Improving stratification and management of patients with atrial fibrillation and heart failure using novel technology

By

Simrat Kaur Gill MPHARM, MBChB, MRCP (UK) Student ID: 601498

A thesis submitted to the University of Birmingham for the degree of Doctor of PHILOSOPHY

Institute of Cardiovascular sciences

School and Medical and Dental Sciences

University of Birmingham

June 2023

Abstract

Atrial fibrillation (AF) and heart failure (HF) commonly co-exist, and exhibit higher rates of morbidity and mortality. Treatment remain unclear due to the lack of robust data regarding treatment efficacy in this specific population, and current ones have unpredictable responses. The use of big data and artificial intelligence (AI) based analysis has allowed us to find better ways to identify and sub-categorise this cohort of patients based on treatment response, that can improve outcomes and quality of life.

This PhD looked at the use of big data and machine learning to improve phenotyping and prognoses in patients with AF and HF. A systematic review and meta-analysis of novel screening techniques for AF detection found smartphone photoplethysmography had a high sensitivity and specificity for AF detection when compared to an ECG (n=28 studies). A meta-analysis of 20 comparisons (n=17 studies; n=6,891; 2,299 with AF) found a pooled sensitivity of 94% (95% CI 92-95%) and specificity of 97% (CI 96-98%), with significant heterogeneity ($p<0.01$). Studies were found to be small, of poor quality and biased.

Following this, a consumer wearable Fitbit device measuring PPG heart rate and physical activity (step count), was embedded within an on-going randomised clinical trial (RATE-AF) to compare the effectiveness of beta blockers and low-dose digoxin in AF rate control. Over 143 million heart rate recordings and 23 million corresponding physical activity intervals were collected over a mean of 20 weeks. No significant difference was seen in heart rate control with beta-blockers or low-dose digoxin, the regression coefficient was 1.22 (95% CI -2.82 to 5.27; $p=0.55$). There remained no difference in heart rate after adjusting for clinical variables ($p=0.75$), and individual physical activity ($p=0.74$). Conventional trial assessments and a wearable neural network trained on sensor data, were used to predict future trial

outcomes. The wearable neural network showed a similar performance for predicting future New York Heart Association class: F1 score 0.55 (95% CI 0.40 to 0.70), versus 0.59 for electrocardiogram and 6-minute walk test (95% CI 0.44 to 0.74; p=0.72 for comparison).

The limitations of conventional statistical analysis in the presence of multi-morbidity were highlighted in the meta-analysis, assessing beta-blocker efficacy in heart failure with reduced ejection fraction (HFrEF) and renal dysfunction. This was only able to detect a weak interaction between beta blocker efficacy and renal function, with low power (p=0.021). An AI based clustering technique was evaluated to overcome this. The clustering technique identified clusters of HFrEF patients with different beta-blocker efficacy in sinus rhythm and AF. A cluster of older patients in sinus rhythm were found to have no benefit from beta-blockers, and a cluster of younger patients with AF were found to have a significant benefit.

These studies have highlighted the potential wearable devices and AI-based analysis can offer in terms of improving patient care.

Dedication

To my husband Pavan and my children Avaani and Aariyan, without you I would have finished a lot quicker, but I am eternally grateful for your support and understanding especially during my write up.

Acknowledgements

The work I have presented in this thesis was carried out during my PhD under the supervision and guidance of: Professor Dipak Kotecha (Professor of Cardiology and Cardiac Imaging and Consultant Cardiologist) from the Institute of Cardiovascular Sciences at the University of Birmingham; and Professor George Gkoutos (Associate Director of Health Data Research UK, and Chair of Clinical Bioinformatics from the Institute of Cancer and Genomic sciences at the University of Birmingham). I would like to thank them for providing me with this exciting opportunity to undertake research in a growing field of cardiovascular research. I am very thankful for their guidance, support and patience throughout my PhD. My lead supervisor Professor Kotecha has been invaluable and has continued to offer guidance and encouragement at each stage of this journey. I would also like to thank Professor Gkoutos for his ongoing support, encouragement and wisdom as I navigated through this process.

I would also like to acknowledge and thank the collaborators who have worked on several of my projects: Dr Karina Bunting as part of the RATE AF trial, for her help and guidance on setting up the wearables sub-study within the trial**.** I would also like to thank the Application of Artificial Intelligence to Routine Healthcare Data to Benefit Patients with Cardiovascular Disease (card*AI*c) team and the Health Data Research UK team based at the University of Birmingham for their help and support with processing the wearable data and carrying out artificial intelligence-based analysis. In particular, Dr Andrey Barsky for his help with the wearable AI-based analysis and Dr Andreas Karwath for his teaching guidance and support with the AI toolkits framework. I would also like to thank the BigData@Heart consortium for supporting my PhD through a grant and for the numerous international collaborations. Finally, I would like to thank my husband Pavan, and my family for their

overwhelming support, guidance and cups of tea, without them this would not have been possible.

