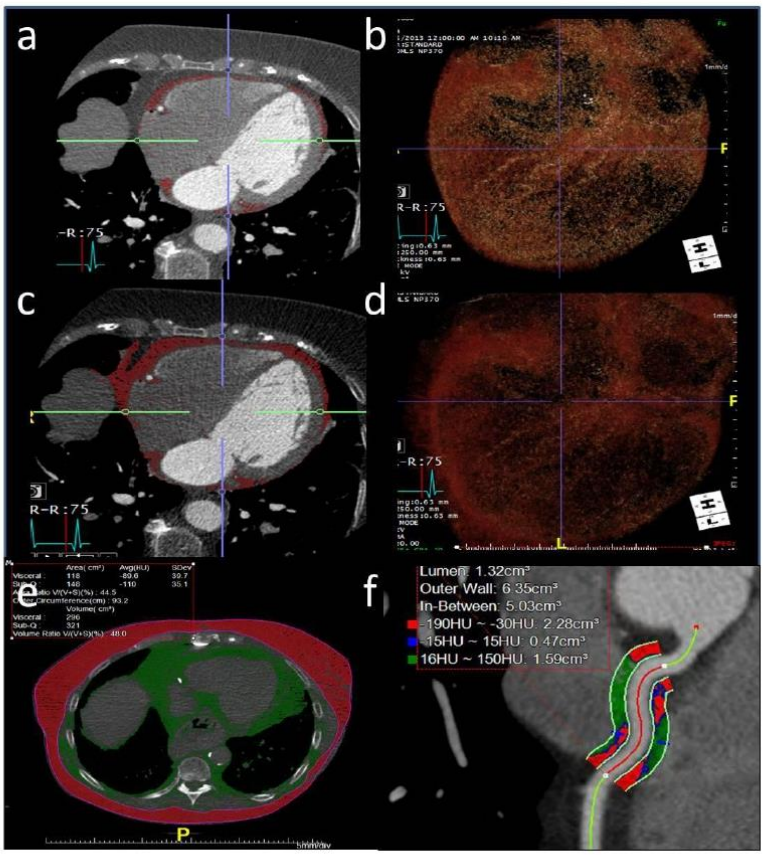


Fig. A3: Example of calculation of adipose tissue volumes using computerised tomography. a) calculation of epicardial adipose tissue and b) volume-rendered reconstruction of the tracked volume, c) measurement of thoracic adipose tissue volume and d) volume-rendered reconstruction of the tracked volume. e) example of calculation of subcutaneous thoracic adipose tissue volume from axial slices, f) a novel tool to quantify the characteristics of perivascular adipose tissue around the right coronary artery. Adipose tissue was defined as voxels between -190 to -30 Hounsfield Units.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

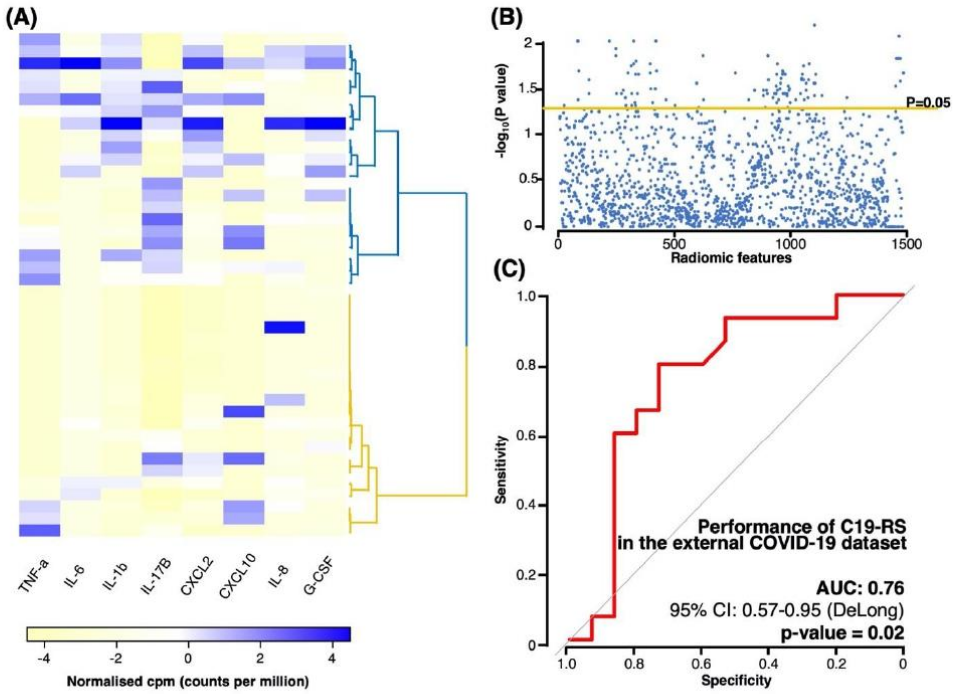


Fig. A4: In pilot data, we have trained a radiotranscriptomic signature of peri-internal mammary artery fat, to detect activation of COVID-19-related inflammatory pathways (cluster activation of TNFa, IL6, IL1b, IL17B, CXCL2, CXCL10, IL8, GCSF) defined by RNA sequencing in internal mammary arteries. We clustered a set of patients into groups of high versus low activation of key inflammatory genes (**A**). In univariate ROC analyses radiomic features extracted from segmentations of adipose tissue surrounding the right internal mammary artery in CCTA images were able to detect vascular inflammatory activation. The dotted line represents the significance level set to $\alpha = 0.05$ (**B**). Subsequently we trained a XGBoost model, the C19-Radiotranscriptomic signature (C19-RS), to differentiate patients with high vascular inflammation seen in COVID-19. We next tested the performance of C19-RS in an external validation dataset of 30 patients (15 COVID-19 positive vs 15 COVID-19 negative) undergoing non-gated CTPA (**C**).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

3.6. Study Hypothesis

We hypothesise that biomarkers of metabolic risk could predict clinical outcomes in patients undergoing computed tomography angiography. Specifically, we hypothesise that:

1) Imaging biomarkers of body adiposity are associated with plasma biomarkers of metabolic risk.

2) Imaging/plasma (e.g.*SFRP5*)/genetic (e.g. *SFRP5* genetic polymorphisms) biomarkers predict vascular disease severity (such as extent of coronary atherosclerosis, vascular calcification, aortic aneurysm size).

3) Imaging/plasma (e.g. *SFRP5*)/genetic (e.g. *SFRP5* genetic polymorphisms) biomarkers predict future cardiovascular events and vascular disease progression in subjects undergoing computed tomography angiography.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

4. OBJECTIVES AND OUTCOME MEASURES

4.1 Primary Objective

To investigate whether biomarkers of metabolic risk can predict major adverse cardiovascular events

Primary outcome measures:

1. Measurements of plasma/imaging markers of cardiometabolic risk
2. Atherosclerosis progression by Computerised Tomography
3. Major adverse cardiovascular events over 10 years

4.2 Secondary Objectives

Secondary objective 1: To identify novel biomarkers able to predict cardiovascular disease pathogenesis and extent of pre-existing vascular disease

Secondary outcome measures (objective 1):

1. Measurement of coronary plaque burden/wall thickness by Computerised Tomography
2. Measurement of coronary calcification by Computerised Tomography
3. Measurement of aortic dimensions by Computerised Tomography
4. Measurement of changes in cardiac structure by Computerised Tomography
5. Measurement of novel indices of adipose tissue expansion and distribution, by using Computerised Tomography and/or ultrasound

Secondary objective 2: Develop and validate imaging biomarkers to predict clinical outcomes in patients with COVID-19 infection.

Secondary outcome measures (objective 2):

1. Measurement of modification to cardiovascular disease pathogenesis and modulation of pre-existing vascular disease that is attributable to infection with coronavirus disease 2019 (COVID-19).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

5. STUDY DESIGN

5.1. Summary of Study Design

This is a multi-centre cohort observational study involving patients undergoing computed tomography angiography (CTA) or who have had a CTA or CT chest, abdomen and pelvis scan where at least 6 months have elapsed and patients attending outpatient clinics.

Study Arms 1, 2 and 3

We will recruit patients attending routine clinical appointments in the cardiology outpatient clinics of the OUH NHS Foundation Trust (Study Arm 1, n=500) or those undergoing routine clinical appointments in multiple centres for CTA scans (Study Arm 2, n=10,000) or those who underwent routine clinical appointments in multiple centres for CTA scans (Study Arm 3, n=5,000).

Participants undergoing a clinical CTA (Study Arm 2) will be asked to participate in a research CTA scan at least 6 months after their enrolment in the study, to assess for cardiovascular disease progression (this would be optional). At this visit, a sample of blood will be collected, and anthropometric measurements taken.

For study arm 3, individuals where at least 6 months have elapsed from a clinical or research CTA scan may also be invited for a repeat CTA scan depending on scan capacity/availability. If agreeable they will be recruited under arm 3a. Patients not having a repeat CTA scan will be invited to participate in the follow-up study arm 3b choosing their preferred level of engagement.

We will ask for the participant's informed consent to:

1. Collect blood samples during their appointment to measure biomarkers with a predictive value in cardiovascular disease development and progression (e.g. markers of cardiometabolic risk) (Study arms 1, 2 and 3a).
2. Access their medical records on an annual basis for a period of 10 years and record information, such as major adverse cardiovascular events (e.g. death, COVID-19 infection/COVID-19 suspected infection acute myocardial infarction, cerebrovascular accident) and hospital admissions (Study Arms 1, 2 and 3) through linkage with local and national databases (including NHS Digital, the National Institute Of Cardiovascular Outcomes Research – NICOR, Sentinel Stroke National Audit Programme – SSNAP, Office for National Statistics – ONS or any other NHS affiliated service or registry that collects patient data).
3. Telephone participants every 2 years to follow up for clinical events, hospitalisations and changes in their medication status (Study Arms 1, 2 and 3).
4. Use clinical images obtained for clinical purposes (e.g. CT angiography scans, ultrasound scans and others) to perform measurements of imaging biomarkers of cardiometabolic risk and assess the severity of vascular and cardiac disease (Study Arms 1, 2 and 3).
5. Take anthropometric measurements during their visit that includes height, weight, and waist-hip-ratio (Study arms 1, 2 and 3a).

One of the long-term aims of the study is to create a unique cohort of participants that will allow the potential for subgroups to be formed and called back for future studies. Participants when consented for the ORFAN study will also be given the option to opt in/out of being recalled for any future studies. Any

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

such studies that are created will be submitted for NRES review. The study will be included also in the NIHR Bioresource.

Study Arm 4

We will look to collect data from patients who have undergone clinical CTAs or unenhanced CT chest, abdomen and pelvis scans, and post-mortem CTs in the centres participating in the study (up to 250000 patients). Images and results will be pseudonymised on site of clinical care or using a cloud gateway like CIMAR/Ambra or equivalent. This study arm focuses on collecting CT images, patient demographics, clinical information including outcomes to aid the development of new image analysis algorithms and software tools (using artificial intelligence and other approaches) to allow automation of the image analysis processes (such as the automated calculation of Fat Attenuation Index) used in aims 1-3. This will also allow us to confirm the specificity, utility and predictive value of the optimised algorithms/software tools. By analysing the CT data from all sites, we will also identify vulnerable patients within an otherwise "low risk" group. Recording the landscape of the distribution of these new imaging biomarkers , will inform a future health economic study that will enable better understanding of their health economic impact for the NHS (e.g. by identifying patients who may not otherwise have been understood to be at risk, guiding deployment of new treatments to improve clinical care and outcomes in the future). This study arm will allow us to have enough statistical power to use AI approaches, including deep-learning automated image analysis and machine learning statistical analysis for the generation of risk scores in order to achieve the secondary objective: i.e. to develop and validate robust novel imaging biomarkers using cardiac CT, that would describe the extent of pre-existing vascular disease.

5.2. COVID-19 research

Following the COVID-19 pandemic, the ORFAN study programme has been designated by the NIHR and the British Heart Foundation as a National Flagship programme (UK COVID-19 Cardiovascular CT Research Consortium or UK C19-CRC) in our concerted efforts to study and treat the disease. We will specifically ascertain COVID-19 status (suspected or confirmed) for arms 2, 3 and 4. The influence of COVID-19 infection on the cardiovascular disease outcome measures of the study will be explored. Arm 4 will be critical to the understanding of the effects of COVID-19 infection on the risk of cardiovascular diseases and also the impact of pre-existing cardiovascular disease on the severity of COVID-19 infection.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Study Flow Diagram

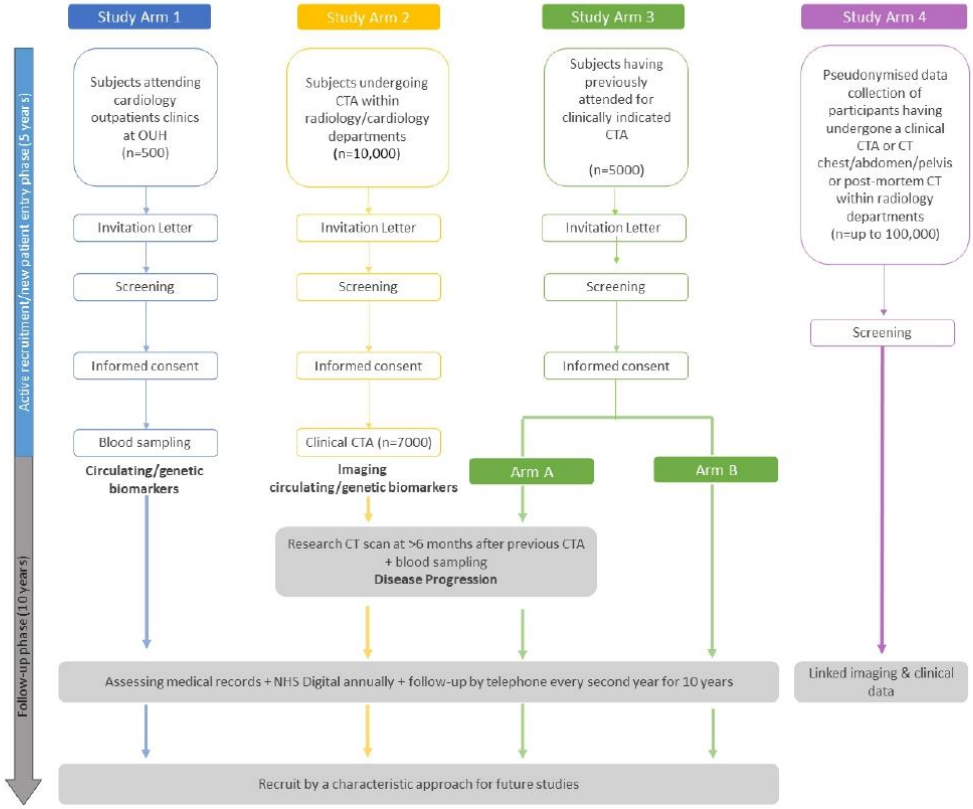


Figure B - Study Design.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

5.3. Clinical Protocol

a) Informed Consent

i) Study Arm 1 (individuals attending the outpatient clinics): A total of 500 participants attending Cardiology Department, Hypertension or University outpatient clinics at the Oxford University Hospitals (OUH) NHS Foundation Trust will be enrolled to the study. Patients attending Hypertension or University outpatient clinics will be mailed a Participant Information Leaflet (PIL) before their clinical appointment. At their appointment, a member of the clinical care team (or researchers acting under arrangements with the responsible care organisation) will obtain verbal consent for a member of the research team to approach them. If they agree, the patient will be asked to sit back in the waiting room once their clinical appointment is over. A study investigator holding an honorary contract with the OUH NHS Foundation Trust will then introduce themselves to the potential research participant and discuss further the ORFAN study. The PIL and the Informed Consent documents will be presented to the participant detailing the exact nature of the study; what it will involve for the participant; the implications and constraints of the protocol; the known side effects and any risks involved in taking part. It will be clearly stated that the participant is free to withdraw from the study at any time for any reason without prejudice to future care, and with no obligation to give the reason for withdrawal. The participant will be allowed as much time as they wish to consider the information, and the opportunity to question the investigator. When the PIL cannot be mailed (e.g. non-elective cases) there will be particular effort to maximise the time available for the participant to consent by providing the information leaflet as soon as possible, giving the patient enough time to read and understand the documents, discuss them with relatives and ask questions. The participants will be allowed to opt out of specific parts of the protocol, such as the participation in genetic studies.

If the patient is happy to participate, informed consent will be obtained where the latest version of the study consent form will be signed and dated by both the participant and the study investigator. A copy of this consent form will be given to the participant, another copy will be filed in the participants medical notes and the original documents will be retained at the study site in a locked cupboard.

ii) Study Arm 2 (patients having CTA for clinical reasons or as part of other research studies): A total of 10000 participants undergoing CT Angiography for clinical reasons will be included in the study. The PIL will be mailed to all patients undergoing a CTA before their clinical appointment. When this is not possible (e.g. in non-elective cases), there will be particular effort to maximise the time available for the participant to consent, by providing the information sheet as soon as possible on the ward prior to the patient going for their scan, and giving the patient enough time to read and understand the documents, discuss them with their relatives and ask questions. After agreeing with a member of the clinical care team (or researchers acting under arrangements with the responsible care organisation) that they are happy to discuss research opportunities, a member of the study team will approach the potential candidate to discuss the ORFAN study. Once the consent form has been completed, the study team will have access to the participant's medical notes to screen for eligibility and include/exclude as necessary. The participants will be allowed to opt out of specific parts of the protocol, such as the participation in genetic studies or the follow up CTA.

iii) Study Arm 3 (individuals having previously attended for a clinical or research CTA scan including patients who had COVID-19 infection requiring hospital admission): A total of up to 5000 patients, who had a CTA scan (as part of either their clinical care or as part of a research study) including patients who had COVID-

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

19 will be the potential candidates for this study arm (Figure B). We would contact these participants initially by letter to ask them to consent to:

a) Their image data being linked with data obtained from their medical records and
b) Partake in the follow up part of the study

Once initial written contact has been made a member of the study team will telephone the potential participant, introduce themselves and explain the study protocol in more detail. Depending on scan capacity/availability, individuals where at least 6 months have elapsed from a clinical or research CTA scan may also be invited for a repeat coronary CTA scan. If agreeable, then a provisional appointment will be made over the telephone and this will be followed by an appointment letter from The Nuffield Health – Manor Hospital, Oxford or one of the various radiology/cardiology departments participating in the study. Consent in this case will be obtained at the time of the appointment (Consent 1).

If the participant is not offered or declines a research coronary CTA (Arm 3a) they will be presented with the option to participate in the follow-up part of the study (Arm 3b). If they decide to participate in the follow-up part of the study, they will be asked to sign and date the consent form (Consent Form 2) and return it in the pre-paid envelope accompanying the invitation letter. The study team member will then sign and date the original consent once returned by the participant and save it into the study folders. A copy of the signed consent will be returned to the participant along with a thank you card and a second copy placed/uploaded into the participant's medical notes.

All potential participants will be referred to the study-specific webpage (www.oxhvf.com/the-orfan-study) for further information. The ORFAN webpage is embedded in the 'The Oxford Heart, Vessels & Fat (OX-HVF) cohort homepage. The OX-HVF cohort consists of a cluster of ethically approved clinical studies (namely ART Vascular, Bypass Grafts, AdipoRedOx, ORFAN and IMPACT), which provide synergistic results allowing the deployment of a multi-level strategy to understand the mechanisms of cardiovascular diseases. As each one of the individual studies approaches the issue of cardiovascular disease pathogenesis from a different angle, the OX-HVF cohort provides a unique and powerful platform that provides direct access to human tissue (vessels, myocardial and fat biopsies, DNA, plasma and others), in combination with extensive non-invasive cardiovascular phenotyping that includes cardiovascular computed tomography angiography, ultrasound and others. The OX-HVF cohort also collects prospective 10-year clinical outcome data and is linked to the NIHR Bioresource.

iv) Study Arm 4 (collection of pseudonymised data):

Individual patient consent will not be sought for Arm 4 as patient identifiable data will not be recorded by the study team in the data collection strategies described below.

UK Sites:

1) The UK participating sites are NHS Trusts. At the participating sites the clinical care team is responsible for the identification of patients who are eligible for inclusion in the study. The clinical care team also works with local researchers, such as research nurses in cardiovascular care units, who operate under the direction of the clinical care team and with appropriate contracts at the local site. All staff operate in keeping with obligations of patient confidentiality and local information governance. The clinical care team at individual participating sites (or, as mentioned, researchers acting under arrangements with the responsible care organisation) identifies the NHS/Hospital numbers of the eligible participants according to the study protocol (i.e. patients who

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

have undergone a specific type of CT scan at their Trust). Where a patient opt-out is in place, their data will not be used.

2) Within the local NHS Trust firewall, a locally populated database that includes the assigned study ID and relevant identifiable information alongside relevant non-identifiable study variables about the participant is constructed. Collected patient identifiable information include NHS number, hospital number, date of birth, date of death if relevant, gender, and address (postcode level).
3) At a local NHS Trust site level, the clinical care team (or researchers acting under arrangements with the responsible care organisation) identifies relevant CT scans of patients on the local PACS. The participating site clinical care team (or researchers acting under arrangements with the responsible care organisation) assigns the corresponding participant study ID to the scan in order to pseudonymise the scans on an ISO/IEC 27001 certified portal, such as CIMAR/Ambra or equivalent. The pseudonymised scan is sent to the study core lab at the University of Oxford via the same portal. Alternatively, the participating site team manually pseudonymises the scans and sends them to the study core lab via encrypted hard drives.
4) Clinical care teams at each participating site (or researchers acting under arrangements with the responsible care organisation) send to the study core lab their locally populated database containing the relevant, non-identifiable study variables after patient identifiable information is deleted. No patient identifiable data is ever shared with the study core lab. The study core lab will only receive study data that is non-identifiable. The individual clinical teams at the NHS Trusts (or researchers acting under arrangements with the responsible care organisation) where the patient was treated and subsequently found to be eligible for inclusion in ORFAN Arm 4 will be the only holders of their site's identifiable information.
5) Section 251 support for UK Arm 4 patient records will be sought as a data linkage method to complement the local linkage procedures described in 1-4. With section 251 support, we will instigate flow of personal identifiable data from participating sites to third parties with approval to hold linked patient data (e.g. NHS digital/NICOR/SSNAP/ONS or any other NHS affiliated service or registry that collects patient data). The third parties will be supplied by the participating site clinical care team (or researchers acting under arrangements with the responsible care organisation) with the database containing site-specific participant study IDs alongside the required identifiable information to perform data linkage (NHS number, Hospital ID number, Date of Birth, gender, Address (postcode level) . Where a patient opt-out is in place, their data will not be used.
6) Once third parties have performed the data linkage (linkage is done on annual basis for 10 years) between their databases and the study site supplied database, patient identifiable data are deleted by the local NHS Trust sites. This will create one database containing the third party extracted variables alongside the participant study ID. This database can then be sent to the study core lab, via secure communication means. The third party that performed the data linkage (e.g. NHS digital/NICOR/SSNAP/ONS or any other NHS affiliated service or registry that collects patient data) will keep the linkage file for 10 years to update the relevant outcomes annually.

With the above processes, the study investigators who are not members of the clinical care team (or researchers acting under arrangements with the responsible care organisation) will not be given access to patient's identifiable information at any stage. The sites (local clinical care team or local researchers acting under arrangements with the responsible care organisation) will destroy the linkage file at the end of the study, when data linkage is completed.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Data received from centres outside of the UK

Collaborators outside of the UK will be provided with the variables the ORFAN study is collecting to accompany CT scan. Their collection and curation of data will be subject to local requirements and regulations. ORFAN will receive data from international centres only in anonymised form. This will either be retrospective only or, where centres are able and willing to provide ongoing data related to cases, will be updated at intervals. The study core lab will receive this de-identified data either via an established portal or via encrypted hard drives.

b) Completion of Study Questionnaire

A study team member will be responsible for filling out a study questionnaire for each participant in Arms 1, 2 and 3. The process of filling out the study questionnaire will be brief (around 10 minutes) and administered in an area away from other patients/or via telephone. For those participants who had a past CTA and have returned for a research CTA scan (Study Arm 3a), the questionnaire may be completed at their study visit. For participants under Arm 3b this questionnaire may be administered over the phone. Any inconsistencies will be cross-checked with the participants' clinical notes, after they have been consented.

c) Blood Sampling

Whole blood will be used for measurement of plasma/serum biomarkers of metabolic risk (e.g. *SFRP5* or others) and genotyping for variants associated with cardiometabolic risk.

i) For patients attending the outpatient clinics (Study Arm 1): A small quantity of blood (~30ml) will be drawn by a peripheral vein of the arm by an appropriately trained study team member after the consultation.

ii) For patients having CTA (Study Arm 2): A study team member will draw a small quantity of blood (~30ml) from the peripheral vein cannula (when possible) which is routinely inserted for the clinical purposes of the CT angiography. Alternatively, a peripheral vein of the arm will be used for the blood sampling at the end of the CT scan.

iii) For participants who had a previous CTA and are attending for a research CTA scan (Study Arm 3a): Blood samples will be collected only in those who are attending for a research CTA scan, during their CTA scan procedure. A study team member will draw a small quantity of blood (up to30ml) from the peripheral vein cannula (when possible) which is routinely inserted for the clinical purposes of the CT angiography. Alternatively, a peripheral vein of the arm will be used for the blood sampling at the end of the CT scan.

d) Follow-up of Participants, Access of Medical Records and contacting patients for participation in future studies

The participants (from Study Arms 1, 2 and 3) will be contacted by telephone every second year for the 10 years after consent at the initial appointment. They will be followed up for clinical events, hospitalisations and changes in their medication status (Appendix B). For the same purpose, medical records will be assessed on an annual basis for a total of 10 years after enrolment to the study. Access to medical records will be used to record data on demographic characteristics, previous medical history, and current medication status and also to prospectively record occurrences of major adverse cardiovascular events, admissions to hospital; outpatient clinics attended; medication administered. A member of the study team will also collect data, prior to making contact by telephone, from NHS digital, NICOR, SSNAP, ONS or any other NHS affiliated service or registry that collects patient data for the sites within England and Wales or the Information and

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Statistics Division of the NHS in Scotland, the Office for National Statistics (ONS) or equivalent third bodies (non-UK sites), to obtain data on mortality, hospital episode statistics and other adverse events.

For those patients that have imaging data available (e.g. CT angiography, echocardiography data or other) the raw images will be accessed by study investigators to assess indices of cardiometabolic risk. The images will be retrieved from NHS archives or equivalent (non-UK sites), pseudonymised and transferred to specialised workstation for further analyses. Since the study will be under the NIHR Bioresource, the local NIHR coordinator will have access to the identifiable information of those patients who agree to be contacted for future studies (including ethically approved studies under the NIHR Bioresource).

e) Coronary CT Angiography

Participants who enrolled to Study Arms 2 and 3a and consented for participation in the follow up phase of the study, will be invited to attend a research CTA at least 6 months after their initial appointment (Study Arm 2) or at any time from the time of their recruitment, as long as it is at least 6 months after their previous CTA scan (Study Arm 3a). CTA images will be used to assess atherosclerotic progression by comparison with the available baseline CTA images from the clinical /research CTA scan. Participants in Study Arm 2 will be made aware by the research team of this optional additional CTA at their initial clinical appointment.

Computed Tomography Angiography at the research study visit (Study Arms 2 and 3a) will be performed by using a standard protocol that is routinely used in clinical practice. Briefly, participants with an elevated heart rate (>60bpm) will receive intravenous metoprolol to achieve a heart rate ≤60 bpm. Additionally, participants will receive sublingual nitroglycerine (0.4 mg/dose, one to two doses per participant) directly before the test. A calcium score scan will be undertaken followed by a contrast enhanced scan. Vessel opacification for the contrast enhanced scan will be achieved by IV injection of up to 100mls of iodinated contrast through a peripheral vein, followed by a saline flush, as part of the routine clinical practise. For the Study Arm 3a participants, it will be explained that the examination will be performed exclusively for research purposes and therefore they won't be directly informed about the results of the scan. However, in the case of any clinically significant incidental findings, the General Practitioner (GP) will be informed, but only in the cases where the participant has consented for the study investigators to communicate with the relevant physician if appropriate. The opinion of a relevant healthcare specialist will also be sought.

Images collected by a study investigator from the research CTA will be compared to the baseline imaging data available from the initial CTA scan in order to assess changes in imaging biomarkers of cardiometabolic risk and atherosclerosis progression. This will allow us to explore whether baseline information on adipose tissue imaging can predict vascular disease development or progression. Study team members will routinely use the web based Picture Archiving and Communications Systems (PACS) or similar to review the research CTA scans and report any incidental findings to the referring consultant doctor and the GP (where the participant has consented to this).

CT imaging will be undertaken with an estimated effective dose of 14mSv. The additional lifetime cancer risk associated with this effective dose is estimated to be 1:1300.

For participants attending Milton Keynes University Hospital the estimated effective dose will be 18mSv. The additional lifetime cancer risk associated with this effective dose is estimated to be 1:1000.

To compensate participants for their time and effort, those who attend a follow up study visit will be offered a reimbursement for their travel expenses.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

5.4. Experimental Design

a) Measurement of Plasma/Serum Biomarkers (Arms 1, 2 and 3a): Plasma/serum will be collected and stored at -80ºC by a study investigator until it will be assayed by a University researcher to measure biomarkers of cardiometabolic risk (such as circulating *SFRP5*, C-reactive protein and others).

b) DNA Extraction and Genotyping (Arms 1, 2 and 3a): DNA will be extracted from whole blood and used for genotyping of genetic variants that might play a role in cardiometabolic risk or cardiovascular disease pathogenesis (such as single nucleotide polymorphisms affecting *SFRP5* expression and others).

c) *Isolation of Cells (Arms 1, 2 and 3a):* Whole blood will be used for isolation of various cell types (e.g. peripheral blood mononuclear cells, endothelial progenitor cells and others) which will be phenotyped for cellular signatures of cardiometabolic and cardiovascular disease pathogenesis (e.g. *SFRP5* gene expression).

d) Analysis of CT Images (all Arms): CT images will be retrieved from the medical archives, pseudonymised and used by study investigators and potential partner commercial companies (e.g. Caristo Diagnostics) for:

1. Adiposity measurements: quantitative and qualitative measurement of adipose tissue depots, e.g. in perivascular, visceral, subcutaneous, epicardial and pericardial adipose tissue depots.

2. Assessment of vascular disease severity: this will include measurements of vascular disease severity such as the extent of coronary atherosclerosis, atheroma burden, vascular calcification, aortic aneurysm diameter and others.

3. Assessment of cardiac disease severity: This will include measurements of cardiac anatomy, cardiac structure, hypertrophy and others.

e) Analysis of Echocardiography and Other Images (all Arms): Any available cardiac echo images will be retrieved from medical archives, pseudonymised and used for measurements such as assessment of cardiac adiposity by measurement of epicardial fat pad depth, cardiac structure/function, cardiac disease severity (e.g. cardiac hypertrophy) etc. When other types of medical images are available (e.g. intravascular ultrasound, optical coherence tomography images and more), these may also be used to gain further information on the extent of cardiovascular disease.

f) Prospective Record of Cardiovascular Events: Medical records will be assessed on an annual basis for a total of 10 years (by a study team member) after enrolment to the study, to record major adverse cardiovascular events, admissions to hospital, outpatient clinic attendance. Data will also be collected from the ONS and NHS Digital (a central source, sponsored by the Department of Health, to provide national information, data and IT services for patients, clinicians, commissioners and researchers), NICOR, SSNAP or any other NHS affiliated service or registry that collects patient data or equivalent third body (non-UK sites). Identification of adverse events will allow the investigators to assess the value of biological markers for vascular, adipose and myocardial tissue in predicting the long-term outcome of patients. For those in Arms 1, 2 and 3 a telephone follow-up to the participant every 2 years (*Appendix A*) will enable investigators to increase the granularity of events ascertained by record linkage with the aforementioned organisations.

All analysis and reporting will be carried out on the pseudonymised data and samples by researchers at the University of Oxford.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

6. PARTICIPANT IDENTIFICATION

6.1. Study Participants

The study participants will be subjects aged 18-99 years, attending OUH NHS Foundation Trust outpatient clinics (Study Arm 1) or the radiology / cardiology department of the selected hospital for a Computed Tomography angiogram (Study Arm 2) or subjects who had a CTA or CT chest, abdomen and pelvis scan at the selected radiology / cardiology department and where at least 6 months have elapsed (Study Arm 3). Potential participants will be informed of the study prior to their visit in the form of the Participant Information Leaflet (PIL) that will be posted to them in advance. A study investigator, holder of a clinical contract with the respective NHS Trust or similar (non-UK sites), will be responsible for screening the medical records for eligibility criteria once the participant has consented to the study. The consented participants will be also screened by the clinical care team (or researchers acting under arrangements with the responsible care organisation) to identify those who have had COVID-19 infection. When it is not feasible to send out the PIL in advance (e.g. non-elective cases) then the potential participant will be approached on the day of the clinical appointment by a member of the clinical care team (or researchers acting under arrangements with the responsible care organisation) and if they indicate they are willing to receive further information they will be introduced to a member of the study team.

Study arm 4 is a pseudonymised cohort and participants will not be sought for consent. Data will be from patients between 18-99 years who have attended one of the participating sites' CT departments for a CTA or CT chest, abdomen and pelvis scan or post-mortem CT and are not participants of study arms 1, 2 or 3. COVID-19 is one of the clinical diagnoses that will also be collected through the national databases, e.g. NHS digital, in many Arm 4 patients.

6.1.1. Inclusion Criteria for study Arms 1, 2 and 3

- Participant is willing and able to give informed consent for participation in the study.
- Male or Female, aged 18 -99 years.

6.1.2. Inclusion Criteria for study Arm 4

- Male or Female, aged 18 -99 years.

6.1.3. Exclusion Criteria for Study Arms 1, 2 and 3

- Unable or unwilling to consent
- Active cancer

6.1.4. Exclusion Criteria for Study Arm 4

- Participation in Study Arms 1, 2 or 3
- Existing opt-out from use of data for research purposes

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

7. STUDY PROCEDURES

7.1. Recruitment

All eligible patients referred for a Computed Tomography Angiogram (Study Arm 2) at the selected hospital or attending the outpatient clinics at the OUH NHS Foundation Trust (Study Arm 1) will receive the study PIL by post along with their clinical appointment letter, and will then be approached initially by a member of the clinical care team (or researchers acting under arrangements with the responsible care organisation), and if they express interest in the research, by a study investigator to gain informed consent. Consent will be obtained either prior to their CT scan (Study Arm 2 and 3a) or during their clinic appointment (before or after their consultation, depending on the timings of their appointment – Study Arm 1). When the PIL cannot be mailed (e.g. non-elective cases) there will be particular effort to maximise the time available for the participant to consent by providing the information sheet as soon as possible on the ward.

For Study Arm 3, patients who have undergone a CTA at the selected hospital site including patients who were hospitalised with COVID-19 infection and at least 6 months have elapsed, will be identified by the clinical care team (or researchers acting under arrangements with the responsible care organisation) and sent a PIL and an invitation letter to participate in the study. The letter will include a pre-paid envelope. Following this initial contact, a study investigator will contact the participants by phone or email in order to answer any queries they may have. They may then be invited to participate in the follow up part of the study that includes an optional Coronary/CTA at least 6 months after their previous CTA (study Arm 3a) and 10 year clinical follow up (Arm 3). If they agree to the CT scan, then they will be consented at the time of their visit by a study investigator. If the potential participant declines/is not offered a CT scan but would like to be prospectively followed-up, they will be asked to sign, date the consent form (Consent form 2) and return it in the pre-paid envelope. The returned consent form will then be signed by a study investigator and a copy returned to the participant.

For study Arm 4, screening will be performed using an electronic search of CT scans and electronic case records at each hospital site and suitable cases will also be identified by the clinical care team. Patients who had a CTA in the past and are not part of Arms 1, 2 and 3 will be study Arm 4 participants. A member of the study team who is also a part of the clinical care team will check for duplication. The same process will be followed to identify and screen patients who had COVID-19 and a previous CTA.

7.2. Screening and Eligibility Assessment

All subjects 18-99 years of age, will be eligible to enter the study (see also inclusion/exclusion criteria for different study arms). A study investigator, holder of a clinical contract with the respective NHS Trust or equivalent (non-UK sites), will be responsible for screening the medical records of participants.

7.3. Baseline Assessments

a) ***Blood Samples:*** Plasma/serum will be collected and stored in a pseudonymised form at -80°C by a study investigator until it will be assayed by a University researcher to measure biomarkers of cardiometabolic risk (such as circulating sfrp5, C-reactive protein and others). DNA will be extracted from whole blood and used for genotyping of genetic variants that might play a role in cardiometabolic risk or cardiovascular disease pathogenesis (such as single nucleotide polymorphisms affecting SFRP5 expression and others). Whole blood will be used for isolation of

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

various cell types (e.g. peripheral blood mononuclear cells, endothelial progenitor cells and others) which will be phenotyped for cellular signatures of cardiometabolic and cardiovascular disease pathogenesis (e.g. sfrp5 gene expression). These measurements will be correlated with any imaging data undertaken at the time.

b) ***Analysis of CT Angiography Images:*** CT images will be retrieved from the medical archives, pseudonymised and used by study investigators and potential partner commercial companies (e.g. Caristo Diagnostics) for various measurements including the assessment of vascular disease severity and cardiac disease severity and also adiposity measurements.

c) ***Analysis of Echocardiography and Other Images:*** Any available cardiac echo images will be retrieved from the medical archives, pseudonymised and used for measurements such as assessment of cardiac adiposity. When other types of medical images are available (e.g. intravascular ultrasound, optical coherence tomography images and more), these may also be used to gain further information on the extent of cardiovascular disease.

7.4. Subsequent Visits and Long-Term Follow Up

Participants who have consented to the study will be followed up every second year via telephone (Appendix A) for a 10 year period post initial appointment or from obtaining their consent. They will be followed up for clinical events, hospitalisations and changes in their medication status and used to record patient demographics.

Participants who were consented to Study Arms 2 and 3a will be invited to attend a follow up research Coronary/ CTA scan at least 6 months after their initial appointment and have a sample of blood taken at this visit. The CT images will be analysed as stated in section 5.3.d and the blood sample will be taken to measure plasma/serum biomarkers for their correlation to the image data.

7.5. Sample/Image Handling

Blood (approximately 30ml) will be taken at the baseline stage and at the follow up appointment by one of the study investigators who holds a contract with one of the participating NHS Trusts or equivalent (non-UK sites) or a member of the clinical care team. It will then be centrifuged, the cells isolated and then stored temporarily on site until they are sent coded to the core laboratory at the University of Oxford. Measurements will be taken of the plasma/serum biomarkers of metabolic risk (e.g. SFRP5 or others). The whole blood will be used for isolation of cell types which will be phenotyped for cellular signatures of cardiometabolic and cardiovascular disease pathogenesis and genotyped for variants associated with metabolic risk.

The blood and extracted cells and image data will be stored in a pseudonymised form and processed initially by a local researcher and those linked to future ethically approved studies. However, to better characterise the samples collected in the context of this study, the expertise of scientific collaborators from other centres in the UK or abroad may be sought. Samples and imaging data collected for research purposes may be sent to other centres for specific analyses, according to the availability of specialist techniques that are not available to the local research team. Moreover, pseudonymised data and/or biological samples may be sent to other centres participating in the National Institute for Health Research (NIHR) Bioresource for further processing. As all stored samples/data are identified by study code only, this does not raise confidentiality issues; where necessary material transfer agreements will be put in place.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

7.6. Discontinuation/Withdrawal of Participants from Study

Each participant has the right to withdraw from the study at any time. In addition, the Investigator may discontinue a participant from the study at any time if the Investigator considers it necessary for any reason including: Withdrawal of Consent.

The date and reason for withdrawal will be recorded on the study database.