The RATE-AF trial was funded by the National Institute for Health & Care Research (NIHR; CDF-2015-08-074). The wearable devices sub-study received funding from the European Union/European Federation of Pharmaceutical Industries and Associations Innovative Medicines Initiative 2 Joint Undertaking (BigData@Heart; 116074). The card*AI*c team at the University of Birmingham and University Hospitals Birmingham NHS Foundation Trust have received support and funding from the NIHR Birmingham Biomedical Research Centre (NIHR203326), MRC Health Data Research UK (HDRUK/CFC/01), NHS Data for R&D Subnational Secure Data Environment programme (West Midlands), and a British Heart Foundation Accelerator Award (AA/18/2/34218).

The opinions expressed in this thesis do not represent any of the listed funders; none of the organizations had any role in design or conduct of the study (including collection, analysis and interpretation of the data) or any involvement in preparation, review or approval of this thesis.

Published research

1**. S. K. Gill**, A. Karwath, H. W. Uh, V. R. Cardoso, Z. Gu, A. Barsky, et al. Artificial intelligence to enhance clinical value across the spectrum of cardiovascular healthcare. Eur Heart J 2023 Vol. 44 Issue 9 Pages 713-725. DOI: 10.1093/eurheartj/ehac758

A first author publication and the information can be found in chapter 1.

2. **S. Gill**, K. V. Bunting, C. Sartini, V. R. Cardoso, N. Ghoreishi, H. W. Uh, et al. Smartphone detection of atrial fibrillation using photoplethysmography: a systematic review and meta-analysis. Heart 2022 Vol. 108 Issue 20 Pages 1600-1607. DOI: 10.1136/heartjnl-2021-320417

A first author publication and the information can be found in chapter 3.3. D. Kotecha, K. V. Bunting, **S. K. Gill**, S. Mehta, M. Stanbury, J. C. Jones, et al. Effect of Digoxin vs Bisoprolol for Heart Rate Control in Atrial Fibrillation on Patient-Reported Quality of Life: The RATE-AF Randomized Clinical Trial. Jama 2020 Vol. 324 Issue 24 Pages 2497-2508. DOI: 10.1001/jama.2020.23138

A third author publication and my contribution included: conducting and reviewing patients during trial visits, data acquisition, analysis and interpretation, manuscript drafting and revision. The information can be found in chapter 4. Other author contributions are listed below.

Authors: Kotecha, Mehta, Stanbury, Jones, Haynes, Calvert, Deeks, Steeds, Strauss, Griffith, Lip, Kirchhof contributed to the concept and design of the study.

Authors: Kotecha, Bunting, Mehta, Calvert, Deeks, Steeds, Strauss, Rahimi, Camm, Griffith, Townend, Kirchhof contributed to data acquisition and interpretation.

Authors: Kotecha, Bunting, Mehta, Steeds contributed to manuscript drafting.

Authors: Kotecha, Bunting, Gill, Stanbury, Jones, Haynes, Calvert, Deeks, Steeds, Strauss, Rahimi, Camm, Griffith, Lip, Townend, Kirchhof contributed to manuscript revision.

Authors: Mehta, Deeks, Strauss conducted the statistical analysis

4. **Gill SK**, Barsky A, Bunting KV, Karwath A, Tica O, Stanbury M, Haynes S, Folarin A, Dobson R, Kurps J, Asselbergs FW, Grobbee D, Camm J, Eijkemans M, Gkoutos G and Kotecha D. Wearable devices to evaluate dynamic heart rate control with digoxin versus beta-blockers in the RATE-AF randomised trial

A first author manuscript currently in submission. The information can be found in chapter 5.

5. D. Kotecha, **S. K. Gill**, M. D. Flather, J. Holmes, M. Packer, G. Rosano, et al. Impact of Renal Impairment on Beta-Blocker Efficacy in Patients with Heart Failure. J Am Coll Cardiol 2019 Vol. 74 Issue 23 Pages 2893-2904. DOI: 10.1016/j.jacc.2019.09.059

A second author publication and my contribution included: data interpretation and manuscript drafting. The information can be found in chapter 6. Other author contributions are listed below.

On behalf of the Beta-Blockers in Heart Failure collaborative group: authors Kotecha, Flather, Holmes, Packer, Rosano, Bohm, McMurray, Wikstand, Anker, van Veldhuisen, Manzano, von Lueder, Rigby, Andersson, Kjekshus , Wedel, Ruschitzka, Cleland, Damman, Redon, and Coats contributed to concept and design of the study. As well as data interpretation and manuscript drafting.

6. A. Karwath, K. V. Bunting, **S. K. Gill**, O. Tica, S. Pendleton, F. Aziz, et al. Redefining β-blocker response in heart failure patients with sinus rhythm and atrial fibrillation: a machine learning cluster analysis. Lancet 2021 Vol. 398 Issue 10309 Pages 1427-1435. DOI: https://doi.org/10.1016/S0140-6736(21)01638-X

A third author publication and my contribution included: concept and design of the study and manuscript drafting. The information can be found in chapter 6. Other author contributions are listed below.

Authors: Karwath and Kotecha contributed to concept and design of study, data acquisition and analysis.

Authors Karwath, Bunting, Gill, Tica, Pendleton, Aziz, Barsky, Chernbumroong, Duan, Mobley, Cardoso, Slater, Williams, Bruce, Wang, Flather, Coats, Gkoutos and Kotecha, KB, contributed to manuscript drafting and editing the manuscript for intellectual content.