7.7. Definition of End of Study

The end of the study is defined as the time point when the 10-year follow up of the last participant of the study is completed.

7.8. Definition of Serious Adverse Events

A serious adverse event is any untoward medical occurrence that:

- Results in death
- Is life-threatening
- Requires inpatient hospitalisation or prolongation of existing hospitalisation
- Results in persistent or significant disability/incapacity
- Consists of a congenital anomaly or birth defect

Other 'important medical events' may also be considered serious if they jeopardise the participant or require an intervention to prevent one of the above consequences.

NOTE: The term "life-threatening" in the definition of "serious" refers to an event in which the participant was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe.

The study's procedures are limited to blood sampling and a CTA, so the risk for an adverse event is minimal.

7.9. Reporting Procedures for Serious Adverse Events

A serious adverse event (SAE) occurring to a participant will be reported to the REC that gave a favourable opinion of the study where in the opinion of the Chief Investigator the event was 'related' (resulted from administration of any of the research procedures) and 'unexpected' in relation to those procedures. Reports of related and unexpected SAEs will be submitted within 15 working days of the Chief Investigator becoming aware of the event, using the HRA report of serious adverse event form (see HRA website).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

8. STATISTICS AND ANALYSIS

8.1. Description of Statistical Methods

Exploratory data analysis will be performed for all variables derived from this study to obtain the appropriate summary statistics. Where necessary, transformation of data will be performed for further statistical testing. Between patient relationships of plasma/imaging biomarkers will be assessed using appropriate hierarchical or multilevel data analysis.

For the participants attending outpatient clinics: Subgroup analysis within the population will be performed to explore the correlations of the plasma biomarker of interest with specific clinical risk factors (e.g. in diabetic patients' vs non-diabetic, hypertensive vs non-hypertensive etc.) using t-tests for continuous variables between groups, chi-square tests for categorical variables or other tests as appropriate. Linear regression models will be used to search whether circulating plasma metabolic biomarkers (e.g. SFRP5 plasma levels) predict presence of certain disease traits (e.g. diabetes) independently of other classic risk factors.

For participants undergoing CT angiography: Linear regression models will be used to search whether biomarkers of body adiposity are predictors of vascular disease severity. Each measure of vascular disease severity (e.g. atheroma burden, vascular calcification) will be used as a dependent variable. Independent variables will include factors such as imaging or plasma biomarkers of metabolic risk, and patients' demographic characteristics (e.g. age, gender, hypertension, body mass index, diabetes, smoking, and hypercholesterolemia, COVID-19/suspected COVID-19). Known cardiovascular risk factors (regardless of p-value) plus any other variable which show a possible association with vascular disease severity in bivariate analysis ($P<0.15$) will be initially selected for the construction of the regression models. Appropriate diagnostics will be performed to ensure the assumptions are met for the regression model of choice. Comparisons of continuous variables (e.g. vascular calcification) between participants below the lowest tertile and above the highest tertile of perivascular adipose volume will also be attempted. When splitting variables lead to more than two groups, continuous variables will be compared across the groups by using a multiple comparisons test such as one-way ANOVA (followed by Bonferroni post-hoc correction for individual comparisons as appropriate).

Table: Example of correlations that will be investigated

Correlation of:	***With:***
Pericardial adipose tissue volume	1. coronary calcification 2. coronary atheroma burden 3. plasma sfrp5 (plus other adipokines)
Epicardial adipose tissue volume	1. coronary calcification 2. coronary atheroma burden 3. plasma sfrp5 (plus other adipokines)
Pericoronary adipose tissue volume	1. coronary calcification 2. coronary atheroma burden
Periaortic adipose tissue volume	1. aorta calcification

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

In the whole population of the study: The participants attending outpatients' clinics will be used as a pool of low cardiovascular risk individuals (control group) to explore the predictive value of any identified biomarkers in this population. In the total study population identified biomarkers will be correlated with known cardiovascular risk factors (e.g. systolic/diastolic blood pressure levels, plasma lipids, waist circumference etc.). Adjustments for multiple testing will be applied. Furthermore, in the total study population and in each group of participants separately the risk of major adverse cardiovascular events (such as myocardial infarction, cardiovascular mortality, and total mortality) during the follow up will be assessed by using appropriate Kaplan Meier curves. The log-rank test will be used to compare Kaplan-Meier curves between groups stratified by groups (e.g. median or tertiles) of the biomarker of interest. Appropriate Cox proportional hazards models will then be applied by using the imaging/plasma biomarker of interest as a continuous variable, and its effect on the risk for cardiovascular mortality or other endpoints will be adjusted for other covariates, and the respective hazard ratios (HR)[95%CIs] will be calculated. Similar approach will also be used to test the prognostic value of other variables on clinical outcome.

The identified biomarkers and the constructed prognostic models will be calibrated by examining their ability to correctly estimate the likelihood of a future event across the whole range of prognostic estimates e.g. by grouping participants into deciles of estimated probabilities of coronary events and calculating for each decile the sum of predicted and observed events and applying the Hosmer–Lemeshow test. Discrimination measures of the constructed prognostic models will be also assessed by calculating the C-statistics (C-index).

Finally, the prognostic value of any identified biomarkers will be sought to be further validated in other prospective cohorts within the collaboration research network of the study (e.g. within NIHR Bioresource). This will allow us to test the prognostic value of the identified biomarkers in other populations too (e.g. in high risk subjects such as diabetic subjects or patients with established CVD such as coronary patients, post-acute myocardial infarction etc.).

8.2. The Level of Statistical Significance

All p-values will be based on two-sided tests and compared to a significance level of 5%.

8.3. Procedure for Accounting for Missing, Unused and Spurious Data

Subjects with missing values in specific variables will be excluded from the analysis of those particular variables.

8.4. Power Calculations

For power calculations we have used recent data from our studies that have revealed a low incidence of adverse cardiac events among individuals referred for clinically-indicated CTA. Based on data from similar cohorts the incidence of cardiac mortality is now estimated at 3.5 events per 1000 person-years of follow-up. By splitting the study population in two groups based on the 90th percentile of a given biomarker (e.g. Fat Attenuation Index of perivascular adipose tissue), we estimated that we need a total of 96 events to detect a Hazard Ratio of 3 between the high- and low- risk patient groups (as defined based on the biomarker), with a power of 90% and a (alpha) set at 0.05. Using the expected incidence rates, and assuming

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

a censoring rate of 20% and a mean length of follow-up of 5 years, it was calculated that the required study population size in Arms 2 & 3a is 6455 participants. This number has been increased to 8500 to account for missing data and participants lost to follow-up.

Arm 4 has a projected total sample of 250000 participants in total (200000 within the UK; 50000 international). This number will enable the power to facilitate exploration of the impact of COVID-19 on those with pre-existing cardiovascular disease and the effects of COVID-19 on the development of cardiovascular disease, including the ethnic diversity necessary for the development of globally applicable imaging biomarkers. We require this large sample size to capture enough participants to fall into all necessary groups for the development of novel risk assessment tools and biomarkers as per the study objectives. These groups are (inclusive only of those who meet the core inclusion criteria for ORFAN Arm 4 - 18-99 y.o and having had a clinically indicated CTA or CT Chest):

1. Those with pre-existing cardiovascular disease (for example atherosclerotic coronary artery disease and/or atrial cardiomyopathy) and who had COVID-19 infection previously or during follow-up.
2. Those with pre-existing cardiovascular disease and who did not have COVID-19 infection previously or during follow-up.
3. Those without pre-existing cardiovascular disease and who had COVID-19 infection previously or during follow-up.
4. Those without pre-existing cardiovascular disease who did not have COVID-19 infection previously or during follow-up.

The total number of participants who will be within each group is not known and will change over the course of the ORFAN Arm 4- 10 year follow up period. For example, a participant may initially be within group 4 (as in, received a CT scan that rendered them eligible for inclusion in Arm 4 but the participant did not have any detectable cardiovascular disease and had never previously suffered from COVID-19 infection). Were this participant to subsequently suffer from COVID-19 after first inclusion in the study they would move to group 3. Were this participant to then go on to suffer from a stroke following COVID-19 infection they would move to group 1. These groups are not separate cohorts but are for demonstrative purposes to explain the size of the sample required to ensure enough participants are included to have statistical power to identify differences between all such groups.

Figures reported in the scientific literature as at July 2021 indicate that up to 9.98% (5.23 million) of the UK adult population (52.4 million total) had experienced COVID-19 infection between February 2020 and July 2021(23). If infection rates are assumed to stay similar for the period of initial enrolment for ORFAN Arm 4, then it is likely that at least some 20000 participants will be enrolled within ORFAN Arm 4 who have already had COVID-19 (i.e. group 1 or 3), however it is not possible to estimate the future percentage of the sample population who will go on to have COVID-19 infection during the 10 year follow up. Given the reported rates of adult heart disease in the UK (7.6 million adults living with known atherosclerotic heart disease as at March 2021(24)) it is reasonable to assume that some 14.5% of the ORFAN Arm 4 cohort will have known cardiovascular disease at the time of enrolment in the study, not counting sub-clinical or non-atherosclerotic heart disease. Assuming that cardiovascular disease patients are equally likely as healthy patients to be infected with COVID-19, we can reasonably expect some 2900 patients at the time of initial study enrolment will have suffered from both COVID-19 and had known heart disease. This figure will rise substantively with subsequent exposure to disease, both COVID-19 and heart disease, over the 10-year follow up.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Power calculations have been performed to demonstrate the sample size required to address the scientific objectives of ORFAN Arm 4. For example, following the same approach as for Arms 2 and 3a, the important scientific question of detecting clinically relevant change in the Fat Attenuation Index of perivascular adipose tissue, and hence likelihood of a coronary artery disease related event, given prior COVID-19 infection in those with pre-existing heart disease can be calculated assuming 2.9% of the study population will have had both COVID-19 and known heart disease and 11.6% having known heart disease only, as outlined above. By splitting these two groups based on the 90th percentile of the Fat Attenuation Index we estimated that we need a total of 121 events in the group with both COVID-19 infection and heart disease to detect a hazard ratio of 1.4 between the high- and low- risk patients in each group (as defined based on the biomarker). This is assuming power of 90% and α (alpha) set at 0.05. Using the expected incidence rates of coronary artery disease, and assuming a censoring rate of 20% and a mean length of follow-up of 10 years, it was calculated that the required study population size in Arm 4 is 205500 participants to enrol enough patients both with and without previous COVID-19 infection. This number has been increased to 250,000 to account for missing data and participants lost to follow-up which is likely to occur in Arm 4 given the high number of clinical sites participating and the quantity of necessary data being collected at each site.

8.5 Interim Analysis

Interim analysis for the ORFAN Arm 2/3 outcome data will be done when we complete 500 patients with paired CT scans (baseline and after at least 3 years).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

9. DATA MANAGEMENT

9.1. Access to Data

Direct access will be granted to authorised representatives from the Sponsor or host institution for monitoring and/or audit of the study to ensure compliance with regulations.

9.2. Data Recording and Record Keeping

Study Arms 1, 2 and 3

An internal study ID number will identify participants, and will not contain personal data such as names or initials. The personal data that links a participant's study ID number with their personal information will be kept by the study investigator on each selected site.

Pseudonymised data from imaging studies will be stored in restricted access areas on each selected site as well as in the core laboratory at the University of Oxford, where access to the building is carefully controlled, with both swipe card and key code entry systems in place. Codes are only given to authorised study investigators or authorised permanent members of staff.

Study information will be stored on secure computers of the study investigators which are all individually password protected. Information stored on laptop computers of the study investigators will have encrypted hard disks and contain additional password protection pertaining to any relevant documents; this will be in addition to all computers being password protected. Data will not contain subject-identifiable material, and will be stored in a pseudonymised form.

With the consent of the participant, personal data (namely participant's NHS number, Date of Birth and post code) will be shared securely between the local UK site, the main study site at the University of Oxford and NHS digital/ONS/NICOR/SSNAP or any other NHS affiliated service or registry that collects patient data via secure email in order to request health outcome data from participants. Data will be transferred from each of the local UK sites to the main study site at the University of Oxford which will act as data controller and request such information from NHS digital/ONS/NICOR/SSNAP or any other NHS affiliated service or registry that collects patient data. NHS digital/ONS/NICOR/SSNAP or any other NHS affiliated service or registry that collects patient data will return the requested information to the data controller at the main site in pseudonymised form which is only identifiable to the data controller.

Additionally, pseudonymised patient data will be recorded in eCRF, encrypted and stored in Castor, a cloud-based platform (https://www.castoredc.com). All sites will be able to access only their own cohort of patients, whereas the main study site at the University of Oxford will have a full access and control of all cohorts. One of the Castor's servers is located within the UK and this server will be used for the study. Castor EDC is compliant with the University of Oxford Information Security Policy and IT security requirements. Access to the patients' identifiable information will be given to authorised study investigators and the local NIHR coordinator (for those patients who agreed to be approached for participation in future ethically approved studies).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Study Arm 4

Patients (including those with COVID-19) who have had CTA or CT chest, abdomen and pelvis examinations as part of their routine clinical care will be identified by the clinical care team (or researchers acting under arrangements with the responsible care organisation) at each recruiting site through a search of the radiology information system of participating study sites (UK and internationally). The clinical care team (or researchers acting under arrangements with the responsible care organisation) will assign unique participant study IDs to these patients and will keep these in a site-specific code file. The clinical care team (or researchers acting under arrangements with the responsible care organisation) will then proceed to pseudonymise the CT scans by assigning them this unique participant study ID and send them to the University of Oxford.

For the UK, the clinical care team (or researchers acting under arrangements with the responsible care organisation) will provide identifiable data alongside a unique participant study ID (code file) to an authorised third party with approval to hold linked patient data (e.g. NHS digital/ONS/NICOR/SSNAP or any other NHS affiliated service or registry that collects patient data) via secure email to facilitate linkage with relevant national databases. Third parties (e.g. NHS digital/ONS/NICOR/SSNAP or any other NHS affiliated service or registry that collects patient data) will send the data extraction with the unique participant study ID only to the University of Oxford to allow linkage of the imaging data with their data extract.

All study data will be entered on a password protected and encrypted database – Castor EDC. The participants will only be able to be identified by study specific participant number and/or code in any database. The name and any other identifying details will not be included in this study.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

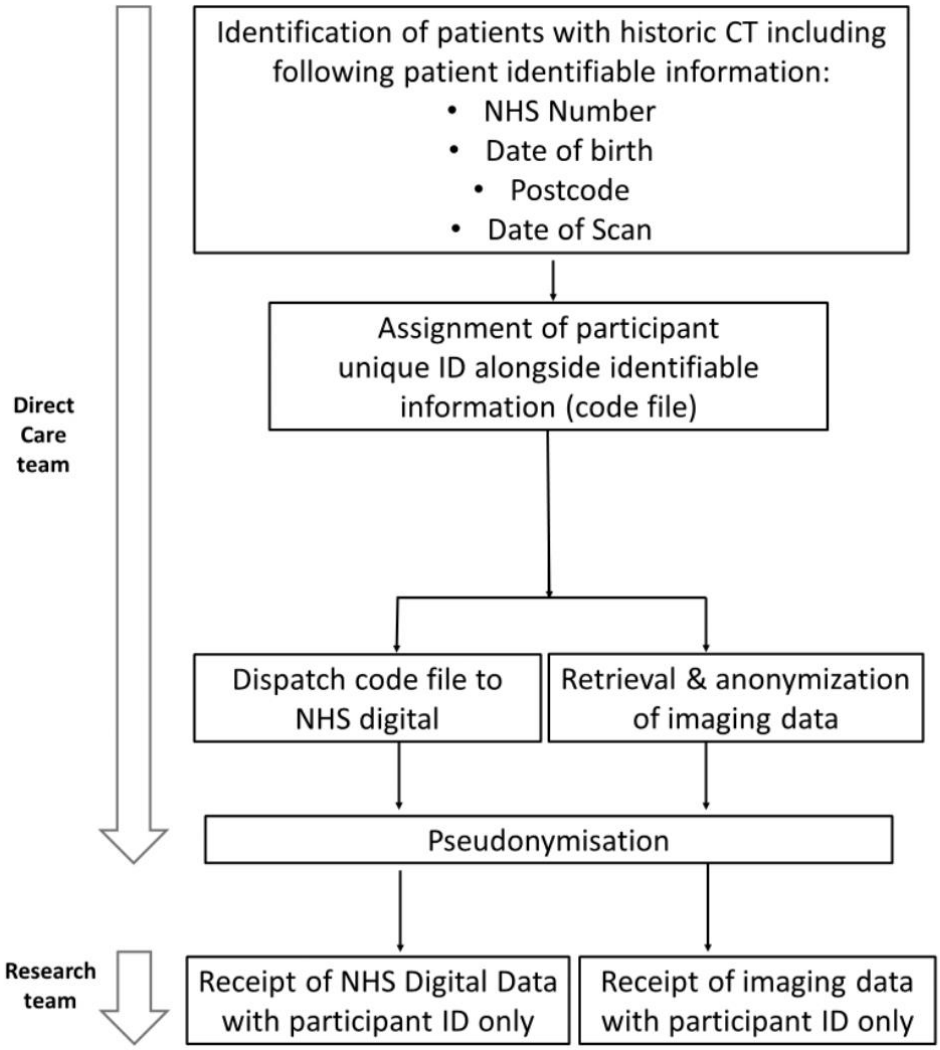


Figure C- Arm 4 Data flow for UK sites

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

9.3 Image Data Anonymisation

Image data from internal referrals will be transferred from the OUH NHS Trust PACS system via Clear Canvas and pseudonymised at the point of transfer. Pseudonymised scans from non-OUH sites will be transferred to the University of Oxford either on password-protected encrypted hard drives or via a HIPAA-compliant image transfer system (e.g. CIMAR gate run by Caristo Diagnostics, Ambra or equivalent; this is applicable to all study arms). Scans will then be uploaded to the University of Oxford server and will be scanned to remove any possible left-over series containing screen shots / captured images or patient identifiable information not removed by the anonymization tools.

For arm 4 specifically, image data from external sources will be scanned for patient identifiable information prior to upload onto the Oxford University server via a DICOM viewer. Any data containing patient identifiable information will be deleted immediately and the sender will be contacted.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

10. QUALITY ASSURANCE PROCEDURES

The study may be monitored, or audited in accordance with the current approved protocol, ICH GCP, relevant regulations and standard operating procedures.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

11. ETHICAL AND REGULATORY CONSIDERATIONS

11.1. Declaration of Helsinki

The Investigator will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki.

11.2. ICH Guidelines for Good Clinical Practice

The Investigator will ensure that this study is conducted in full conformity with relevant regulations and with the ICH Guidelines for Good Clinical Practice (CPMP/ICH/135/95) July 1996.

11.3. Section 251 Support

Section 251 refers to section 251 of the National Health Service Act 2006 and its current Regulations, the Health Service (Control of Patient Information) Regulations 2002. Section 251 support for Arm 4 UK patient records will be used to instigate flow of personal identifiable data from participating sites to third parties with approval to hold linked patient data (e.g. NHS digital/NICOR/SSNAP/ONS or any other NHS affiliated service or registry that collects patient data). The third parties will be supplied by the participating site clinical care team (or researchers acting under arrangements with the responsible care organisation) with the database containing site-specific participant study IDs alongside the required identifiable information to perform data linkage (NHS Number, Hospital ID Number, Date of Birth, Gender, Address (Postcode Level).

Once third parties have performed the data linkage between their databases and the study site supplied database, patient identifiable data are deleted. This will create one database containing the third party extracted variables alongside the participant study ID. This database can then be sent to the study core lab, via secure communication means. The third party that performed the data linkage (e.g. NHS digital/NICOR/SSNAP/ONS or any other NHS affiliated service or registry that collects patient data) will keep the linkage file to update the relevant outcomes in the future.

With the above processes, the study investigators who are not members of the clinical care team (or researchers acting under arrangements with the responsible care organisation) will not be given access to patient's identifiable information at any stage. The sites (local clinical care team or local researchers acting under arrangements with the responsible care organisation) will destroy the linkage file after 10 years.

11.4. Approvals

The protocol, informed consent form, participant information sheet and any proposed advertising material will be submitted to an appropriate Research Ethics Committee (REC), and host institution(s) for written approval. The Investigator will submit and, where necessary, obtain approval from the above parties for all substantial amendments to the original approved documents.

11.5. Reporting

The CI shall submit once a year throughout the study, or on request, an Annual Progress report to the REC Committee, host organisation and Sponsor. In addition, an End of Study notification and final report will be submitted to the same parties.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

11.6. Participant Confidentiality

The study staff will ensure that the participants' anonymity is maintained. The processing of the personal data of participants will be minimised by making use of a unique participant study number only on all study documents and any electronic database(s), with the exception of the CRF or eCRF, where participant initials may be added. All documents will be stored securely and only accessible by study staff and authorised personnel. The study will comply with the General Data Protection Regulation (GDPR) for European sites and associated Data Protection Act 2018, which requires data to be pseudonymised as soon as it is practical to do so. For US sites, the study will comply with the Health Insurance Portability and Accountability Act (HIPAA).

Participants will have the right to withdraw at any point of the study.

All information collected for the purposes of the study, will be information already filed in the medical records of the participants reported, and approved by specialist physicians and the responsible clinical care team. Therefore, reporting of incidental findings is not applicable.

11.7. Approval for Sending Samples to External Specialised Centres for Processing

To better characterise the samples collected in the context of this study, the expertise of scientific collaborators from other centres in the UK or abroad may be sought. Samples and imaging data collected for research purposes may be sent to other centres for specific analyses (e.g. NCIMI centres), according to the availability of specialist techniques that are not available to the local research team. Moreover, pseudonymised data and/or biological samples may be sent to other centres participating in the National Institute for Health Research (NIHR) bioresource for further processing. As all stored samples/data are identified by study code only, no identifying information will leave the site ensuring that no confidentiality issues will arise. Study investigators and other authorised investigators (including the local NIHR coordinator) will have access to the participants' identifiable information, but any contact with study participants for future ethically approved studies will require the prior approval of the study's Principal Investigator, and will be carried out by authorised investigators.

Participants will be informed in the PIL and sign in the consent form (if they approve) that their pseudonymised samples may be sent to external centres for specialist processing. It will be clearly stated that no identifying information will be sent with these samples.

11.8. Expenses and Benefits

Reasonable travel expenses for those consented to study arm 2 and 3a will be reimbursed on production of receipts.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

12. FINANCE AND INSURANCE

12.1. Funding

This study is funded by the Oxford Biomedical Research Centre, The British Heart Foundation, Innovate UK, National Consortium of Intelligent Medical Imaging and the European Commission, National Institute for Health Research (NIHR), and NHSX.

12.2. Insurance

The University has a specialist insurance policy in place which would operate in the event of any participant suffering harm as a result of their involvement in the research (Newline Underwriting Management Ltd, at Lloyd's of London, policy numbered: WD1200463). NHS indemnity operates in respect of the clinical treatment which is provided.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

13. PUBLICATION POLICY

The Investigators will be involved in reviewing drafts of the manuscripts, abstracts, press releases and any other publications arising from the study. Authors will acknowledge that the study was funded by Oxford Biomedical Research Centre, British Heart Foundation, Innovate UK, the European Commission and National Institute for Health Research (NIHR), and NHSX. Authorship will be determined in accordance with the ICMJE guidelines and other contributors will be acknowledged. Participants can access any published findings created from this study by contacting the Principal Investigator.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

14. REFERENCES

1. Frayn KN, Karpe F, Fielding BA, Macdonald IA, Coppack SW. Integrative physiology of human adipose tissue. Int J Obes Relat Metab Disord 2003;27:875-88.
2. Antoniades C, Antonopoulos AS, Tousoulis D, Stefanadis C. Adiponectin: from obesity to cardiovascular disease. Obesity reviews : an official journal of the International Association for the Study of Obesity 2009;10:269-79.
3. Antonopoulos AS, Margaritis M, Coutinho P et al. Reciprocal effects of systemic inflammation and brain natriuretic peptide on adiponectin biosynthesis in adipose tissue of patients with ischemic heart disease. Arteriosclerosis, thrombosis, and vascular biology 2014;34:2151-9.
4. Margaritis M, Antonopoulos AS, Digby J et al. Interactions between vascular wall and perivascular adipose tissue reveal novel roles for adiponectin in the regulation of endothelial nitric oxide synthase function in human vessels. Circulation 2013;127:2209-21.
5. Marinou K, Hodson L, Vasan SK et al. Structural and functional properties of deep abdominal subcutaneous adipose tissue explain its association with insulin resistance and cardiovascular risk in men. Diabetes Care 2014;37:821-9.
6. Iacobellis G. Relation of epicardial fat thickness to right ventricular cavity size in obese subjects. The American journal of cardiology 2009;104:1601-2.
7. Ito T, Nasu K, Terashima M et al. The impact of epicardial fat volume on coronary plaque vulnerability: insight from optical coherence tomography analysis. European heart journal cardiovascular Imaging 2012;13:408-15.
8. Mahabadi AA, Berg MH, Lehmann N et al. Association of epicardial fat with cardiovascular risk factors and incident myocardial infarction in the general population: the Heinz Nixdorf Recall Study. Journal of the American College of Cardiology 2013;61:1388-95.
9. Nakazato R, Dey D, Cheng VY et al. Epicardial fat volume and concurrent presence of both myocardial ischemia and obstructive coronary artery disease. Atherosclerosis 2012;221:422-6.
10. Fujimoto WY, Bergstrom RW, Boyko EJ et al. Visceral adiposity and incident coronary heart disease in Japanese-American men. The 10-year follow-up results of the Seattle Japanese-American Community Diabetes Study. Diabetes care 1999;22:1808-12.
11. Kamimura MA, Carrero JJ, Canziani ME, Watanabe R, Lemos MM, Cuppari L. Visceral obesity assessed by computed tomography predicts cardiovascular events in chronic kidney disease patients. Nutrition, metabolism, and cardiovascular diseases : NMCD 2013;23:891-7.
12. Nakanishi R, Rajani R, Cheng VY et al. Increase in epicardial fat volume is associated with greater coronary artery calcification progression in subjects at intermediate risk by coronary calcium score: a serial study using non-contrast cardiac CT. Atherosclerosis 2011;218:363-8.
13. Ouchi N, Higuchi A, Ohashi K et al. Sfrp5 is an anti-inflammatory adipokine that modulates metabolic dysfunction in obesity. Science 2010;329:454-7.
14. Murphy RA, Register TC, Shively CA et al. Adipose tissue density, a novel biomarker predicting mortality risk in older adults. The journals of gerontology Series A, Biological sciences and medical sciences 2014;69:109-17.
15. Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LF. The trinity of COVID-19: immunity, inflammation and intervention. Nature Reviews Immunology 2020:1-12.
16. Center JHCR. COVID-19 Map. 2020.
17. Guzik TJ, Mohiddin SA, Dimarco A et al. COVID-19 and the cardiovascular system: implications for risk assessment, diagnosis, and treatment options. Cardiovasc Res 2020.
18. Kreutz R, Algharably EAE, Azizi M et al. Hypertension, the renin-angiotensin system, and the risk of lower respiratory tract infections and lung injury: implications for COVID-19. Cardiovasc Res 2020.
19. Adao R, Guzik TJ. Inside the heart of COVID-19. Cardiovasc Res 2020;116:e59-e61.
20. Oikonomou EK, Marwan M, Desai MY et al. Non-invasive detection of coronary inflammation using computed tomography and prediction of residual cardiovascular risk (the CRISP CT study): a post-hoc analysis of prospective outcome data. Lancet (London, England) 2018;392:929-939.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

21. Oikonomou EK, Williams MC, Kotanidis CP et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. Eur Heart J 2019.
22. Smeeth L, Thomas SL, Hall AJ, Hubbard R, Farrington P, Vallance P. Risk of myocardial infarction and stroke after acute infection or vaccination. N Engl J Med 2004;351:2611-8.
23. ONS. Coronavirus (COVID-19) latest insights. ONS, 2021.
24. British Heart Foundation. UK Factsheet. 2021.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

15. APPENDIX A: TELEPHONE QUESTIONNAIRE

Telephone Questionnaire (to be answered by participant, filled in by study investigator)	
Participant ID	
Study Year (please circle)	2 4 6 8 10
Date of assessment	
Treatment for high blood pressure	☐ Yes home BP: ☐ No
Breathlessness symptoms by NYHA Class (I-V)	
Chest pain symptoms by CCS Class (1-4)	
Smoking Status (0=never 1=ex-smoker 2=active)	
Death (Received from the ONS prior to telephone call)	☐ Yes Date: ☐ No Cause of death (free text):
Myocardial Infarction Have you had a heart attack in the last year?	☐ Yes Date: ☐ No Details (free text):
Cerebrovascular Accident Have you had a stroke or a minor stroke (TIA) in the last year?	☐ Yes Date: ☐ No Details (free text):
Revascularisation Have you had heart surgery, stenting or coronary care in the last year?	☐ Yes Date: ☐ No Surgery Coronary stenting (PCI) Details (free text):
Other serious adverse event	☐ Yes Date: ☐ No Details (free text):
Number of admissions to hospital	Details (free text):
Reason for admission	Details (free text):
Total number of nights in hospital	

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Medications			**Exact medication, dose and starting date (free text)**
Aspirin	☐ Yes	☐ No	
Clopidogrel	☐ Yes	☐ No	
Warfarin	☐ Yes	☐ No	
Heparin	☐ Yes	☐ No	
Beta Blocker	☐ Yes	☐ No	
Calcium Channel Blocker ('pine')	☐ Yes	☐ No	
Nitrate	☐ Yes	☐ No	
Statin	☐ Yes	☐ No	
Ace Inhibitor ('pril')	☐ Yes	☐ No	
Angiotensin Receptor Blocker ('sartan')	☐ Yes	☐ No	
Diuretic	☐ Yes	☐ No	
Digoxin	☐ Yes	☐ No	
Amiodarone	☐ Yes	☐ No	
Insulin	☐ Yes	☐ No	
Oral Hyperglycaemic	☐ Yes	☐ No	
Other cardiovascular medicine	☐ Yes	☐ No	
Other medication	☐ Yes	☐ No	

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

AMENDMENT HISTORY

Amendment No.	Protocol Version No.	Date issued	Author(s) of changes	Details of Changes made
1	V3.0	Jan 2017	Sheena Thomas	Addition of Web Page Addition of Study Investigator Addition of Thank You letter Changes in the PIL, Protocol, Consent Form1 and Invitation Letters.
2	V4.0	October 2017	Sheena Thomas	Addition of Study Investigators. Change in the duration of the study; change in the age range for recruitment; change in the volume of blood being donated
3	V5.0	October 2018	Sheena Thomas	Addition of Study Investigators. Addition of multiple external sites to the study. Increase in sample size. Change in radiation dose. Inclusion of additional funding body and the involvement of a commercial partner and consequent revision to the conflict of interest statement.
4	V6.0	June 2019	Sheena Thomas	Addition of Study Arm 4 Addition of sites and study investigators to the study. Increase in sample size. Inclusion of additional funding body.
5	V7.0	November 2019	Maria Lyasheva Sheena Thomas	Addition of study investigators in Protocol. Addition of Information and Statistics Division of the NHS in Scotland to NHS Digital in PIL (External), Protocol and Consent form 1 OUH. Inclusion of PASS in PIL (External).

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

				Use of eCRF (Castor) in Protocol, PILs (OUH and External) and GDPR Privacy Notice. Addition of the Health Insurance Portability and Accountability Act (HIPAA) in Protocol. Changed the time period between the baseline CTA and the follow up CTA in Protocol, PILs (OUH and External), Study Arm 3 Letter of invitation to participate – OUH and Study Arm 3 Letter of invitation to participate – External Sites. Addition of wording to allow for blood to be drawn from a peripheral vein in PIL (External). Change of the location where the study is conducted in Study Arm 2 Letter of invitation to participate – OUH and Study Arm 3 Letter of invitation to participate – OUH. Addition of the website to the header in Study Arm 3 Letter of invitation to participate – OUH. addition of the wording 'for sites within England and Wales or the Information and Statistics Division of the NHS in Scotland' to clause 4 in Consent Form 1 OUH. Change of wording: removed the term "UK" in Protocol; blood amount taken from the participants is 'up to' 30 ml in PILs (OUH and External); added "/cardiology" to "the various radiology department" in Protocol.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

6	V 8.0		Rafail A Kotronias Maria Lyasheva Sheena Thomas	Addition of COVID-19 infection status. Addition of a new COVID objective and a measure. Addition of study investigators in Protocol. Splitting of Arm 3 to 3a and 3b. Previous Arm 3 becomes 3a with 3b representing an arm without prospective imaging and blood sampling. Update of Consent Form 2. New invitation letter for 3b. New PIL "follow-up PIL" for 3b (including external sites). Addition of third parties used for data collection. Change of arm 4 from anonymised retrospective data to pseudonymised data and outcomes, with Section 251 support. Update of patient facing documents to conform with I-GARD suggestions.
7	V9.0		Henry West Maria Lyasheva	Update of Arm 4 participant numbers with additional scientific justification and power calculation. Amendment of the medical physics section (Section D - Clinical Assessment: Q D1) of the IRAS NHS REC form following the risk re-assessment.
8	V10.0		Maria Lyasheva Cheng Xie	Addition of the post-mortem CT to the Study Arm 4. Addition of the new scanner to the protocol and the PIL. Addition of a study investigator to the Protocol.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

				Addition of a new participating site.

© Copyright: The University of Oxford and Oxford University Hospitals NHS Foundation Trust 2017

Appendix 5. ORFAN Arm 4 Study variables

Local ORFAN Arm 4 site case record file (database)

Column Code	Column Name	Definition/Explanation	Data Type	Importance (1-4) 1 = critical
Record_ID	**Study ID**	The Study ID in the following format: <ORFAN Study abbreviation [F]> <Arm Name [4]> <Abbreviation of your Organisation [LP]><Participant Number> For instance, F4LP00001 for the Liverpool Heart and Chest Hospital NHS Foundation Trust.	String [F4LP00001]	1
CT_Type_Arm4	**Scan type**	To determine what type of chest CT scan the patient underwent. Note, regular ORFAN Arm 4 is for patients who undergo CTCA. Arm 4 Covid arm is any chest imaging.	Number [1 - 6] where 1 = CT Coronary Angiogram (CTCA), 2 = CT Pulmonary Angiogram (CTPA), 3 = non-contrast cardiac CT, 4 = CT chest with contrast, 5 = CT chest without contrast, 6 = other type of chest scan, NA = unknown	1
Date_CT_scan_Arm4	**Date of scan**	To determine when the scan has been done (can be extracted from CRIS)	Number [DD/MM/YYYY], allow more than 1 date to be entered.	1
Comments_about_scan_Arm4	**Scan request**	To determine the CT scan request history, that is, the reasons for the CT scan (can be extracted from CRIS) , i.e. original scan is due to XX; the scan request as written by the requesting doctor	String [Free text]	1
CT_Report_Arm4	**Clinical Reports**	Clinical reports accompanying the CT scan i.e. the scan report/findings. (Can be extracted from CRIS)	String [Free text]	1
Ca_score_non_contrast_scan_Arm4	**Ca score**	To determine Ca score from a non-contrast cardiac CT scan if performed	Number	3
Date_Ca_Score_non_contrast_scan_Arm4	**Date of Ca Score**	To determine the date of the Ca Score on a non-contrast caridac CT scan	DD/MM/YYYY	3
CLINICAL HISTORY				
Gender_Arm4	**Gender**	To determine participants' gender.	Number [0 / 1/ 2], where 0 = Female, 1 = Male, 2=Transgender	1
Ethnicity_Arm4	**Ethnicity**	To determine ethnicity (if known).	Number [1 / 2 / 3 / 4], where 1 = Caucasian, 2 = Black, 3 = Asian, 4 = Other	1
Ethnicity_Other_Arm4	**Ethnicity (Other)**	To determine ethnicity if Other.	String [Free text]	3
Age_Arm4	**Month and Year of Birth**	To determine patient age by month and year of birth. Do not collect day of birth.	MM/YYYY	1
Blood_pressure_Arm4	**Blood pressure**	To determine patient blood pressure at the time of the CT scan (ideally, as close to the CT scan as possible)	Number (systolic/diastolic in mmHg)	1
Heart_Rhythm_Arm4	**Heart Rhythm**	To determine heart rhythm of participants at the moment of CT scan, or as close to as possible.	Number [1 / 2 / 3 / 4], where 1 = Sinus Rhythm, 2 = Atrial Fibrillation, 3 = Permanent pacemakers, 4 = other arrhythmia, NA = unknown	1
Smoking_Arm4	**Smoking**	To determine whether the participant smoke or ever smoked	Number [0 / 1 / 2], where 0 = Non-smoker, 1 = Ex-smoker*, 2= Active Smoker**, NA = unknown * Ex-smoker if smoked up to 1 month ago ** Active Smoker if smoking to smoked in the last month	1
Family_History_Arm4	**Family History**	To determine whether any of participants' parents/siblings have/had any problems with the heart (if available).	Number [0 / 1] where 0 = No, 1 = Yes, NA = unknown	3
Family_History_Comments_Arm4	**Family History Comments**	To determine which heart problems participants' parents/siblings have/had.	String [Free text]	4
Height_Arm4	**Height**	To determine height of patient	Number [cm]	1
Weight_Arm4	**Weight**	To determine height of patient	Number [kg]	1
BMI_Arm4	**BMI**	To determine Body Mass Index (BMI) at time of CT scan. Can use formula if height and weight is complete, but if BMI alone is avaliable and not height and weight, manually input BMI	Number [kg/m2]	1

MEDICAL HISTORY - DIAGNOSES UP TO TIME OF CTCA - ICD-10 codes provided if used at your site				
AF_Arm4	**Atrial Fibrillation**	To determine whether participants have atrial fibrillation. ICD-10: I48.0. ICD-10: https://icd.who.int/browse10/2016/en#/I48.0 ICD-10: I48.1. ICD-10: https://icd.who.int/browse10/2016/en#/I48.1	Number [0 / 1 / 2 / 3 / 4], where 0 = No, 1 = Paroxysmal atrial fibrillation, 2 = persistent AF, 3 = Chronic AF, 4 = AF of unknown type.	1
Hypertension_Arm4	**Hypertension**	To determine whether participants have hypertension.	Number [0 / 1], where 0 = No, 1 = Yes, NA = unknown	1
Hypercholesterolemia_Arm4	**Hypercholesterolaemia**	To determine whether participants have hypercholesterolemia. ICD-10: E78.0 Pure hypercholesterolaemia ICD-10: https://icd.who.int/browse10/2016/en#/E78.0	Number [0 / 1], where 0 = No, 1 = Yes, NA = unknown	1
Diabetes_Arm4	**Diabetes mellitus**	To determine whether participants have diabetes mellitus. ICD-10: E10-E14 ICD-10: https://icd.who.int/browse10/2016/en#/E10-E14 E10: Type 1 diabetes mellitus E11: Type 2 diabetes mellitus E12: Malnutrition-related diabetes mellitus E13: Other specified diabetes mellitus E14: Unspecified diabetes mellitus	Number [0 / 1 / 2/ 3], where 0 = No, 1 = Type 1, 2 = Type 2, 3 = other types of diabetes mellitus, NA = unknown	1
Congestive_Heart_Failure_Arm4	**Congestive Heart Failure**	To determine whether participants have had congestive heart failure. ICD-10: I42 and I50.0-I50.9 ICD-10: https://icd.who.int/browse10/2016/en#/I50 I50.0: Congestive heart failure I50.1: Left ventricular failure I50.9: Heart failure, unspecified	Number [0 / 1] where 0 = No, 1 = Yes, NA = unknown	2
Valvular_Heart_Disease_Arm4	**Valvular Heart Disease**	To determine whether participants have/had valvular heart disease. ICD-10: I34-I39 https://icd.who.int/browse10/2019/en#/I34 I34: Nonrheumatic mitral valve disorders I35: Nonrheumatic aortic valve disorders I36: Nonrheumatic tricuspid valve disorders I37: Pulmonary valve disorders I38: Endocarditis, valve unspecified I39: Endocarditis and heart valve disorders in diseases classified elsewhere	Number [0 / 1] where 0 = No, 1 = Yes, NA = unknown	2
Previous_Stroke_Arm4	**Previous Stroke**	To determine whether participants have had a haemorrhagic and/or ischaemic stroke prior to the CTCA https://icd.who.int/browse10/2019/en#/, I61, I62 I61: Intracerebral haemorrhage I62: Other nontraumatic intracranial haemorrhage https://icd.who.int/browse10/2019/en#/I63	Number [0 / 1 / 2] where 0 = No, 1 = Ischaemic stroke, 2 = haemorrhagic stroke, NA = unknown	1
Previous_ACS_Arm4	**Previous ACS**	To determine whether participants have suffered from Unstable Angina (chest pain of cardiac origin at rest), NSTEMI and/or STEMI prior to the CTCA If multiple, select most severe. ICD-10: https://icd.who.int/browse10/2016/en#/I20-I25	Number [0 / 1 / 2 / 3] where 0 = No, 1 = Unstable angina, 2 = NSTEMI, 3 = STEMI, NA = unknown	1