Contents

Abstract 2

Dedication 4

Acknowledgements 5

Published research 7

Contents 10

Extent of personal contributions 13

List of tables 14

List of figures 16

List of abbreviations 18

1. Introduction 22

1.1 Atrial Fibrillation 23

Table 1: EHRA symptom score assessing effect of AF on daily life 25

Table 2: Risk score for predicting ischaemic stroke in AF 26

1.2 Heart failure 28

Table 3: New York Heart Association functional classification 29

1.3 Atrial fibrillation and heart failure 31

1.4 Big data and cardiovascular disease 37

Table 4: The strengths and limitations of data sources used for big data applications 41

Table 5: Opportunities of big data analytics in cardiovascular research 49

1.5 Wearable technology 52

Table 6: Application of wearable technology in cardiovascular care 56

1.6 Stepwise workflow for the application of artificial intelligence techniques 59

Table 7: Type of data modality and data collected 62

Table 8: Patient and public perspective on artificial intelligence 68

2. Research aims and hypotheses 70

3. A systematic review and meta-analysis of smartphone detection of atrial fibrillation using photoplethysmography....73

3.1 Introduction....73

3.2 Methods....79

Table 9: Inclusion and exclusion criteria....80

Table 10: Calculating sensitivity, specificity positive predictive value and negative predictive value for AF detection using PPG signals from smartphone camera....81

Table 11: Summary of full text publications....85

Table 12: Summary of characteristics for conference abstracts....87

Table 13: QUADAS tool assessment questions....91

3.3 Results....93

Table 14: Baseline demographics of participants included in systematic review....96

Table 15: Risk of bias and applicability concerns for full text studies....98

Table 16: Risk of bias and applicability concerns for conference abstracts....99

Table 17: Summary of sensitivity, specificity and accuracy of PPG....101

3.4 Discussion....107

3.5 Conclusion....112

4 Design and rationale of the RATE trial and wearables substudy....113

Table 18: Inclusion criteria for RATE AF trial....115

Table 19: Exclusion criteria for RATE AF Trial....116

Table 20: Schedule of assessments for RATE AF trial....119

4.2 RATE-AF main trial results....120

4.3 Design and rationale of the RATE-AF wearables substudy....121

Table 21: Android smartphone Photoplethysmography performance....127

4.4 Statistical analysis plan....143

Chapter 5: The use of wearable technology to evaluate dynamic heart rate control in the RATE-AF randomised trial....147

Table 22 : Baseline characteristics of patients included or declined participation....151

Table 23: Baseline characteristics by randomised group 153

Table 24: Compliance with technology at interim review 155

Table 25: Usage of wearable device at interim review 156

Table 26: Clinical characteristics at the end of the sub-study 157

Table 27: Usage of smartphone and applications at end of sub-study 158

Table 28: Usage of wearable device measured at end of sub-study 160

Table 29: Sensor data collection from wearables 162

Table 30: Wearable sensor parameters of heart rate and physical activity over sequential five-week periods 165

Chapter 6: Other approaches to phenotype and determining prognosis in atrial fibrillation and heart failure 175

Table 31: National Kidney foundation staging 180

Table 32: Baseline characteristics by eGFR category for patients in sinus rhythm 181

Table 33: Hazard ratio according to renal function for beta-blocker versus placebo in sinus rhythm 185

Table 34: Hazard ratio according to renal function for beta-blocker versus placebo in atrial fibrillation 186

Table 35: Baseline characteristics 193

Table 36: Cluster-specific all-cause mortality results for sinus rhythm and atrial fibrillation 196

Table 37: Validation protocol in the sinus rhythm cohort 199

Table 38: Validation protocol in the atrial fibrillation cohort 199

Chapter 7: Discussion 204

References 210

Appendices 231

Extent of personal contributions

For the RATE-AF trial, I conducted the medical examinations and reviews at clinical visits for the latter part of the trial. For the sub-study, I designed, set up and incorporated the wearables sub-study into the RATE-AF clinical trial. I was responsible for identifying and contacting participants for enrolment, consent and setting up the hardware and software. I also arranged and performed all study and ad-hoc visits. If a sub-study visit coincided with a RATE-AF clinical trial appointment, then the electrocardiogram and serum NT-proBNP sample was obtained by the research nurse Patience Domingos. I designed the electronic case reporting system using REDCap (Research Electronic Data Capture; Vanderbilt/US National Institutes of Health) and completed all uploads following the appointment. Periodic data checks were carried out with the assistance of Dr Saisakul Chernbumroong, a data scientist from our card*AI*c team.

All statistical analyses were performed by me, with guidance from Professor Dipak Kotecha. Neural network-based analyses were carried out by Dr Andrey Barsky, a machine learning expert in our card*AI*c team. For the artificial intelligence framework, I worked as part of an international collaboration through the Big Data@Heart Innovative Medicines Initiative and took on the lead role of organising individual collaborations through regular meetings and drafted the manuscript. Assoc Professor Andreas Karwath, machine learning lead in the card*AI*c team, provided support and guidance with the manuscript, and performed the machine learning aspects discussed in Chapter 6.

Professors Dipak Kotecha and George Gkoutos both provided help and guidance with the drafting of manuscripts.