MEDICATIONS - AT TIME OF SCAN OR FROM WITHIN 6 MONTHS OF THE CT SCAN				
Medications_Arm4	**Medications**	To determine what medications the particioant was on at the time of the scan.	String [Free text]. List all medications seperated by comma. Scientific names preferred over brand names.	1
PATHOLOGY INVESTIGATIONS - 6 MONTHS PRIOR TO OR POST THE CT SCAN (can be from multiple blood draws)				
Date_of_blood_test_Arm4	**Date of blood test**	To determine the date of a recent blood test (up to 3 months prior- / post CT).	DD/MM/YYYY	1
Total_Cholesterol_Arm4	**Total Cholesterol**	To determine total Cholesterol level.	Number [mmol / L]	2
HDL_C_Arm4	**HDL-C**	To determine HDL-C level.	Number [mmol / L]	1
LDL_C_Arm4	**LDL-C**	To determine LDL-C level.	Number [mmol / L]	1
Triglycerides_Arm4	**Triglycerides**	To determine Triglycerides level.	Number [mmol / L]	1
C_reactive_protein_Arm4	**C Reactive Protein**	To determine C Reactive Protein level.	Number [mg / L]	2
Troponin_I_Arm4	**Troponin**	To determine Troponin level.	Number [ng / L]	3
Haemoglobin_A1c_DCCT_ aligned_blood_Arm4	**HbA1C**	To determine HbA1C level.	Number [%]	2
BNP_Arm4	**BNP**	To determine B-type natriuretic peptide (BNP) level.	Number [pmol / L]	2
Glucose_level_plasma_Arm4	**Glucose (Plasma)**	To determine Glucose level in plasma (formal laboratory level)	Number [mmol / L]	2

NHS Digital data variables

<u>NHS Digital data for linkage with ORFAN Arm 4</u>

The data from the datasets held and managed via NHS Digital that is to be linked to patient identifiers submitted by clinical care teams on behalf of the ORFAN Study investigators. The identifiers will not be shared with the ORFAN study team.

There are multiple datasets managed by NHS Digital for which we intend to link with. These are listed individually in tables that provide a brief introduction to the dataset and state the data controller and total number of variables requested from each dataset. We list each individual variable we wish to link with, and provide a comment as to its justification.

We request data linkage with the following datasets (each set out in subsequent pages)

- Hospital Episode Statistics (HES) – Accident and Emergency dataset
- HES – Admitted patient care dataset
- HES - Outpatients
- HES – Critical care
- Emergency Care Data Set (ECDS)
- Medicines dispensed in Primary Care (NHSBSA data)
- Civil Registration (Deaths) - Secondary Care Cut
- COVID-19 Hospitalization in England Surveillance System
- HES:Civil Registration (Deaths) bridge

Emergency Care Dataset (ECDA)	
Data controller: NHS Digital	
Introduction: The ECDS is the national dataset for urgent and emergency care. Linkage to this dataset will enable the inclusion of critical diagnostic and treatment data in relation to cardiovascular disease outcomes that are managed in the emergency setting such as heart attacks and strokes.	
Number of variables requested: We request 18 variables for linkage. No high risk variables or identifiable variables are requested.	
Periods: October 2017 to March 2018 Final Data 2018/19 2019/20 Annual data releases Latest available data	
Field code and Field name ***(Same in the case of this database)***	**Comment in regards to ORFAN Study**
ARRIVAL_DATE	To know when in relation to the CT scan of relevance the episode of care occurred
ARRIVAL_MODE	To know how the patient arrived at the emergency department to inform analysis of severity
ARRIVAL_TIME	To know when the patient arrived for the care episode, so timing of investigations such as CT scans can be correctly ordered for the study
ASSESSMENT_DATE	To adjust analysis for the date of assessment so that timing of investigations such as CT scans can be correctly ordered for the study
ASSESSMENT_TIME	To adjust analysis for the time of assessment so that timing of investigations such as CT scans can be correctly ordered for the study
ATTENDANCE_CATEGORY	To know the category of attendance type for adjustment to limit attendances to appropriate emergency care types
ATTENDANCE_SOURCE	To adjust for where the patient presented from, an important variable for consideration of patient disease risk
CHIEF_COMPLAINT	To know what the patient's own chief complaint was, such as chest pain, which is an important variable for adjustment for analysis focused on symptoms
COMORBIDITIES	To know which cardiovascular diseas risk factors the patient had for adjustment of analysis
CONCLUSION_DATE	To know when this episode of care ended so that timing of investigations such as CT scans can be correctly ordered for the study
CONCLUSION_TIME	To know when this episode of care ended so that timing of investigations such as CT scans can be correctly ordered for the study
DEPARTURE_DATE	To know when this episode of care ended so that timing of investigations such as CT scans can be correctly ordered for the study
DEPARTURE_TIME	To know when this episode of care ended so that timing of investigations such as CT scans can be correctly ordered for the study

DIAGNOSIS_CODE	Critical variable to know what health event occurred and what diagnosis was assigned to the patient for this attendance at the emergency department
DISCHARGE_DESTINATION	To know and to adjust for the discharge destination of the patient in analysis, a marker of disease severity
DISCHARGE_STATUS	To know the outcome of this attendance to the emergency department, a key variable for adjustment
INVESTIGATION_CODE	To know which investigations were performed during this episode of care, a key variable for analysus
TREATMENT_CODE	To know what treatments were performed during this episode of case, a key variable of analysis for adjusting severity of the condition

COVID-19 Hospitalization in England Surveillance System	
Data controller: NHS Digital	
Introduction: The COVID-19 Hospitalization in England Surveillance System is the national dataset for covid-19 hospitalisation. Linkage to this dataset will enable the inclusion of critical diagnostic and treatment data in relation to covid-19, critical data for the ORFAN Study to address scientific questions concerning the effect of covid-19 infection and disease severity on later cardiovascular disease outcomes such as heart attacks and strokes.	
Number of variables requested: We request 65 variables for linkage. No high risk variables or identifiable variables are requested.	
Periods: Annual data releases	
Field code/Field name	**Comment in regards to ORFAN Study**
AdmissionCOVID19/Was the admission Covid-19 related?	To adjudicate if the admission was related to covid-19, a key variable for selection of patients for analysis
AdmissionFlu/Was the admission flu related?	A key variable to limit analysis to the correct patients with covid-19, not other conditions
AdmissionRSV/Was the admission RSV related?	A key variable to limit analysis to the correct patients with covid-19, not other conditions
AdmittedFrom/Admitted From	To adjust for where ethe patient was admitted form, which influences analysis if the patient was admitted from another healthcare facility
AntiCovid19Treatment/Received specific anti-Covid-19 treatment	Important variable to adjust for those who specifically received covid-19 specific treatments, which may change later cardiovascular disease outcomes
AsthmaRequiring/Asthma (underlying condition)	To adjust analysis for those with asthma, a major risk factor for covid-19 severity and may influence other outcomes
AsymptomaticTesting/Asymptomatic Testing	To know if the patient was originally asymptomatic, a major variable for severity assessment
CauseOfDeath/Cause of Death	A critical variable to know how the patient died if they were covid-19 positive. This could include heart disease precipitated by infection.
ChronicHeart/Chronic heart disease (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes
ChronicLiver/Chronic liver disease (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes
ChronicNeurological/Chronic neurological disease (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes

ChronicRenal/Chronic renal disease (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes
ChronicRespiratory/Chronic respiratory disease (excluding asthma) (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes
Covid19/Result of laboratory test Covid-19?	A key variable to adjust for those patient who have laboratory confirmed infection beyond symptom confirmed infection
DateAdmittedICU/Date Admitted to ICU/HDU	A key variable to assess severity of infection due to intensive therapy for adjustment
DateLeavingICU Date Leaving ICU/HDU	A key variable to assess severity of infection due to intensive therapy for adjustment
DateUpdated/Date Updated	An administrative variable for validation of the data and when it was updated
DiabetesType/Diabetes Type	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes
EstimatedDateOnset/Estimated date of onset of symptoms	A key variable to know how long the patient may have had covid-19 beyond hospital admission data
FinalOutcome/Final Outcome	A key variable to adjust analysis for outcomes from covid-19
FinalOutcomeDate/Final Outcome Date	A key variable to know when the final outcome, such as discharge or death, was met, so that other relevant analysis can be ceased
Highflownasaloxygen/Patient required high flow nasal oxygen?	A key variable for disease severity adjustment
HospitalAdmissionDate/Date of admission to hospital	Critical variable to know temporality between admission and investigations that may have occurred during the admission
HospitalAdmissionHours/HospitalAdmissionHours	A key variable to assess severity of infection due to inpatient therapy required for adjustment
HospitalAdmissionMinutes/HospitalAdmissionMinutes	A key variable to assess severity of infection due to inpatient therapy required for adjustment
HoursAdmittedICU/HoursAdmittedICU	A key variable to assess severity of infection due to intensive therapy for adjustment
Hypertension/Hypertension (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular

	disease outcomes
ImmunosuppressionDisease/Immunosuppression (disease related) (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes
ImmunosuppressionTreatment/Immunosuppression (treatment related) (underlying illness)	A critical variable to adjust analysis for patients with underlying disease states that can influence later cardiovascular disease outcomes
InfectionSwabDate Swab/Specimen Date	Key variable to know when the test swab was performed to inform analysis of disease severity
InfluenzaAH1N1pdm2009/Result of laboratory test A/H1N1pdm2009?	To adjust analysis for the presence of other respiratory pathogens that confound covid-19 severity
InfluenzaAH3N2/Result of laboratory test A/H3N2?	To adjust analysis for the presence of other respiratory pathogens that confound covid-19 severity
InfluenzaAnonsubtyped/Result of laboratory test A/non-subtyped?	To adjust analysis for the presence of other respiratory pathogens that confound covid-19 severity
InfluenzaAunsubtypable/Result of laboratory test A/unsubtypable?	To adjust analysis for the presence of other respiratory pathogens that confound covid-19 severity
InfluenzaB/Result of laboratory test InfluenzaB?	To adjust analysis for the presence of other respiratory pathogens that confound covid-19 severity
Invasivemechanicalventilation/Patient required invasive mechanical ventilation respiratory support?	A key variable for adjustment, as such treatment and complications need to be taken into account when undertaking analysis of disease severity
IsARDSComplication/ARDS complication during admission	A key variable for adjustment, as such treatment and complications need to be considered when undertaking analysis of disease severity
IsDiabetes/Diabetes (underlying illness)	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
IsOtherCoInfectionsComplication/Other co-infections complication during admission	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
IsOtherComplication/Other complication during admission	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
IsSecondaryBacterialPneumoniaComplication/Secondary	A key variable for adjustment, as such

bacterial pneumonia complication during admission	treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
IsUnknownComplication/Unknown complication during admission	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
IsViralPneumoniaComplication/Viral pneumonia complication during admission	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
LabTestDate/Laboratory Test Date	
Mechanicalinvasiveventilationduration /Mechanical invasive ventilation duration	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
MinutesAdmittedICU/Minutes Admitted ICU	
NoninvasiveMechanicalventilation/Patient required non-invasive ventilation respiratory support?	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
ObesityBMI/Obesity (BMI)	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
Obesity/Clinical Obesity (Clinically apparent)	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
OrganismName/Organism Name	A key variable for adjustment, as such complications and their virulence need to be considered when undertaking analysis of disease severity
Other/Underlying condition - other	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
OtherResult/Other Result	A key variable for adjustment, as such treatment and complications and underlying conditions need to be considered when undertaking analysis of disease severity
Oxygenviacannulaeormask/Patient required Oxygen via cannulae or mask	A key variable for adjustment, as such treatment and complications and

	underlying conditions need to be considered when undertaking analysis of disease severity
PatientStillOnICU/Is patient still on ICU/HDU?	To adjust analysis for unresolved clinical episodes at the time of data extraction, to limit false analysis without outcomes.
RespiratorySupportECMO/Patient required ECMO respiratory support?	A key variable for adjustment, as such advanced treatment need to be considered when undertaking analysis of disease severity
RespiratorySupportNone/Patient did not require additional respiratory support	A key variable for adjustment, as such treatment needs to be considered when undertaking analysis of disease severity
RespiratorySupportUnknown/Unknown respiratory support	A key variable for adjustment, as such treatment needs to be considered when undertaking analysis of disease severity
RSV/Result of laboratory test RSV?	To adjust analysis for the presence of other respiratory pathogens that confound covid-19 severity
SBDate/Secondary Bacterial - date	To adjust analysis for the presence of other respiratory pathogens that confound covid-19 severity and the date of which the secondary infection became know
TransferDestination/If transferred destination	To adjust for the destination of transfer following admission, to adjust for disease impact and effect of disease on function
Typeofspecimen/Type of Specimen	To adjust for the nature of the specimen sample used for covid-19 testing, as not all samples are as accurate
WasthepatientadmittedtoICU/Was the patient admitted to ICU/HDU?	To adjudicate severity of infection based upon a key indicator of severity: ICU admission
WeekNo/WeekNo	To adjust analysis for appropriate time frames and to know when admissions occurred in respect to other events such as CT imaging
WeekofAdmission/WeekofAdmission	To adjust analysis for appropriate time frames and to know when admissions occurred in respect to other events such as CT imaging
YearofAdmission/YearofAdmission	To adjust analysis for appropriate time frames and to know when admissions occurred in respect to other events such as CT imaging

Medicines dispensed in Primary Care (NHSBSA data)	
Data controller: NHS Digital	
Introduction: The NHSBSA is the national dataset for electronic and paper prescriptions submitted to the NHSBSA for reimbursement each month. It covers all prescriptions issued by all prescribers in general practice, community clinics, hospital clinics, dentists and community nursing. Linkage to this dataset will enable the inclusion of critical treatment data in relation to cardiovascular disease outcomes. For example, if a patient has a CT scan of relevance to the ORFAN Study in 2018, however they are commenced on a statin medication in 2019 by their general practitioner, this is critical information to adjust analysis for risk of death due to heart attack, as the statin will influence the degree of inflammation present around the coronary artery, and our analysis will otherwise over predict disease risk if we were not able to adjust for the commencement of the medication.	
Number of variables requested: We request 17 variables for linkage. No high risk variables or identifiable variables are requested.	
Periods: Annual data releases	
Field code/Field name	**Comment in regard to ORFAN Study**
DispensedPharmacyType/DISPENSER_TYPE	To know from where the medication was dispensed, such as hospital or community pharmacies for adjustment
PaidBNFCode/DISPENSED_BNF_CODE	To know the national formulary code for the medication, to ensure correct medications are identified
PaidBNFName/DISPENSED_BNF_NAME	To know the national formulary name for the medication, to ensure correct medications are identified
PaiddmdCode/DISPENSED_DMD	To know which dictionary of medicines and devices code the dispensed drug was
PaidDrugStrength/DISPENSED_STRENGTH	To know and adjust for what strength was supplied to the patient, a key variable
PaidFormulation/DISPENSED_FORMULATION	To know and adjust for what formulation was supplied to the patient
PaidQuantity/DISPENSED_QUANTITY	To adjust for the quantity of the medication supplied to know how long a course of therapy was initiated for
PaidSupplierName/DISPENSED_SUPPLIER_NAME	To know the supplier brand for medications where this can be relevant, such as biologically active anti-inflammatory medications
PatientAge/AGE	To adjust for patient age in years at the time of the script dispensing
PrescribedBNFCode/PRESCRIBED_BNF_CODE	To know the originally prescribed national formulary code for the medication
PrescribedBNFName/PRESCRIBED_BNF_NAME	To know the originally prescribed national formulary name for the prescribed medication
PrescribeddmdCode/PRESCRIBED_DMD	To know which dictionary of medicines and devices code the prescribed drug was
PrescribedFormulation/PRESCRIBED_FORMULATION	To know in which formulation the prescriber intended to the medication to be for adjustment of analysis
PrescribedMedicineStrength/PRESCRIBED_STRENGTH	To know in which strength the prescriber intended to the medication to be for

	adjustment of analysis
PrescribedQuantity/PRESCRIBED_QUANTITY	To know in what quantity the prescriber intended to the medication to be for adjustment of analysis
PrescribedSupplierName/PRESCRIBED_SUPPLIER_NAME	To know which supplier the prescriber intended the medication to be from, which can be relevant to analysis for biologically active therapies
PrescriberType/PRESCRIBER_TYPE	To know the type of prescriber to adjust for the possible therapeutic intentions of the dispensing
ProcessedPeriod/MONTH_OF_PROCESSING	Critical variable to know the date of the script being processed, to know temporality of the dispensing in relation to other events such as a CT scan of relevance to the study
ProcessingPeriodDate/DATE_OF_PROCESSING	Critical variable to know the date of the script being processed, to know temporality of the dispensing in relation to other events such as a CT scan of relevance to the study

Hospital Episode Statistics – Accident and Emergency dataset		
Data controller: NHS Digital		
Introduction: The HES Accident and Emergency dataset contains critical data for linking hospital events that occur at emergency departments to ORFAN relevant outcomes and events.		
Number of variables requested: We request 34 variables for linkage of the 160 variables within this dataset. We request no high risk or identifiable variables.		
Field	**Field name**	**Comment in regards to ORFAN Study**
Category: Attendances		
AEARRIVALMODE	Arrival mode	Critical variable for understanding and adjusting for patient urgency in relation to ambulance/walk-in arrival mode.
AEATTENDCAT	Attendance category	To adjust for initial or follow-up attendances for the same medical problems
AEATTENDDISP	Attendance disposal	Key variable to adjust for end of care, including death in department, onwards referral or discharged.
AEPATGROUP	Patient group	Key variable to group analysis by types of presentations, particularly those related to medical events.
AEREFSOURCE	Source of referral for A&E	A valuable variable to adjust analysis for the origin of the referral to the GP, key for analysis of evets that occur at other health care facilities
ARRIVALDATE	Arrival date	Key variable to adjust for the temporality of events in relation to study relevant scans or other events.
ARRIVALTIME	Arrival time	Key variable to adjust for the temporality of events in relation to study relevant scans or other events.
CONCLTIME	Conclusion time	Key variable to adjust for the temporality of events in relation to study relevant scans or other events.
DEPTIME	Departure time	Key variable to adjust for the temporality of events in relation to study relevant scans or other events.
TRETTIME	Time see for treatment	Key variable to adjust for the temporality of events in relation to study relevant scans or other events.
Category: Clinical diagnoses		
DIAG_NN	A&E Diagnosis	Critical variable to understand the health issues that led to the attendance at the emergency department.
DIAGSCHEME	Diagnosis scheme in use	Administrative data on the diagnostic classification utilised.
Category: Clinical investigations		
INVEST_NN	A&E Investigation	Key variable to understand the and adjust for key investigations ordered

		during this episode
Category: Clinical Treatments		
NOTREATS	Number of treatments	Key variable to adjust for severity of condition
TREAT_NN	A&E Treatment	Key variable to adjust for treatments that influence subsequent care and disposition for other ORFAN relevant clinical events
Category: General fields		
MATCH_RANK	MATCH Rank	Administrative matching
STUDY_ID	STUDY_ID	For linkage
Category: Healthcare resource groups (HRG) data		
DOMPROC	Dominant procedure	Key variable to adjust for the main procedure that occurred in this episode of care
Category: Patient data		
ARRIVALAGE	Age on arrival	Key code to know patient age in years on arrival to the department and adjust for age related risks
Category: System data		
AEKEY	Record identifier	To link with other HES datasets