List of tables

Table 1: EHRA symptom score assessing effect of AF on daily life 25

Table 2: Risk score for predicting ischaemic stroke in AF 26

Table 3: New York Heart Association functional classification 29

Table 4: The strengths and limitations of data sources used for big data applications 41

Table 5: Opportunities of big data analytics in cardiovascular research 49

Table 6: Application of wearable technology in cardiovascular care 56

Table 7: Type of data modality and data collected 62

Table 8: Patient and public perspective on artificial intelligence 68

Table 9: Inclusion and exclusion criteria 80

Table 10: Calculating sensitivity, specificity positive predictive value and negative predictive value for AF detection using PPG signals from smartphone camera 81

Table 11: Summary of full text publications 85

Table 12: Summary of characteristics for conference abstracts 87

Table 13: QUADAS tool assessment questions 91

Table 14: Baseline demographics of participants included in systematic review 96

Table 15: Risk of bias and applicability concerns for full text studies 98

Table 16: Risk of bias and applicability concerns for conference abstracts 99

Table 17: Summary of sensitivity, specificity and accuracy of PPG 101

Table 18: Inclusion criteria for RATE AF trial 115

Table 19: Exclusion criteria for RATE AF Trial 116

Table 20: Schedule of assessments for RATE AF trial 119

Table 21: Android smartphone Photoplethysmography performance 127

Table 22 : Baseline characteristics of patients included or declined participation 151

Table 23: Baseline characteristics by randomised group 153

Table 24: Compliance with technology at interim review 155

Table 25: Usage of wearable device at interim review 156

Table 26: Clinical characteristics at the end of the sub-study 157

Table 27: Usage of smartphone and applications at end of sub-study.....158

Table 28: Usage of wearable device measured at end of sub-study 160

Table 29: Sensor data collection from wearables 162

Table 30: Wearable sensor parameters of heart rate and physical activity over sequential five-week periods 165

Table 31: National Kidney foundation staging.....180

Table 32: Baseline characteristics by eGFR category for patients in sinus rhythm 181

Table 33: Hazard ratio according to renal function for beta-blocker versus placebo in sinus rhythm.....185

Table 34: Hazard ratio according to renal function for beta-blocker versus placebo in atrial fibrillation 186

Table 35: Baseline characteristics.....193

Table 36: Cluster-specific all-cause mortality results for sinus rhythm and atrial fibrillation 196

Table 37: Validation protocol in the sinus rhythm cohort 199

Table 38: Validation protocol in the atrial fibrillation cohort.....199

List of figures

Figure 1: Shared pathophysiological mechanisms in heart failure and atrial fibrillation 33

Figure 2: Effect of treatment in patients with HF and reduced ejection fraction from double blind randomised trials.....34

Figure 3: Volume versus depth for different data types 43

Figure 4: Simplified overview of a neural network 48

Figure 5: Typical PPG waveforms seen from wearable devices 59

Figure 6: Artificial intelligence framework for clinical application 61

Figure 7: Examples of analytic steps based on data modality 66

Figure 8: Smartphone camera fingertip photoplethysmography 77

Figure 9: Comparing PPG, blood pressure and ECG signals 78

Figure 10: Flow diagram for systematic review 94

Figure 11: Assessment of bias, weighted funnel plot and bivariate box plot 100

Figure 12: Forest plot of the sensitivity and specificity of smartphone PPG vs ECG.....103

Figure 13: Forest plot of positive and negative prediction of smartphone PPG vs ECG 104

Figure 14: Summary receiver operator characteristic plot.....105

Figure 15: A graphical summary of the main findings within this systematic review and meta-analysis.....106

Figure 16: RATE AF trial schema…….118

Figure 17: The My AF application 129

Figure 18: Quality of life scoring on My AF app 130

Figure 19: Quality of life and symptom scoring on My AF app 131

Figure 20: RADAR platform deployment in the wearables sub-study.....133

Figure 21: The REDCAP interface 135

Figure 22: RADAR Application screen 138

Figure 23: Schematic of data review process 142

Figure 24: Neural network architecture 146

Figure 25: RATE-AF wearables sub-study enrolment flowchart 150

Figure 26: Wearable device data for heart rate and physical activity over a 24-hour period 162

Figure 27: Wearable device data for heart rate and physical activity 163

Figure 28: Wearable device heart rate according to randomised therapy 166

Figure 29: Prediction of clinical progress with wearable data neural network 168

Figure 30: Graphical abstract summary of the wearables sub-study 170

Figure 31: Impact of renal dysfunction on mortality in heart failure with reduced ejection fraction 184

Figure 32: Beta-blocker efficacy according to baseline renal function in sinus rhythm 186

Figure 33: Beta-blocker efficacy according to baseline renal function in atrial fibrillation……...190

Figure 34: Study flow chart……191

Figure 35: Clustering for all-cause mortality and beta-blocker efficacy in sinus rhythm 197

Figure 36: Clustering for all-cause mortality and beta-blocker efficacy in atrial fibrillation. 198

List of abbreviations

AF	Atrial fibrillation
HF	Heart failure
ECG	Electrocardiogram
ESC	European Society of Cardiology
CAD	Coronary artery disease
EHRA	European Heart Rhythm Association
ATRIA	Anticoagulation and Risk factors in atrial fibrillation
ABC	Age, Biomarkers, Clinical history
EAST	Early Rhythm-Control Therapy in Patients with Atrial Fibrillation
LVEF	Left ventricular ejection fraction
HFrEF	Heart failure with reduced ejection fraction
HFmEF	Heart failure with mildly reduced ejection fraction
HFpEF	Heart failure with preserved ejection fraction
CO	Cardiac output
NHYA	New York Heart Association
NTproBNP	N-terminal B-type natriuretic peptide
ARB	Angiotensin receptor blocker
ACEI	Angiotensin converting enzyme inhibitor
ARNI	Angiotensin receptor-neprilysin inhibitor
MRA	Mineralocorticoid receptor antagonist
SGLT2	Sodium glucose co-transporter inhibitor
CRT	Cardiac resynchronization device
LBBB	Left bundle branch block
CRT-P	Cardiac resynchronization device pacemaker
CRT-D	Cardiac resynchronization device defibrillator

MRI	Magnetic resonance imaging
HR	Hazard ratio
CI	Confidence interval
RAAS	Renin-angiotensin-aldosterone system
RCT	Randomised controlled trial
CHARM	Candesartan in heart failure assessment of reduction in mortality and morbidity
EHR	Electronic health record
CT	Computer Tomography
ML	Machine learning
CVD	Cardiovascular disease
HERMES	Heart failure Molecular Epidemiology for Therapeutic targets
GWAS	Genome wide associated studies
AI	Artificial intelligence
ROC	Receiver operating characteristic
ANN	Artificial neural network
CNN	Convolutional neural networks
HCM	Hypertrophic cardiomyopathy
GPS	Global positioning system
PPG	Photoplethysmography
LED	Light emitting diode
NICE	National Institute for Health and Care Excellence
RATE AF	RAte control therapy evaluation in permanent Atrial Fibrillation
IPAQ	International Physical Activity Questionnaire
TIA	Transient ischaemic attack
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PPV	Positive predictive value

NPV	Negative predictive value
SROC	Summary receiver operating characteristic
CRMA	Cardiio rhythm mobile application
PROBE	Prospective, randomised open label blinded endpoint
PPI	Patient and public involvement
SF-36	Short form-36
AFEQT	Atrial fibrillation Effect on Quality of Life
EQ-5D-5L	European Quality of Life Five Dimension
6MW	6-minute walk
PIL	Patient information leaflet
BCTU	Birmingham Clinical trials Unit
SD	Standard deviation
RADAR	Remote Assessment of Disease and Relapses
FPS	Frames per second
CATCH ME	Characterizing Atrial fibrillation by Translating its Causes into Health Modifiers in the Elderly
RADAR-CNS	Remote Assessment of Disease and Relapse – Central Nervous System
pRMT	Passive remote monitoring technologies
aRMT	Active remote monitoring technologies
AWS	Amazon Web Service
RECap	Research Electronic Data Capture
IQR	Inter quartile range
GLS	Generalised least squares regression method
6MWD	6-minute walk distance
SEM	Standard error of the mean
eGFR	Estimated glomerular filtration rate
IPD	Individual patient data

CIBS I	Cardiac Insufficiency Bisoprolol Study
HR	Hazard ratio
CI	Confidence interval
OR	Odds ratio
ARI	Adjusted Rand Index
CONSORT-AI	Consolidated standards of reporting trials-artificial intelligence
FAIR	Findability, Accessibility, Interoperability and Reusability
GDPR	General Data Protection Regulation

1. Introduction

Over the past 50 years, cardiovascular research has provided significant insights into our understanding of disease pathophysiology, leading to the development of treatments that have improved prognosis. Despite this progress, common cardiac conditions such as atrial fibrillation (AF) and heart failure (HF) continue to exhibit high rates of mortality and re-hospitalisation, alongside signs and symptoms that adversely impact quality of life. In recent years it has been recognised that disease definitions are outdated and do not account for changes at the patient level, as well as the effect of underlying aetiologies, co-morbidities and varied prognoses. Additionally, current clinical practice still reflects the findings of landmark research trials where therapy is guided by conventional clinical markers and trial end points. (1, 2) There is a clear need to develop more targeted and personalised treatments in healthcare.

Over the last decade the digitalisation of information routinely collected during healthcare visits has led to a substantial increase in the amount of data available in healthcare, giving rise to the term 'big data'. This coupled with advances in data management and machine learning techniques for analysis has led to new opportunities to support a personalised approach to medicine, that has the potential to improve the delivery of healthcare, both in terms of patient outcomes and cost effectiveness.

1.1 Atrial Fibrillation

1.1.1 Definition of atrial fibrillation

AF is a supraventricular arrhythmia that occurs when abnormal electrical activity in the sino-atrial node, resulting in uncoordinated atrial contraction. This causes a reduction in stroke volume due to loss of atrial kick.(3) At present the diagnosis of AF requires a 12-lead electrocardiogram (ECG), or a single lead ECG of at least 30 seconds duration, with no obvious P waves and irregular R to R intervals.(4) The clinical presentation of AF is variable and can range from entirely asymptomatic individuals to those with disabling symptoms. Due to its effect on haemodynamics and risk of thromboembolic complications such as stroke, AF can result in significant morbidity, mortality and can impair patient quality of life.(5, 6)