Hospital Episode Statistics - Admitted Patient Care	
Data controller: NHS Digital	
Introduction: The HES admitted patient care database is a critical linkage database for ORFAN. Linkage to this dataset will enable the inclusion of critical, diagnostic data and treatment data in relation to myriad relevant conditions that influence heart health.	
Number of variables requested: We request 37 variables for linkage, with no high risk or identifiable fields requested	
Periods: 2005/06 2006/07 2007/08 2008/09 2009/10 2010/11 2011/12 2012/13 2013/14 2014/15 2015/16 2016/17 2017/18 2018/19 2019/20 Annual data releases Latest available	
Field code/Field name	**Comment regarding ORFAN Study**
ADMIDATE/Date of admission	Critical variable to adjust for when hospital admissions occurred in regard to other events such as study relevant CT scans
ADMIMETH/Method of admission	Key variable to know how a patient is admitted to hospital, as in a booked/planned/emergency/transfer etc, a key variable to understand the nature of the admission
ADMISORC/Source of admission	To adjust for where the patient was immediately prior to admission, a key variable for adjustment for certain conditions where transfer may have occurred from other health facilities
AEKEY/Record identifier	To facilitate linkage between A&E datasets acquired from NHS Digital
BEDYEAR/Bed days within the year	To adjust for the duration of a care episode in days, a key adjustment for severity
CAUSE/Cause code	To adjust for external causes of injury, to exclude admissions for non-relevant diagnoses
CLASSPAT/Patient classification	A key variable to identify the type of admission (day case/ordinary/night etc) for adjustment to only relevant admissions

DIAG_NN/All Diagnosis codes	A critical variable to know which conditions the patient was admitted for this episode.
DISDATE/Date of discharge	To adjust for the conclusion of an episode of care in relation to possible study relevant CT scans
DISDEST/Destination on discharge	To adjust for patients transferred to other health care settings at the conclusion of this episode of care
DISMETH/Method of discharge	Critical variable to adjust for discharge types such as death, or with medical advice.
DISREADYDATE/Discharge ready date	A more accurate variable for adjudicating times of when an episode of care has concluded from a health perspective, but when discharge is delayed.
DOMPROC/Trust derived dominant procedure	A key variable for the main procedure undertaken at this admission, a key variable for procedures such as elective PCI.
ELECDATE/Date of decision to admit	A variable for adjustment of care from when a consultant agreed to admit the patient, commencing care.
EPIDUR/Episode duration	A key adjustment for episode duration, a marker of severity
EPIEND/Date episode ended	A key variable of temporality for study relevant CT scans
EPIKEY/Record identifier	The HES record key for linkage
EPIORDER/Episode order	The number of the episode within the current spell, important for ensure proper temporality of episodes and spells in relation to other study relevant events such as CT scans
EPISTART/Date episode started	A key variable to know the commencement of the relevant episode for adjudicating temporality to study relevant CT scans
EPISTAT/Episode status	To ensure episode of care that are ongoing at the time of the HES reporting year are correctly managed, and data included in analysis are accurate.
EPITYPE/Episode type	A key code to group episodes of care into relevant bundles such as general episodes or maternity episodes.
FIRSTREG/First regular day or night admission	An important variable to indicates whether the episode falls within a sequence of regular day and night admissions and, if so, whether it is the first or subsequent episode within the sequence. Important for proper accuracy of data.
INTMANIG/Intended management	The variable to define what the planned admission was meant to be in regards to length of stay, important to understand when complications have occurred and length of stay has been extended.
MAINSPEF/Main specialty	Key variable to limit data to appropriate

	specialities such as cardiology, a very useful field for analysis
MATCH_RANK/MATCH_RANK	Administrative matching ID
OPDATE_NN/Date of operation	Key variable to know when operations have occurred in regards to other study relevant occurrences
OPERSTAT/Operation status code	To adjust analysis for patients with operations during the admission
OPERTN_NN/Primary Operative Procedure Codes	To adjust for specific procedures that occurred during the admission, an important variable for analysis
OPDATE_NN/Date of procedure	To know the date of procedures that occured during the admission, a key variable for assessment of temporality
POSOPDUR/Post-operative duration	An important variable that defines the difference in days between the date of the main procedure and the date the episode ended, a marker of severity.
PREOPDUR/Pre-operative duration	An important variable that defines the difference in days between the date of the start of the admission and the main procedure, a marker of complexity and possible complication.
SPELBGIN/Beginning of spell	A key variable to define if an episode is the first in a series of admissions (a spell), critical for correct understanding of the admissions in the right order.
SPELDU/Duration of spell	A key variable to define the period of a spell, critical for correct understanding of the admissions in the right order.
SPELEND/End of spell	A key variable to understand when a spell ends and what the concluding episode is.
STARTAGE/Age at start of episode	A key variable for the patients age in years for the start of the episode, a critical variable for analysis.
STUDY_ID/STUDY_ID	For linkage
SUSRECID/SUS record id	Secondary uses services record identifier for administration
TRETSPEF/Treatment specialty	A critical variable to define the nature of care, for analysis adjustment by specialty type.

Hospital Episode Statistics - Critical Care	
Data controller: NHS Digital	
Introduction: The HES Critical Care database contains all information related to the occurrence and duration of admissions to critical care facilities. Linkage to this dataset will enable the inclusion of critical treatment data for ORFAN patients who suffer severe disease events such as heart attacks and strokes necessitating life support.	
Number of variables requested: We request 24 variables for linkage, with no high risk or identifiable fields requested	
Periods: 2008/09 2009/10 2010/11 2011/12 2012/13 2013/14 2014/15 2015/16 2016/17 2017/18 2018/19 2019/20 Annual data release Latest available	
Field code/Field name	**Comment regarding ORFAN Study**
Acardsupdays/Advanced cardiovascular support days	A key variable to adjust for severity of systems based life support required.
Aressupdays/Advanced respiratory support days	A key variable to adjust for severity of systems based life support required.
Bcardsupdays/Basic cardiovascular support days	A key variable to adjust for severity of systems based life support required.
Bestmatch/Best match flag	An administrative variable stating whether the row represents the best match between the critical care and episode start and end dates for the period.
Bressupdays/Basic respiratory support days	A key variable to adjust for severity of systems based life support required.
Ccadmisorc/Critical care admission source	A key variable to understand where the patient came from prior to admission to critical care, an important variable for adjustment.
Ccadmitype/Critical care admission type	The admission type will be critical for adjustment for the nature of the admission being planned or unplanned.
Ccdisdate/Critical care discharge date	A key variable to understand the time frames for critical care admission and to adjudicate when events occur in regards to study relevant CT scans or other events
Ccdisdest/Critical care discharge destination	A key variable to understand patient disposition at the conclusion of the episode.
Ccdisloc/Critical care discharge location	A key variable to understand patient health

	status at the end of the episode, to adjust for severity
Ccdisstat/Critical care discharge status	A key variable to understand and adjust for patient status at discharge, This will inform analysis of risk if a patient dies or survives the episode.
Ccdistime/Critical care discharge time	A key variable to understand the time frames for critical care admission and to adjudicate when events occur in regards to study relevant CT scans or other events
cclev2days/Critical care level 2 days	A key variable to adjust for the length of support required for different levels of care in the ICU, a key marker of severity
cclev3days/Critical care level 3 days	A key variable to adjust for the length of support required for different levels of care in the ICU, a key marker of severity
Ccsorcloc/Critical care source location	A key timing variable to know where patients flowed to the critical care unit from, for assessment of severity
Ccstartdate/Critical care start date	A key timing variable to know when the episode started in relation to other events such as study relevant CT scans
Ccstarttime/Critical care start time	A key timing variable to know when the episode started in relation to other events such as study relevant CT scans
Dermsupdays/Dermatological support days	A key variable to adjust for severity of systems based life support required.
Gisupdays/Gastro-intestinal support days	A key variable to adjust for severity of systems based life support required.
Liversupdays/Liver support days	A key variable to adjust for severity of systems based life support required.
Neurosupdays/Neurological support days	A key variable to adjust for severity of systems based life support required.
Orgsupmax/Organ support maximum	A key variable for severity with the maximum number of organ support systems in place.
Rensupdays/Renal support days	A key variable to adjust for severity of systems based life support required.
Susrecid/SUS record ID	Administrative linkage

Hospital Episode Statistics - Outpatients	
Data controller: NHS Digital	
Introduction: The HES Outpatients database contains all information related to the occurrence of outpatient appointments. These appointments are the source of data for many specialist level diagnoses of critical relevance to the ORFAN study.	
Number of variables requested: We request 16 variables for linkage, with no high risk or identifiable fields requested	
Periods: 2005/06 2006/07 2007/08 2008/09 2009/10 2010/11 2011/12 2012/13 2013/14 2014/15 2015/16 2016/17 2017/18 2018/19 2019/20 Annual data releases Latest available	
Field code/Field name	**Comment regarding ORFAN Study**
ACTIVAGE/Age at activity date	Critical variable for understanding patient age at the time of the outpatient event
APPTAGE/Age on day of appointment	Critical variable to adjust for the age of the patient in years for the time of the appointment, key variable to understand the relationship between diagnoses and study relevant events.
APPTDATE/Appointment date	The date that the appointment was scheduled, for adjustment of temporality to study events
ATENTYPE/Attendance type	Critical variable to adjust for if the patient actually attended the appointment
ATTENDED/Attended or did not attend	A simpler variable for if the patient did attend or did not attend the outpatient appointment
DIAG_NN/Diagnosis	A critical variable to discern what conditions the patient had at the time of the appointment
FIRSTATT/First attendance	Key variable to adjust for
IMD04RK/IMD Overall Ranking	A key variable to adjust for the IMD ranking to understand deprivation effects on care in the analysis.
MAINSPEF/Main specialty	A key variable to adjust for specialty of

	the outpatient clinic appointment.
MATCH_RANK/MATCH_RANK	Administrative linkage
OPERTN_NN/Operative procedure	To adjust to operative procedure codes that patients may have had
PRIORITY/Priority type	A key variable for adjustment of patient priority
REFSOURC/Source of referral	A key variable to understand the origin of the patients referral to specialist services, to adjust for acuity.
STUDY_ID/STUDY_ID	Study linkage
SUSRECID/SUS record id	Administrative linkage
TRETSPEF/Treatment specialty	A key variable to adjust for the specialty of the outpatient clinic

Civil Registration (Deaths) - Secondary Care Cut	
Data controller: NHS Digital	
Introduction: The Civil Registration (Death) database links to all deaths recorded by ONS for adjudication of deaths in the cohort	
Number of variables requested: We request 3 variables for linkage	
Periods: Annual data releases	
Field code/Field name	**Comment regarding ORFAN Study**
cause_of_death/Original Underlying Cause of Death	Critical cause of death variable
Dod/Date of Death	The date of death of the patient, in format MM/YYYY for reduced risk of identification
Match rank/Match rank	Administrative elinkage

HES: Civil Registration (Deaths) bridge dataset	
Data controller: NHS Digital	
Introduction: The HES: Civil Registration (Death) bridge dataset allows linage of HES and civil registration deaths datasets,	
Periods: Annual data releases	
Field code/Field name	**Comment regarding ORFAN Study**
All fields	This dataset does not have variables but allows linkages of HES and civil registration datasets.

NICOR data variables

National Institute for Cardiovascular Outcomes Research (NICOR) data for linkage with ORFAN Arm 4

The data from the datasets held and managed via NICOR that is to be linked to patient identifiers submitted by clinical care teams on behalf of the ORFAN Study investigators. The identifiers will not be shared with the ORFAN study team.

There are multiple datasets managed by NICOR for which we intend to link with. These are listed individually in tables that provide a brief introduction to the dataset and state the data controller and total number of variables requested from each dataset. We list each individual variable we wish to link with, and provide a comment as to its justification.

We request data linkage with the following datasets:

- The National Adult Cardiac Surgery Audit (NACSA) dataset
- The Myocardial Ischaemia National Audit Project (MINAP) dataset
- The National Audit of PCI (NAPCI) dataset
- National Audit of Heart Failure (NAHF) dataset
- National Cardiac Rhythm Management (CRM) – electrophysiological studies (EPS) dataset
- National Cardiac Rhythm Management (CMR) - pacemaker and implantable cardiac device (PMICD) dataset
- The National Congenital Heart Disease Audit (NCHDA) dataset
- The Transcatheter Aortic Valve Implantation (TAVI) dataset

National Adult Cardiac Surgery Audit (NACSA) dataset		
Data controller: HQIP		
Introduction: The NACSA dataset contains data on all major heart operations carried out in NHS hospitals and a selection of private hospitals throughout the UK. These data are relevant to the ORFAN Arm 4 project as we need to know if a participant underwent cardiac surgery for ischaemic heart disease or valvular disease that would influence our analysis of their CT scan.		
Number of variables requested: We request 34 variables for linkage of the 171 variables within this dataset (v5.1.0)		
Data variable description	**Code in NACSA Dataset version 5.1.0**	**Comment in regards to ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.09	For linkage, not for extraction
Number of previous MIs	2.03	Required to adjust for the severity of coronary artery disease
Previous PCI	2.05	Required to adjust analysis for previous PCI stends within the coronary artery tree at the time of image analysis
Previous cardiac surgery	2.07	Required to adjust analysis for previous disturbance of the pericardial structures surrounding the heart at time of CT scan analysis.
Diabetes management	2.09	Required to adjust analysis for patients suffering from diabetes, which severely influences epicardial tissues, particularly epicardial adipose tissue.
History of hypertension	2.10	Required to adjust analysis for patients suffering from hypertension as a major cardiovascular disease risk factor.
Cigarette smoking history	2.11	Required to adjust analysis for patients with this common risk factor for coronary artery disease
History of pulmonary disease	2.13	Required to adjust analysis for patients with this common risk factor for atrial cardiac disease including atrial fibrillation
Pre-operative heart-rhythm	2.18	Required to adjust analysis for patients reported to be in non-sinus rhythm, which would lead to re-classification of the patient in analysis of CT scans for atrial fibrillation risk
Extent of coronary vessel disease	2.21	Required to stratify patients for analysis by the actual anatomical severity of coronary artery disease, which will be utilised to test the ability of CT image interpretation algorithms to predict disease severity

Left main stem disease	2.22	Required to stratify patients by actual anatomical occurrence of LMS disease, which will be utilised to test the ability of CT image interpretation algorithms to predict the occurrence of LMS disease.
Severity of Aortic Valve Stenosis (EOA in cm2)	2.24.1	Required to stratify patients by underlying aortic valve stenosis severity for analysis of automated CT prediction of aortic stenosis severity
Severity of Aortic Valve Stenosis (Gradient mmHg)	2.24.2	Required to stratify patients by underlying aortic valve stenosis severity for analysis of automated CT prediction of aortic stenosis severity
Category of aortic valve stenosis	2.24.3	Required to stratify patients by underlying aortic valve stenosis severity for analysis of automated CT prediction of aortic stenosis severity
Left Ventricular Function	2.27	Required to adjust analysis of CT images and associated clinical data for underlying left ventricular function, a critical metric for overall heart status
Number of previous heart operations	2.36	To stratify patients by number of cardiac surgeries, which makes automated CT image analysis more difficult due to repeated disturbance of the heart tissues and in particular the thoracic wall.
Height	2.37	To enable calculation of BMI and total body surface area, two key cardiovascular disease risk factors that all analysis must be adjusted for
Weight	2.38	To enable calculation of BMI and total body surface area, two key cardiovascular disease risk factors that all analysis must be adjusted for
Date and time of operation	3.02	To know if the surgery was performed prior to or after the CT scan of relevance, a critical metric.
CABG	3.11.1	To know what surgery was performed and if it was a coronary artery bypass graft for coronary artery disease, a key variable
Valve	3.11.2	To know what surgery was performed and if it was a valve replacement/repair
Major aortic	3.11.3	To know what surgery was performed and if it was major aortic surgery
Other cardiac procedures	3.11.4	To know what surgery was performed and if it an ungrouped cardiac

		procedure
Other actual cardiac procedures	3.12	To know what surgery was performed and if it was an ungrouped actual cardiac procedure
Total number of distal coronary anastomoses	3.14	To know how many grafts the patient required if they had a CABG, which is a key metric as the automated analysis of CT images must be adjusted for patients with multiple bypass grafts, as the variation in the anatomy becomes very significant
Number of valves replaced/repaired	3.18	To know how many valves were replaced/repaired if the patient underwent a valve surgery, as this variation in anatomy becomes very significant in automated CT analysis, particularly if metallic valves were inserted.
Aortic valve procedure	3.43	To know if the procedure was on the aortic valve, as analysis must be adjusted in different ways for each valve, as some valves cause greater interference with CT scans.
Mitral valve procedure	3.44	To know if the procedure was on the mitral valve, as analysis must be adjusted in different ways for each valve, as some valves cause greater interference with CT scans.
Tricuspid valve procedure	3.45	To know if the procedure was on the tricuspid valve, as analysis must be adjusted in different ways for each valve, as some valves cause greater interference with CT scans.
Pulmonary valve procedure	3.46	To know if the procedure was on the pulmonary valve, as analysis must be adjusted in different ways for each valve, as some valves cause greater interference with CT scans.
Aortic valve implant type	3.47	To know if a metallic valve was inserted, as this will greatly influence CT scan quality
Mitral valve implant type	3.48	To know if a metallic valve was inserted, as this will greatly influence CT scan quality
Patient status at discharge	4.05	To know if the patient died during or after the surgery
Date of discharge/date of death in hospital	4.06	To know the date of death if the patient died in hospital for this admission, relevant to cease follow up

		and to adjust analysis to reflect death

Myocardial Ischaemia National Audit Project (MINAP) dataset		
Data controller: HQIP		
Introduction: The MINAP dataset contains data about the care provided to patients who are admitted to hospital with acute coronary syndromes (heart attacks, both NSETMI and STEMI). These data are highly relevant to the ORFAN Arm 4 project as we need to know if a participant has suffered from coronary disease before or after the CT scan of relevance to the ORFAN Study in order to accurately use the scan for generating risk assessments for coronary disease.		
Number of variables requested: We request 26 variables for linkage of the 142 variables within this dataset (v10.3.2)		
Data variable description	**Code in MINAP Dataset version 10.3.2**	**Comment regarding ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.10	For linkage, not for extraction
Previous MI	2.05	To adjust analysis for this cardiovascular disease risk factor. Previous MI is used in many risk scores we will need to calculate for individual patient risk.
Previous angina	2.06	To adjust analysis for this cardiovascular disease risk factor.
Hypertension	2.07	To adjust analysis for this cardiovascular disease risk factor. Hypertension is used in many risk scores we will need to calculate for individual patient risk.
Hypercholesterolaemia	2.08	To adjust analysis for this cardiovascular disease risk factor. Hypercholesterolaemia is used in many risk scores we will need to calculate for individual patient risk.
Peripheral vascular disease	2.09	To adjust analysis for this cardiovascular disease risk factor. PVD is used in many risk scores we will need to calculate for individual patient risk.
Cerebrovascular disease	2.10	To adjust analysis for this cardiovascular disease risk factor. Previous cerebrovascular is used in many risk scores we will need to calculate for individual patient risk.
Heart Failure	2.13	To adjust analysis for this cardiovascular disease risk factor. Underlying presence of heart failure is used in many risk scores we will need to calculate for individual patient risk.
Serum Cholesterol	2.15	A key investigation finding that is

		incorporated in some classical cardiovascular risk scores which we will calculate
Current smoker	2.16	To adjust analysis for this common disease risk factor.
Diabetes	2.17	To adjust analysis for this common disease risk factor.
Previous Intervention	2.18	To know if the patient had suffered from a previous heart event that required a PCI, for risk stratification analysis
Previous CABG	2.19	To know if the patient had suffered from a previous heart event that required a ACBG, for risk stratification analysis and to guide CT analysis given the disruption to heart anatomy from cardiac surgery
Systolic blood pressure	2.20	A key examination finding that is incorporated into some classical cardiovascular risk scores which we will calculate.
ACE or ARB use	2.25	To know if the patient was taking anti-hypertensive medications, a variable that is critical for risk scores for coronary artery disease
Statin use	2.26	A critical variable given our analysis to detect inflammation around the coronary arteries from CT, which is highly influenced by statin usage
Height	2.29	To calculate BMI and total body surface area
Weight	2.30	To calculate BMI and total body surface area
Infarction site	2.36	To tailor CT analysis by heart location of vessel disease, this is a very important variable for CT analysis of patients who have had heart attacks
Arrival at Hospital	3.06	Date of arrival at the hospital for this event, critical to know in regard to the timing of subsequent CT scans
Initial re-perfusion treatment	3.39	Required to know how the STEMI was initially managed at the initial hospital to inform subsequent analysis of data
Final diagnosis	4.02	The final diagnosis that the patient received for this admission, critical to know for analysis
Death in hospital	4.04	To know if this cardiac event caused death in hospital, a major outcome for

		the ORFAN study
ACE Inhibitor	4.06	To know if the patient was prescribed an ACE Inhibitor anti-hypertensive on discharge, a key medication for clinical risk scores of coronary artery disease risk
Statin	4.07	To know if the patient was prescribed a statin on discharge, a key medication for our CT analysis of coronary inflammation
Coronary Angiography	4.13	To know if NSTEMI patients were referred for CCTA, the investigation of primary relevant to this study
Coronary intervention	4.14	To know if NSTEMI patients were referred for coronary intervention, an investigation of high relevant to this study