1.1.2 Epidemiology and risk factors for Atrial fibrillation

Worldwide, AF is the commonest rhythm disturbance encountered by clinicians. Increasing age is the strongest risk factor; less than 1% of individuals aged under 60 have AF, this figure rises sharply to between 10% and 17% in those aged 80 years and above.(7) The current prevalence of AF is between 2% and 4%, and the lifetime risk for developing it in those of European descent (at age 55) is 1 in 3. The number of individuals with AF is expected to rise significantly in the foreseeable future due to improved treatment and mortality outcomes from co-morbidities. By 2060 the prevalence of AF is predicted to reach 14 million in Europe.(8, 9)

There is an array of predisposing risk factors that can lead to AF. Increasing age and male gender are key drivers, as is the presence of certain co-morbidities including heart failure (HF), hypertension, coronary artery disease (CAD), valvular heart disease, diabetes mellitus,

chronic kidney disease, thyroid disorders, obstructive sleep apnoea and obesity.(5, 10-13) In addition, the presence of protein biomarkers such as a raised N-terminal Pro B-type Natriuretic peptide (NT-proBNP) and decreased Insulin-like growth factor 1 (IGF1) are associated with the development of AF. (14)

1.1.3 Clinical presentation and management of atrial fibrillation

AF is usually classified by time-dependent duration. Paroxysmal AF is usually self-limiting and either spontaneously reverts to sinus rhythm (usually within the first 48 hours) or with some form of intervention such as cardioversion within 7 days. AF that persists beyond 7 days and requires termination through cardioversion (pharmacological or electrical) is classified as *persistent*. The classification changes to *longstanding persistent AF* when it has been present continuously for over 12 months and a rhythm control strategy is still adopted. AF is considered *permanent* when a joint decision by patient and physician is made to no longer pursue attempts to restore sinus rhythm.(15-17) The presentation of AF is variable, with up to two-thirds of patients describing some form of symptoms that affect their day-to-day life. Other common symptoms include; palpitations, dyspnoea, chest pain, dizziness, light headedness, syncope, sleep disturbance and reduced effort tolerance.(17-19) AF symptom severity can be characterised by clinicians using the modified European Heart Rhythm Association (EHRA) symptom scale, which uses six symptoms (palpitations, fatigue, dizziness, chest pain, dyspnoea and anxiety) to assess the effect they have on daily life (see **table 1**). (20)

Table 1: EHRA symptom score assessing effect of AF on daily life

Score	Symptoms	Description
1	None	AF causes no symptoms
2a	Mild	Activities of daily living are not affected by symptoms of AF
2b	Moderate	Activities of daily living are not affected by symptoms of AF, but patient troubled by AF symptoms
3	Severe	Activities of daily living are affected by symptoms of AF
4	Disabling	Normal daily activity discontinued due to symptoms of AF

Adapted from Wynn G. et al, 2021

AF increases the chance of suffering a thromboembolic stroke by up to five-fold. The risk increases with age and the presence of certain risk factors(21). There are a multitude of risk scoring tools that have been developed to predict the risk of thromboembolic stroke in AF.(22, 23) Examples of these scoring systems and the year they were first published are shown below in **table 2.** Overall, each of these scoring systems has a modest ability to predict stroke. Hence, there use can result in overprediction or underprediction of stroke and thromboembolism. Newer scoring systems have been developed but are less commonly used in research and lack external validation.(24, 25)

Table 2: Risk score for predicting ischaemic stroke in AF

Year of publication	Name of risk score	Risk factors included and points scored for each	Anticoagulation
2001	$CHADS_2$ score	Age ≥75(1), hypertension (1), diabetes mellitus (1), CHF (1), TIA or stroke (2).	Recommended if score >2
2010	CHA_2DS_2-VASc	CHF (1), hypertension (1), Age ≥ 75 (2), diabetes mellitus (1), stroke/TIA/ TE history (2); vascular disease (1); Age 65-74 (1); Female (1)	Low risk if 0 (men) or 1 (women) Low-moderate risk if 1 (men) or 2 (women), consider anticoagulation. Moderate-high risk if ≥2 (men) or ≥3 (women) anticoagulation recommended
2012	R_2CHADS_2/ $R_2CHA_2DS_2$-VASc	Above scores plus CrCl<60mL/min and GFR <60ml/min/1.73m^2 (2)	Recommended if score >2
2013	ATRIA	Stroke, age, (<65, 65-74, 75-84, ≥85), female, diabetes mellitus, CHF, hypertension, GFR <45ml/min/1.73m^2 or end stage renal disease	High risk of stroke if Score >6
2016	ABC stroke risk score	Age, Biomarkers (TNI and NT-proBNP,) and prior stroke or TIA	Categorised as low medium and high risk

ATRIA=Anticoagulation and Risk factors in atrial fibrillation; ABC=Age, Biomarkers, Clinical history; CHF= congestive heart failure; TIA=transient ischaemic stroke; TE=thromboembolism; CrCl=creatinine clearance; GFR=glomerular filtration rate; TNI= troponin; NT-proBNP=N-terminal B-type natriuretic peptide

Currently, a 3-step approach is advised when considering anticoagulation. Step 1 involves using the CHA_2DS_2-VASc to identify those patients who are low risk (score of 0 in men and 1 in females) and do not require anticoagulation. The second step involves identifying those with at least 1 risk factor (non-gender) for stroke (score of 1 in men and 2 in females), in

these patients' anticoagulation should be considered (class IIa evidence) and any bleeding risk factors should be addressed. In CHA_2DS_2-VASc scores ≥ 2 (men) and ≥ 3 (women), anticoagulation is recommended for stroke prevention (Class 1A). Direct oral anticoagulants are recommended as first line therapy for anticoagulation.(16, 26)