National Audit of Heart Failure (NAHF) dataset		
Data controller: HQIP		
Introduction: The NAHF dataset contains data about the care provided to patients who are admitted to hospital with heart failure, including congestive and ischaemic cardiomyopathies. These data are highly relevant to the ORFAN Arm 4 project as we need to know if a participant has suffered from heart failure before or after the CT scan of relevance to the ORFAN Study in order to accurately use the scan for generating risk assessments for causes of heart failure such as ischaemic heart disease and atrial fibrillation.		
Number of variables requested: We request 27 variables for linkage of the 144 variables within this dataset (v4.2.1)		
Data variable description	**Code in NAHF Dataset version 4.2.1**	**Comment regarding ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.09	For linkage, not for extraction
Date of admission	2.00	To know the date of when the patient suffered an admission due to heart failure, critical to know to understand temporality with relevant CT scan
Ischaemic Heart Disease (IHD)	4.00	To adjust analysis for this common risk factor for heart failure
Device therapy (prior to or during this admission)	4.07	To understand if the patient was in a paced rhythm prior to the event, which influences analysis for risk of atrial fibrillation
Hypertension	4.12	To adjust analysis for this common risk factor for heart failure
Diabetes	4.14	To adjust analysis for this common risk factor for heart failure
Cerebral vascular accident (CVA)	4.15	To adjust analysis for this common event related to heart failure and atrial fibrillation
Smoking history	5.02	To adjust analysis for this common risk factor for heart failure
ACE inhibition	7.01	To adjust analysis for this anti-hypertensive if the patient were taking it
ARB	7.04	To adjust analysis for this anti-hypertensive if the patient were taking it
Beta blocker	7.06	To adjust analysis for this anti-hypertensive if the patient were taking it
Aspirin	7.16	To adjust analysis for this anti-platelet if the patient were taking it
CCB	7.22	To adjust analysis for this anti-hypertensive if the patient were taking

		it
Statin	7.24	To adjust analysis for this cholesterol lowering medication if the patient were taking it, critical for our CT analysis
Warfarin	7.26	To adjust analysis for this anti-coagulant medication if the patient were taking it, critical for our CT analysis to predict stroke risk
Other oral anticoagulant	7.28a	To adjust analysis for this anti-coagulant medication if the patient were taking it, critical for our CT analysis to predict stroke risk
Height	8.01	To calculate BMI, a key risk factor
Weight (admission)	8.02a	To calculate BMI, a key risk factor
Systolic blood pressure (admission)	8.06a	To calculate cardiovascular disease risk scores, required for comparison of risk scores between novel scores developed in the study
NT-proBNP	9.14	To adjust for heart failure severity using this marker of atrial strain
ACE inhibition (discharge)	11.01	To adjust analysis to the use of this anti-hypertensive
ARB (discharge)	11.04	To adjust analysis to the use of this anti-hypertensive
Beta blocker (discharge)	11.06	To adjust analysis to the use of this anti-hypertensive
Aspirin (discharge)	11.16	To adjust analysis for this anti-platelet medication, critical for prediction of stroke risk
CCB (discharge)	11.22	To adjust analysis to the use of this anti-hypertensive
Statin (discharge)	11.24	To adjust analysis for this anti-cholesterol medication, critical for any CT analysis to predict cardiovascular disease risk based upon inflammation
Warfarin (discharge)	11.26	To adjust analysis for this anti-coagulant, critical to know for any prediction of future stroke risk
Other oral anticoagulant (discharge)	11.28a	To adjust analysis for this anti-coagulant, critical to know for any prediction of future stroke risk
Confirmed diagnosis of heart failure	14.00	To allow analysis to include patients with confirmed heart failure only

National Audit of Percutaneous Coronary Intervention (NAPCI) dataset		
Data controller: HQIP		
Introduction: The NAPCI dataset contains data about the care provided to patients who are admitted to hospital for PCI. These data are highly relevant to the ORFAN Arm 4 project as we need to know if a participant has suffered from a heart attack that necessitates a PCI before or after the CT scan of relevance to the ORFAN Study in order to accurately use the scan for generating risk assessments for ischaemic heart disease.		
Number of variables requested: We request 27 variables for linkage of the 124 variables within this dataset (v5.6.6)		
Data variable description	**Code in NAHF Dataset version 5.6.6**	**Comment regarding ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.10	For linkage, not for extraction
Clinical syndrome	2.01	To facilitate inclusion of symptom derived identification of coronary disease severity (stable vs acute event)
Indication for intervention	2.02	To facilitate inclusion of diagnostic derived identification of coronary disease severity necessitating PCI
Arrival in first hospital of chain (ACS)	2.08	To know the date when the coronary event occurred, important to place before or after the CT scan of relevance to the study
Presenting ECG	2.10	To understand the nature of the heart attack that led to the PCI to inform CT scan analysis of such patients
Previous MI	2.13	To understand the previous occurrence of important cardiovascular risk factors to adjust analysis
Previous CABG	2.14	To understand the previous occurrence of important cardiovascular risk factors to adjust analysis, particularly for CT analysis when the heart structures have been disturbed by cardiac surgery
Previous PCI	2.15	To understand the previous occurrence of important cardiovascular risk factors to adjust analysis
Diabetes	2.16	To understand the previous occurrence of important cardiovascular risk factors to adjust analysis
Height	2.17	To enable calculation of BMI
Weight	2.18	To enable calculation of BMI
Date and time of operation	3.01	Critical to know if the PCI occuredn before or after the CT of relevance to the study
Vessels attempted	3.09	To understand where PCI occurred

		within the heart to tailor CT analysis
Number of vessels attempted	3.10	To understand the distribution of coronary lesions within vessels requiring stenting, a variable that can influence CT analysis
Number of lesions attempted	3.11	To understand the distribution of coronary lesions requiring stenting, a variable that can influence CT analysis
Number of stents used	3.15	To adjust for the number of stents deployed within the coronary vessels, which can influence CT analysis
Number of drug-eluting stent(s) used		To adjust for the number of drug eluting stents deployed within the coronary vessels, which can influence CT analysis
Diagnostic device(s) used during the procedure	3.19	To understand if additional equipment was required to visualise the coronary lesions for stenting, a valuable variable for CT angiography analysis
Left main stem stenosis (post PCI)	3.27	To know the flow in the coronary artery post the PCI, which is critical for adjusting CT analysis techniques if the CT was performed following the PCI
LAD proximal stenosis (post PCI)	3.28	To know the flow in the coronary artery post the PCI, which is critical for adjusting CT analysis techniques if the CT was performed following the PCI
LAD other stenosis (post PCI)	3.29	To know the flow in the coronary artery post the PCI, which is critical for adjusting CT analysis techniques if the CT was performed following the PCI
RCA stenosis (PCI)	3.30	To know the flow in the coronary artery post the PCI, which is critical for adjusting CT analysis techniques if the CT was performed following the PCI
LCx stenosis (PCI)	3.31	To know the flow in the coronary artery post the PCI, which is critical for adjusting CT analysis techniques if the CT was performed following the PCI
Number of lesions successful	3.32	To know the overall utility of the PCI procedure for the treatment of the coronary lesions, an important value to adjust for in analysis
Status at discharge	4.03	To know if the patient died during this admission
Cholesterol	5.02	To know the total serum cholesterol, a key value for calculation of important cardiovascular risk scores

Smoking status	5.03	To know if the patient was smoking and hence to adjust risk scores for such exposure.
Medical history	5.05	To determine if the patient had relevant cardiovascular risk factors such as hypercholesterolaemia, hypertension, PVD or a previous stroke, for critical adjustment of risk analysis.

National Cardiac Rhythm Management (CRM) – electrophysiological studies (EPS) ablation dataset		
Data controller: HQIP		
Introduction: The CRM EPS ablation dataset contains data on all admissions related to the occurrence of electrophysiological studies. These data are critical to understanding the treatment of heart arrhythmias to adjust for ablation procedures that influence the occurrence of common conditions such as atrial fibrillation, which in turn is critical to the assessment of stroke risk.		
Number of variables requested: We request 19 variables for linkage of the 73 variables within this dataset (v3.05)		
Data variable description	**Code in CRM EPS Dataset version 3.05**	**Comment in regards to ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.09	For linkage, not for extraction
Underlying heart disease	2.01	To enable adjustment for underlying cardiovascular disease conditions that influence risk for outcome such as atrial fibrillation
Previous cardiac surgery or transcutaneous intervention	2.02	To allow adjustment for previous surgery that can influence the occurrence of heart arrhythmia
Document AF pre-procedure	2.04	Critical variable to know if the patient was in AF prior to the procedure for understanding the time relationship between the arrhythmia and the CT scan
Other documented arrhythmias pre-procedure	2.05	Critical variable to know if the patient was in another arrhythmia prior to the procedure for understanding the time relationship between the arrhythmia and the CT scan
Ablation indication	2.06	To adjust analysis for the reason the ablation if occurring, such as symptoms
Previous ablation	2.07	To adjust for previous ablation occurrence, a key risk factor for subsequent relapse of arrhythmia
Procedure date	3.01	The critical variable to understand if the ablation occurred prior to or after the CT scan of relevance to the study
Ablation procedures(s)	3.12	A critical variable to know what ablation procedure occurred to know if it is of relevance to the CT scan analysis
Ablation attempted	3.19	To clarify if ablation was actually performed during the procedure
Ablation energy source(s)	3.21	To know what technique of ablation was used, a key variable for CT analysis, as different techniques (RF, cryo, laser)

		influence image analysis)
Success	3.16	To know if the ablation was immediately successful, a critical variable for analysis
LA size	4.01	A critical value for adjustment of AF and stroke risk, will be imperative for stroke risk assessment in CT for those who have had ablation
LA volume	4.02	A critical value for adjustment of AF and stroke risk, will be imperative for stroke risk assessment in CT for those who have had ablation
Rhythm at start of procedure	4.03	The best variable to know how longstanding the AF has been ongoing, which is critical for adjustment of stroke risk
Death on same admission	5.02	To rule patient dead based upon this admission data
Months since procedure (follow up)	6.01	To enable assessment of follow up data for participants
Complications post discharge	6.02	To adjust for procedural complications and also death due to the procedure and other causes, a critical variable to limit analysis on
Outcomes: frequency of palpitations	6.03	The critical outcome variable, to rule if the ablation was successful or not, a key variable for analysis of tools to predict ablation success
Outcomes: duration of palpitation	6.04	A critical outcome variable, to rule if the ablation was successful or not, a key variable for analysis of tools to predict ablation success

National Cardiac Rhythm Management (CMR) - pacemaker and implantable cardiac device (PMICD) dataset		
Data controller: HQIP		
Introduction: The CRM PMICD dataset contains data on all admissions related to the insertion and management of pacemakers and implantable cardiac devices. These data are critical to understanding the treatment of heart arrhythmias and cardiomyopathies to adjust for procedures that influence the occurrence of common conditions such as atrial fibrillation, which in turn is critical to the assessment of stroke risk.		
Number of variables requested: We request 7 variables for linkage of the 117 variables within this dataset (v5.07)		
Data variable description	**Code in CMR PMICD Dataset version 5.07**	**Comment in regards to ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.08	For linkage, not for extraction
Date of first implant	2.01	A key variable to know when the first ever device was implanted for this patient
Pre-device aetiology	2.02	The critical variable to know what pathology the patient has had to require a cardiac device
Pre-device ECG	2.04	To adjust for the underlying rhythm that may have pre-disposed the patient to heightened stroke risk prior to the device insertion
Procedure date	3.01	To know if the device was inserted before or after the CT scan of relevance to the study
Intervention category	3.11	To know which procedure was actually performed for analysis
Acute complications prior to discharge	5.01	To adjust for complication in hospital including the death of the participant
Late complications from discharge to 1 year	5.02	To adjust for relevant complications following discharge, including death related to the device.

The National Congenital Heart Disease Audit (NCHDA) Dataset		
Data controller: HQIP		
Introduction: The NCHDA dataset contains data on all admissions related to the surgical management of congenital heart disease, both in children and adults. The ORFAN study only includes data collected in those over the age of 18, so childhood data will not be included in this linkage. These data are critical to understanding the surgical treatment of anatomic heart disease in order to adjust for procedures that influence the occurrence of common conditions such as atrial fibrillation, which in turn is critical to the assessment of stroke risk.		
Number of variables requested: We request 8 variables for linkage of the 66 variables within this dataset (v6.1)		
Data variable description	**Code in NCHDA Dataset version 6.1**	**Comment in regard to ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.10	For linkage, not for extraction
Diagnosis	2.01	The critical variable for the diagnosis of congenital conditions and subsequent adjustment of analysis
Previous procedure	2.02	A critical variable for the adjustment of CT analysis techniques for disturbance of the heart structures
Comorbidity present	2.06	To provide a guide to the presence of comorbid conditions as listed in the below variable (2.07)
Comorbid Conditions	2.07	To adjust for comorbid conditions complicating cardiac disease status
Procedure Date	3.01	To know when the procedure occurred in reference to the relevant CT scan for this study
Procedure performed	3.09	To know what surgery was performed for adjustment of analysis
Date of death	4.02	To adjust for date of death in analysis if occurred in hospital
Discharge status	4.03	To know if the patient was discharge from hospital or dies during this admission

The Transcatheter Aortic Valve Implantation (TAVI) Dataset		
Data controller: NICOR		
Introduction: The TAVI dataset contains data on all admissions related to the surgical management of congenital heart disease, both in children and adults. The ORFAN study only includes data collected in those over the age of 18, so childhood data will not be included in this linkage. These data are critical to understanding the surgical treatment of anatomic heart disease in order to adjust for procedures that influence the occurrence of common conditions such as atrial fibrillation, which in turn is critical to the assessment of stroke risk.		
Number of variables requested: We request 26 variables for linkage of the 110 variables within this dataset (v4.09)		
Data variable description	**Code in TAVI Dataset version 4.09**	**Comment in regard to ORFAN Study**
NHS Number	1.03	For linkage, not for extraction
Date of Birth	1.06	For linkage, not for extraction
Postcode	1.09	For linkage, not for extraction
Diabetes	3.01	To adjust analysis for this common cardiovascular disease risk factor
Smoking status	3.02	To adjust analysis for this common cardiovascular disease risk factor
Previous MI and interval between procedure and last MI	3.05	To adjust analysis for this common cardiovascular disease risk factor
History of neurological disease	3.08	To adjust analysis for this common cardiovascular disease risk factor, includes previous strokes, a key variable for TAVI analysis
Extracardiac arteriopathy	3.09	To adjust analysis for this common cardiovascular disease risk factors which includes carotid artery disease, a key risk variable for complications in TAVI
Pre-operative heart rhythm	3.11	To adjust for patients who may have been in atrial fibrillation or other rhythms prior to the procedure and to relevant CT scan for the study
Previous cardiac surgery	4.01	To adjust for previous disruption of cardiac structures that make CT image analysis more difficult
Previous TAVI	4.023	To adjust for prior intervention on the aortic valve, a risk factor for complication of the subsequent procedure
Date of previous TAVI	4.024	To adjudicate if the previous TAVI was prior to or after the relevant CT scan for the study
Height	5.01	To calculate BMI, a key cardiovascular disease risk factor.
Weight	5.02	To calculate BMI, a key cardiovascular

		disease risk factor.
Aortic valve mean gradient (mmHg) (pre-procedure)	6.014	To adjust for pre-procedure aortic valve disease severity, a key variable.
Aortic valvular pathology	6.06	A key variable to know what disease is requiring the aortic valve under go this procedure
Aortic valvular aetiology	6.07	A key variable to know what has led to the aortic valve being disease, a key adjustment for analysis, as not all disease are relevant to the study
LV function	6.08	To adjust for underlying heart functionality, a key predictor of procedural success
Extent of coronary vessel disease	6.09	To adjust for severe coronary artery disease, a key predictor of procedural success and recover
Left main stem disease	6.10	To adjust for left main stem disease, a severe form of coronary artery disease, a key predictor of procedural success and recover
Date and time of operation	7.01	The key variable to know if the procedure occurred prior to or after the CT scan of relevance to the study
Valve successfully deployed	8.01	To know if the procedure actually proceeded to plan, otherwise all other variables are meaningless
Post deployment aortic valve mean gradient (mmHg)	8.023	To adjust for if the procedure provided immediate benefit to the patient or if no improvement was evident
CVA up to Dx	9.081	To know if strokes were a result of the procedure, a significant complication of TAVI which would confound study analysis relevant to the occurrence of strokes if it was not know that the stroke was caused by the procedure.
Death to hospital discharge	9.141	To adjust for if the patient died as a result of this admission, a key variable
Date of discharge or death	10.01	To know the date that the patient died or was discharged following this admission.
Drugs at discharge - antithrombotic	10.031	To adjust analysis for anticoagulant medications, a key adjustment for all prediction of stroke risk in the study
Drugs at discharge - antiplatelet	10.032	To adjust analysis for antiplatelet medications, a key adjustment for all prediction of stroke risk in the study

SSNAP data variables

Sentinel Stroke National Audit Programme (SSNAP) data for linkage with ORFAN Arm 4

The data from the SSNAP Core Dataset 4.0.0 that is to be linked to patient identifiers submitted by clinical care teams on behalf of the ORFAN Arm 4 Study investigators. The identifiers will not be shared with the ORFAN study team.

The data controller for SSNAP is HQIP.