The management of AF consists of anticoagulant therapy if indicated as described above, alongside rate and rhythm control strategies. Rate control therapy is commenced to slow down the ventricular rate and improve associated symptoms. If symptoms remain troublesome despite adequate rate control, a rhythm control strategy using anti-arrhythmic pharmacological therapy, electrical cardioversion, catheter or surgical ablation can be employed to restore sinus rhythm.(16) The recent EAST (Early Rhythm-Control Therapy in Patients with Atrial Fibrillation) trial has shown that rhythm control can be beneficial if initiated promptly following the diagnosis of AF. This trial compared the use of a rhythm control strategy using anti-arrhythmic drugs or catheter ablation in patients with early AF that was present with concurrent cardiac conditions, to standard rate control therapy. The early rhythm control treatment group was found to be superior and was associated with a reduction in cardiovascular death, stroke and hospitalization.(27)

As a cardiologist a recognised real-world limitation of the rhythm control strategy is access to a limited service. Ablation services are specialised and often have long waiting lists. As a result, rate control strategy may be used in its place.

1.2 Heart failure

1.2.1 Definition and pathophysiology of heart failure

HF is a syndrome that occurs when the heart cannot meet the metabolic demands of the patient. Cardiac dysfunction gives rise to a group of symptoms that include fatigue, ankle swelling and breathlessness (dyspnoea).(28) The underlying aetiology of HF is heterogenous and often due to abnormalities that arise from structures within the heart. HF can also occur from conduction disease and rhythm disturbances (brady/tachyarrhythmias) such as AF. In terms of biomarkers elevated NT-proBNP levels are associated with HF. An elevated troponin can be used as a prognostic marker.(29) Identifying the underlying cause of HF is essential to ensure appropriate and timely therapies are commenced.(28, 30)

HF is categorised into phenotypes based on left ventricular ejection fraction (LVEF), which is measured on echocardiography. Heart failure with reduced ejection fraction (HFrEF) is defined as an LVEF < 40%. Patients with an LVEF of 40%-49% are considered to have heart failure with mildly reduced ejection fraction (HFmrEF). Heart failure with preserved ejection fraction (HFpEF) is the term used to define patients with signs and symptoms of HF, an LVEF ≥50% in the presence of diastolic dysfunction or raised filling pressures.(30, 31)

1.2.2 Epidemiology and risk factors

The estimated prevalence of HF is thought to lie between 1% and 2% and increases with age from <1% in those aged under 55 years, to >10% in those aged 70 years and above.(32, 33) The commonest causes of HF in developed countries are hypertension and CAD. HFrEF and HFmEF are usually more likely to be caused by ischaemic pathology, whereas the aetiology of HFpEF is more complex and is often associated with multimorbidity including hypertension, diabetes mellitus, renal impairment, and obesity.(28, 34)

1.2.3 Clinical presentation and management

Typical symptoms of heart failure consist of breathlessness, orthopnoea, paroxysmal nocturnal dyspnoea, fatigue, ankle swelling, and reduced exercise tolerance. The severity of symptoms is assessed using the New York Heart Association (NHYA) functional classification (see **table 3**). (35) In the presence of symptoms, a diagnosis of HF is more likely in patients with underlying risk factors, abnormal ECG, imaging evidence of cardiac dysfunction and an elevated N-terminal B-type natriuretic peptide (NT-proBNP).(31)

Table 3: New York Heart Association functional classification

Class	Patient symptoms
Class 1	Ordinary physical activity does not cause symptoms and is not limited
Class 2	Ordinary physical activity causes symptoms, and there is slight limitation to physical activity, comfortable at rest
Class 3	Less than ordinary physical activity causes symptoms, and there is marked limitation of physical activity, comfortable at rest
Class 4	Symptoms are present at rest, unable to carry on any physical activity without discomfort

Adapted from Zhang R. et al. 2018

Treatment options in HFrEF aim to reduce mortality, prevent hospitalizations, and improve functional status and quality of life. Pharmacological therapy that reduces mortality include:

- Beta-blockers(36, 37)
- Angiotensin receptor blocker (ARB)/ Angiotensin converting enzyme inhibitor (ACEI) or angiotensin receptor-neprilysin inhibitor (ARNI)(38-40)
- Mineralocorticoid receptor antagonist (MRA)(41, 42)

- Sodium glucose co-transporter (SGLT2) inhibitor(43)

Once optimal medical treatment is commenced the response is monitored for a period of 3-6 months after which the patient is reviewed with repeat imaging to assess LVEF and symptoms. A cardiac synchronization device (CRT) is recommended if patients remain symptomatic from their HF, with an LVEF of ≤ 35%, with the class of evidence indication dependent of the QRS duration and morphology:

- QRS duration ≥150ms with left bundle branch block (LBBB) morphology (class 1A)
- QRS duration 130-149ms with LBBB morphology (class IIa)
- QRS duration ≥150ms with non LBBB morphology (class IIa)
- QRS duration 130-149ms with non LBBB morphology (class IIb) (44-46)