Data variable description	**Code in SSNAP Core Dataset 4.0.0**	**Comment in regards to ORFAN Study**
NHS Number	1.2	For linkage, not for extraction
Date of Birth	1.5	For linkage, not for extraction
What is the diagnosis?	1.9	To know if the patient had a stroke, a TIA or other. Critical for analysis.
Was the patient already an inpatient at the time of the stroke?	1.10	To understand the risk factor of hospitalisation for stroke in regards to the CT scan analysis. This variable may also be used for the exclusion of cases, due to abnormally heightened risk of stroke in hospitalised individuals.
Date/time patient arrived at first hospital	1.13	To know the date of the event, which is critical for knowing if the event occurred before or after the relevant CT scan for the ORFAN study
Comorbidities prior to admission. Including: Congestive heart failure, hypertension, atrial fibrillation (AF), diabetes, stroke/TIA. If the patient was on anti-platelet medication prior to admission if the patient had AF If the patient was on anticoagulant medication prior to admission. What anticoagulant the patient was on prior to admission, if they were taking an anticoagulant.	2.1 (including 2.1.1, 2.1.2, 2.1.3, 2.1.4, 2.1.5, 2.1.6, 2.1.7(a), and 2.1.8)	Critical information to understand the comorbidities that predispose a patient to stroke and would change how the CT image is analysed within the ORFAN study, with particular importance found in the presence of AF. The use of anti-platelet and anti-coagulant medication also change how the analysis will be carried out, as this changes the risk of a

If a new diagnosis of AF was made on admission.		stroke and how this can be predicted from a CT scan.
What the patient's modified Rankin Scale score before this stroke	2.2	Important to be able to adjust the CT scan analysis in light of pre-existing disability.
What the patient's NIHSS score was on arrival	2.3	To be able to adjust analysis in light of stroke severity
Date and time of first brain imaging after the stroke, or if the patient was not imaged.	2.4	Information will enable possible confirmation of stroke severity and ischaemic distribution by local clinical teams who are collaborating within the ORFAN study
What was the type of stroke? Infarction or haemorrhagic	2.5	Critical information, as our analysis is only relevant to ischaemic strokes.
The discharge/transfer information including if the patient died, was discharged home, was discharged to a care home, or elsewhere, or was transferred to another inpatient facility. The data of death if death occurred during the admission for this event.	7.1 (including 7.1.1)	Critical information for analysis in regards to stroke severity and death due to stroke.
Is there documented evidence that the patient is in AF on discharge and if yes, was the patient taking anti-coagulation on discharge or discharged with a plan to start anti-coagulation therapy?	7.10 (including 7.10.1)	Critical information for the adjustment of analysis for future stroke risk assessment from CT images.
6-month post admission data: Is the patient in persistent, permanent, or paroxysmal AF?	8.5	Critical information for adjustment of stroke risk assessment.
6-month post admission data: Is the patient taking anti-platelet, anti-coagulant, lipid lowering therapy, anti-hypertensive therapy.	8.6 (including 8.6.1, 8.6.2, 8.6.3, 8.6.4)	Critical information for adjustment of stroke risk assessment.
6-month post admission data: Since the initial stoke, is the patient had any further strokes, myocardial infarctions or other illness requiring hospitalisation?	8.7 (including 8.7.1, 8.7.2, 8.7.3)	Critical information for understanding patient risk for further vascular events, as explicit occurrence of events is asked for. No dates or locations are requested.

Data not requested from SSNAP database:

Part 1 (demographics, onset and arrival data): 1.1, 1.3, 1.4, 1.6, 1.7, 1.8, 1.11, 1.12, 1.14, 1.15,

Part 2 (casemix/first 24 hours data): 2.1.7 (b), 2.6, 2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 2.14, 2.15,

Part 3 (patient assessments in first 72 hours): No data requested

Part 4 (Admission data): No data requested

Part 5 (patient condition in first 7 days data): No data requested

Part 6 (patient assessments by discharge): No data requested

Part 7 (discharge/transfer data): 7.1.2, 7.1.3, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7., 7.8, 7.9, 7.11, 7.12

Part 8 (six-month post admission data): 8.1, 8.2, 8.3, 8.4.

Appendix 6. ORFAN Study investigators and affiliations

ORFAN Study Investigators

Chief Investigator: Professor Charalambos Antoniades[1]

Investigators:

Dr Henry West[1]

Dr Alexios Antonopoulos[1]

Ms Sheena Thomas[1]

Ms Maria Lyasheva[1]

Dr Evangelos Oikonomou[1]

Dr Nikant Sabharwal[2]

Dr Cheerag Shirodaria[2]

Dr Susan Anthony[2]

Dr Andrew Kelion[2]

Prof Adrian Banning[2]

Dr Rafail Angelos Kotronias1,2

Dr Cheng Xie[1,2]

Dr Rajesh Kumar Kharbanda[1,2]

Dr Attila Kardos[1,4]

Dr David Adlam[3]

Dr Amrita Bajaj[3]

Dr Intrajeet Das[3]

Dr Aparna Deshpande[3]

Dr Praveen Rao[3]

Dr Tarun Mittal[5]

Dr Saeed Mirsadraee[5]

Dr Edward Nicol[5]

Dr Jonathan Rodrigues[6]

Dr Benjamin Hudson[6]

Prof John Greenwood[7]

Prof Colin Berry[8,9]

Prof Stephan Achenbach[10]

Dr Mohamed Marwan[10]

Dr Milind Y Desai[11]

Dr Nicholas Screaton[12]

Dr Pál Maurovich-Horvat[13]

Prof Guo-Wei He[14]

Dr Wen-Hua Lin[14]

Dr Li-Juan Fan[14]

Prof Naohiko Takahashi[15]

Dr Hidekazu Kondo[15]

Dr Neng Dai[16]

Prof. Junbo Ge[16]

Prof Bon-Kwon Koo[17]

Dr Gianluca Pontone[18]

Dr Marco Guglielmo[18]

Prof Ron Blankstein[19,20]

Prof Theodora Benedek[21]

Dr Ronak Rajani[22]

Dr Mak Sze Mun[22]

Dr Giulia Benedetti[22]

Dr Rebecca Louise Preston[22]

Dr Elisa McAlindon[23]

Dr Shahzad Munir[23]

Dr Derek Leslie Connolly[24]

Dr William Bradlow[25,26]

Dr Matthias Schmitt[27,28]

Dr Fabiano Serfaty[29]

Dr Ilan Gottlieb[30]

Prof Mario Fritsch T. Neves[31]

Prof David Ernest Newby[32]

Prof Steffen E Petersen[33,34]

Dr Francesca Pugliese[35]

Dr Nehal N Mehta[36]

Prof Stéphane Hatem[37]

Prof Alban Redheuil[38]

Dr Georgios Benetos[39]

Prof Meinrad Beer[40]

Dr Gastón A Rodríguez-Granillo[41]

Prof Joseph Selvanayagam[42]

Prof Bernard Gersh[43]

Dr Francisco Lopez-Jimenez[43]

Dr Ruben De Bosscher[44]

Dr Alain Tavildari[45]

Prof Gemma Figtree[46]

Dr Ibrahim Danad[47]

Dr Ronney Shantouf[48]

Dr Bas Kietselaer[49]

Prof Dimitris Tousoulis[50]

Prof George Dangas[51]

Prof Stefan Neubauer[1]

Prof John Deanfield[52,53]

Prof Keith Channon[1]

ORFAN Study Investigator Institutional Affiliations

1. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, UK
2. Oxford University Hospitals NHS Foundation Trust, Oxford, UK
3. Department of Cardiovascular Sciences & NIHR Leicester Biomedical Research Centre, University of Leicester, UK
4. Milton Keynes University Hospital NHS Foundation Trust, UK
5. Royal Brompton and Harefield NHS Foundation Trust, UK
6. Royal United Hospitals Bath NHS Foundation Trust & Department of Health, University of Bath, UK
7. Leeds Teaching Hospitals NHS Foundation Trust, UK
8. NHS Greater Glasgow and Clyde, NHS Scotland, UK
9. Golden Jubilee National Hospital, NHS Scotland, UK
10. Department of Cardiology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
11. Cleveland Clinic Heart and Vascular Institute, Cleveland, U.S.A
12. Royal Papworth Hospital NHS Trust, Cambridge, UK
13. Department of Radiology, MTA-SE Cardiovascular Imaging Research Group, Budapest, Hungary
14. TEDA International Cardiovascular Hospital, Tianjin, China
15. Oita University, Japan
16. Fudan University, China
17. Seoul National University, Seoul, South Korea
18. Centro Cardiologico Monzino IRCCS, University of Milan, Italy
19. Harvard Medical School, Boston, USA
20. Brigham and Women's Hospital, Boston, USA
21. University of Medicine and Pharmacy of Tirgu Mures, Romania
22. Guy's and St Thomas' NHS Foundation Trust, UK
23. Heart and Lung Centre, New Cross Hospital, Wolverhampton, UK
24. Sandwell & West Birmingham Hospitals NHS Trust, UK
25. University Hospitals Birmingham (UHB) NHS Trust, UK
26. University of Birmingham, UK
27. University Hospital of Manchester Foundation Trust, Manchester, UK
28. University of Manchester, Manchester, UK
29. Serfaty Clinicas, Rio de Janeiro, Brazil
30. Casa de Saúde São José, Rio de Janeiro, Brazil
31. Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
32. University of Edinburgh, Royal Infirmary, Edinburgh, UK
33. The William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, UK
34. Bart's Heart Centre, Barts Health NHS Trust, UK
35. NIHR Barts Cardiovascular Biomedical Research Centre, the William Harvey Research Institute, Queen Mary University of London & Barts Health NHS Trust, UK
36. National Institutes of Health, National Heart, Lung, and Blood Institute, USA
37. Foundation for Innovation in Cardiometabolism and Nutrition (ICAN), Paris, France
38. Sorbonne Université, Faculté de Médecine Pierre et Marie Curie, La Pitié Salpêtrière AP-HP, Paris, France
39. Lefkos Stavros Clinic Athens, Greece
40. Universitätsklinikum Ulm, Germany
41. ENERI Medical Institute, Buenos Aires, Argentina
42. South Australia Health and Medical Research Institute, Adelaide, Australia
43. Mayo Clinic, Rochester, USA
44. University Hospitals Leuven, Belgium
45. Cardiovista, France
46. University of Sydney, Australia
47. Amsterdam University Medical Centers, Netherlands
48. Cleveland Clinic Abu Dhabi, UAE
49. Zuyderland Medical Centre, Heerlen, Netherlands
50. University of Athens, Greece
51. Mount Sinai School of Medicine, New York, USA
52. University College London, UK
53. National Institute of Cardiovascular Outcomes Research (NICOR), UK

Appendix 7. Disease and risk factor ICD-10 groupings

Appendix 7 - Table 1. ORFAN Study risk factor and disease ICD-10 code groupings

ORFAN Study disease and risk factor groupings	*Group Code*	*Single Code*	*WHO definition of single code*
Diabetes mellitus	E11	E11	Type 2 diabetes mellitus
	E11	E11.0	Type 2 diabetes mellitus with coma
	E11	E11.1	Type 2 diabetes mellitus with ketoacidosis
	E11	E11.2	Type 2 diabetes mellitus with renal complications
	E11	E11.3	Type 2 diabetes mellitus with ophthalmic complications
	E11	E11.4	Type 2 diabetes mellitus with neurological complications
	E11	E11.5	Type 2 diabetes mellitus with peripheral circulatory complications
	E11	E11.6	Type 2 diabetes mellitus with other specified complications
	E11	E11.7	Type 2 diabetes mellitus with multiple complications
	E11	E11.8	Type 2 diabetes mellitus with unspecified complications
	E11	E11.9	Type 2 diabetes mellitus without complications
Hypercholesterolaemia	E78	E78	Disorders of lipoprotein metabolism and other lipidaemias
	E78	E78.0	Pure hypercholesterolaemia
	E78	E78.1	Pure hyperglyceridaemia
	E78	E78.2	Mixed hyperlipidaemia
	E78	E78.3	Hyperchylomicronaemia
	E78	E78.4	Other hyperlipidaemia
Hypertension	E78	E78.5	Hyperlipidaemia, unspecified
	I11	I11	Hypertensive heart disease
	I11	I11.0	Hypertensive heart disease with (congestive) heart failure
	I11	I11.9	Hypertensive heart disease without (congestive) heart failure
	I12	I12	Hypertensive renal disease
	I12	I12.0	Hypertensive renal disease with renal failure
	I12	I12.9	Hypertensive renal disease without renal failure
	I13	I13	Hypertensive heart and renal disease
	I13	I13.0	Hypertensive heart and renal disease with (congestive) heart failure
	I13	I13.1	Hypertensive heart and renal disease with renal failure
	I13	I13.2	Hypertensive heart and renal disease with both (congestive) heart failure and renal failure
	I13	I13.9	Hypertensive heart and renal disease, unspecified
	I15	I15	Secondary hypertension
	I15	I15.0	Renovascular hypertension
	I15	I15.1	Hypertension secondary to other renal disorders
	I15	I15.2	Hypertension secondary to endocrine disorders
	I15	I15.8	Other secondary hypertension
	I15	I15.9	Secondary hypertension, unspecified
Angina pectoris	I20	I20	Angina pectoris
	I20	I20.0	Unstable angina
	I20	I20.1	Angina pectoris with documented spasm
	I20	I20.8	Other forms of angina pectoris
	I20	I20.9	Angina pectoris, unspecified
Acute myocardial infarction	I21	I21	Acute myocardial infarction
	I21	I21.0	Acute transmural myocardial infarction of anterior wall

	I21	I21.1	Acute transmural myocardial infarction of inferior wall
	I21	I21.2	Acute transmural myocardial infarction of other sites
	I21	I21.3	Acute transmural myocardial infarction of unspecified site
	I21	I21.4	Acute subendocardial myocardial infarction
	I21	I21.9	Acute myocardial infarction, unspecified
	I22	I22	Subsequent myocardial infarction
	I22	I22.0	Subsequent myocardial infarction of anterior wall
	I22	I22.1	Subsequent myocardial infarction of inferior wall
	I22	I22.8	Subsequent myocardial infarction of other sites
	I22	I22.9	Subsequent myocardial infarction of unspecified site
	I24	I24.9	Acute ischaemic heart disease, unspecified
Chronic coronary syndrome	I25	I25	Chronic ischaemic heart disease
	I25	I25.0	Atherosclerotic cardiovascular disease, so described
	I25	I25.1	Atherosclerotic heart disease
	I25	I25.2	Old myocardial infarction
	I25	I25.3	Aneurysm of heart
	I25	I25.4	Coronary artery aneurysm
	I25	I25.5	Ischaemic cardiomyopathy
	I25	I25.6	Silent myocardial ischaemia
	I25	I25.8	Other forms of chronic ischaemic heart disease
	I25	I25.9	Chronic ischaemic heart disease, unspecified
Valve disease	I34	I34	Nonrheumatic mitral valve disorders
	I34	I34.0	Mitral (valve) insufficiency
	I34	I34.1	Mitral (valve) prolapse
	I34	I34.2	Nonrheumatic mitral (valve) stenosis
	I34	I34.8	Other nonrheumatic mitral valve disorders
	I34	I34.9	Nonrheumatic mitral valve disorder, unspecified
	I35	I35	Nonrheumatic aortic valve disorders
	I35	I35.0	Aortic (valve) stenosis
	I35	I35.1	Aortic (valve) insufficiency
	I35	I35.2	Aortic (valve) stenosis with insufficiency
	I35	I35.8	Other aortic valve disorders
	I35	I35.9	Aortic valve disorder, unspecified
	I36	I36	Nonrheumatic tricuspid valve disorders
	I36	I36.0	Nonrheumatic tricuspid (valve) stenosis
	I36	I36.1	Nonrheumatic tricuspid (valve) insufficiency
	I36	I36.2	Nonrheumatic tricuspid (valve) stenosis with insufficiency
	I36	I36.8	Other nonrheumatic tricuspid valve disorders
	I36	I36.9	Nonrheumatic tricuspid valve disorder, unspecified
	I37	I37	Pulmonary valve disorders
	I37	I37.0	Pulmonary valve stenosis
	I37	I37.1	Pulmonary valve insufficiency
	I37	I37.2	Pulmonary valve stenosis with insufficiency
	I37	I37.8	Other pulmonary valve disorders
	I37	I37.9	Pulmonary valve disorder, unspecified
Atrial fibrillation	I48	I48	Atrial fibrillation and flutter
	I48	I48.0	Paroxysmal atrial fibrillation
	I48	I48.1	Persistent atrial fibrillation
	I48	I48.2	Chronic atrial fibrillation
	I48	I48.3	Typical atrial flutter

	I48	I48.4	Atypical atrial flutter
	I48	I48.9	Atrial fibrillation and atrial flutter, unspecified
Heart failure	I50	I50	Heart failure
	I50	I50.0	Congestive heart failure
	I50	I50.1	Left ventricular failure
	I50	I50.9	Heart failure, unspecified
Haemorrhagic stroke	I61	I61	Intracerebral haemorrhage
	I61	I61.0	Intracerebral haemorrhage in hemisphere, subcortical
	I61	I61.1	Intracerebral haemorrhage in hemisphere, cortical
	I61	I61.2	Intracerebral haemorrhage in hemisphere, unspecified
	I61	I61.3	Intracerebral haemorrhage in brain stem
	I61	I61.4	Intracerebral haemorrhage in cerebellum
	I61	I61.5	Intracerebral haemorrhage, intraventricular
	I61	I61.6	Intracerebral haemorrhage, multiple localized
	I61	I61.8	Other intracerebral haemorrhage
	I61	I61.9	Intracerebral haemorrhage, unspecified
Ischaemic stroke	I63	I63	Cerebral infarction
	I63	I63.0	Cerebral infarction due to thrombosis of precerebral arteries
	I63	I63.1	Cerebral infarction due to embolism of precerebral arteries
	I63	I63.2	Cerebral infarction due to unspecified occlusion or stenosis of precerebral arteries
	I63	I63.3	Cerebral infarction due to thrombosis of cerebral arteries
	I63	I63.4	Cerebral infarction due to embolism of cerebral arteries
	I63	I63.5	Cerebral infarction due to unspecified occlusion or stenosis of cerebral arteries
	I63	I63.6	Cerebral infarction due to cerebral venous thrombosis, nonpyogenic
	I63	I63.8	Other cerebral infarction
	I63	I63.9	Cerebral infarction, unspecified
Stroke unspecified	I64	I64	Stroke, not specified as haemorrhage or infarction
Peripheral vascular disease	I70	I70	Atherosclerosis
	I70	I70.0	Atherosclerosis of aorta
	I70	I70.00	Atherosclerosis of aorta without gangrene
	I70	I70.01	Atherosclerosis of aorta with gangrene
	I70	I70.1	Atherosclerosis of renal artery
	I70	I70.10	Atherosclerosis of renal artery without gangrene
	I70	I70.11	Atherosclerosis of renal artery with gangrene
	I70	I70.2	Atherosclerosis of arteries of extremities
	I70	I70.20	Atherosclerosis of arteries of extremities without gangrene
	I70	I70.21	Atherosclerosis of arteries of extremities with gangrene
	I70	I70.8	Atherosclerosis of other arteries
	I70	I70.80	Atherosclerosis of other arteries without gangrene
	I70	I70.81	Atherosclerosis of other arteries with gangrene
	I70	I70.9	Generalized and unspecified atherosclerosis
	I70	I70.90	Generalized and unspecified atherosclerosis without gangrene
	I70	I70.91	Generalized and unspecified atherosclerosis with gangrene
	I71	I71	Aortic aneurysm and dissection
	I71	I71.0	Dissection of aorta [any part]
	I71	I71.1	Thoracic aortic aneurysm, ruptured
	I71	I71.2	Thoracic aortic aneurysm, without mention of rupture
	I71	I71.3	Abdominal aortic aneurysm, ruptured

	I71	I71.4	Abdominal aortic aneurysm, without mention of rupture
	I71	I71.5	Thoracoabdominal aortic aneurysm, ruptured
	I71	I71.6	Thoracoabdominal aortic aneurysm, without mention of rupture
	I71	I71.8	Aortic aneurysm of unspecified site, ruptured
	I71	I71.9	Aortic aneurysm of unspecified site, without mention of rupture
	I72	I72	Other aneurysm and dissection
	I72	I72.0	Aneurysm and dissection of carotid artery
	I72	I72.1	Aneurysm and dissection of artery of upper extremity
	I72	I72.2	Aneurysm and dissection of renal artery
	I72	I72.3	Aneurysm and dissection of Iliac artery
	I72	I72.4	Aneurysm and dissection of artery of lower extremity
	I72	I72.5	Aneurysm and dissection of other precerebral arteries
	I72	I72.6	Aneurysm and dissection of vertebral artery
	I72	I72.8	Aneurysm and dissection of other specified arteries
	I72	I72.9	Aneurysm and dissection of unspecified site
	I73	I73	Other peripheral vascular diseases
	I73	I73.1	Thromboangiitis obliterans [Buerger]
	I73	I73.8	Other specified peripheral vascular diseases
	I73	I73.9	Peripheral vascular disease, unspecified