The choice between a cardiac resynchronisation therapy pacemaker (CRT-P)) or cardiac resynchronisation therapy defibrillator (CRT-D) is dependent on several patient factors such as age, level of scarring seen on cardiac magnetic resonance imaging (MRI), type of cardiomyopathy, life expectancy, co-morbidities, renal impairment and patient preference. Ultimately the choice made is through a shared decision-making process between the patient and clinician.(44, 47-49)

In HFpEF, treatment options were limited to diuretics and treatment of underlying causes until recently where a large trial showed SGLT2 inhibitors conferred a benefit in terms of reduction in death and hospitalization rates (hazard ratio (HR), 0.79; 95% confidence interval [CI], 0.69 to 0.90; P<0.001).(50) A meta-analysis of 12,251 participants from 2 large clinical trials (ejection fraction over 40 %) confirmed SGLT2 inhibitors reduced mortality and HF admissions (HR 0.80 [95% CI 0.73–0.87]).(51) As such they are now a class 2a

recommendation in the American Heart Association guidelines for the management of HFpEF.(52)

1.3 Atrial fibrillation and heart failure

1.3.1: Epidemiology of atrial fibrillation and heart failure

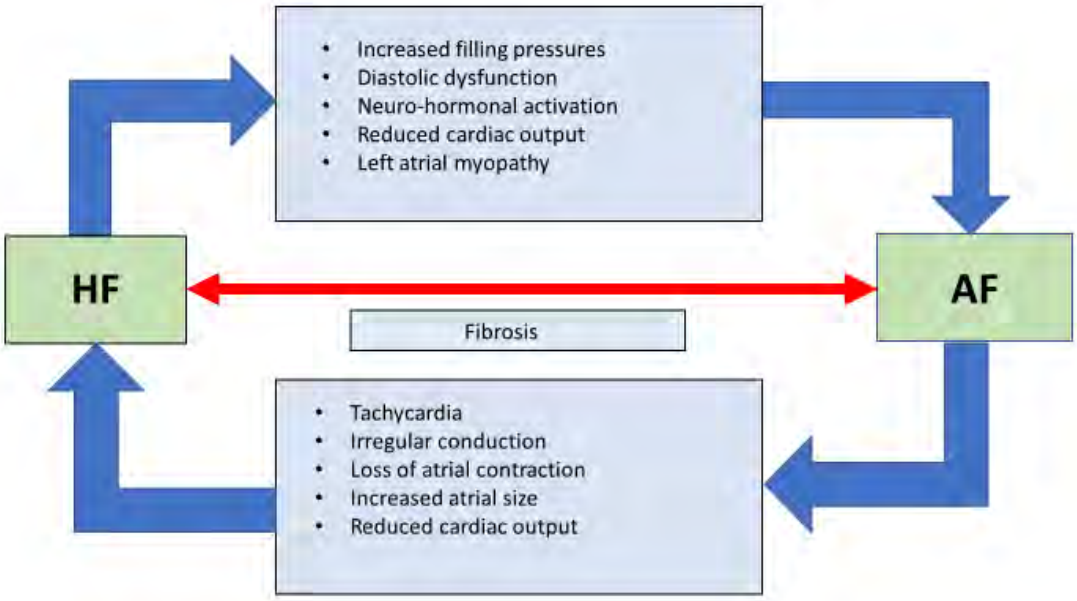
Although AF and HF can occur independently, they are commonly seen together. Their prevalence is rapidly increasing and poses a major public health concern in terms of healthcare delivery and rising costs. Previous findings have shown that over 30% of individuals with HFrEF will develop AF and this increases up to 50% in those with advanced HFrEF and HFpEF.(53) Equally, over 30% of those with a new diagnosis of AF will have HF. These figures are likely to be underestimates as they only account for those with established diagnoses and not for future AF and HF occurrences.(54) Both AF and HF can cause and sustain the other. In AF, the loss of atrial contraction and tachycardia can impair cardiac function causing systolic and diastolic dysfunction. In HF, atrial fibrosis, re-modelling and neurohormonal activation can trigger AF. When both conditions co-exist, patient prognoses are usually worse with higher rates of hospital admissions and a two to three-fold increase in mortality.(54-57)

1.3.2 Pathophysiology of atrial fibrillation and heart failure

AF and HF share many pre-disposing risk factors and both conditions are often present in elderly patients with multiple cardiovascular co-morbidities such as; diabetes, hypertension, obesity, sleep apnoea, and ischaemic heart disease.(58) These risk factors either cause structural changes or create a pro-inflammatory environment in which patients are more susceptible to developing both conditions.(16) The close relationship between the two cannot be explained by shared risk factors alone and is most likely due to common

pathophysiological processes. AF promotes the development of HF in several ways. Firstly, atrial contraction contributes up to 25% of the stroke volume. The loss of this atrial kick is seen in AF and can lead to HF, especially if diastolic dysfunction is present.(59) Tachycardia is often seen in uncontrolled AF. If left untreated it can lead to abnormal calcium signalling which in turn will reduce myocardial contractility; reduce stroke volume, and cause LV dilatation. This results in a condition called tachycardia induced cardiomyopathy.(57) Even when heart rate is well controlled, irregular contractions of the ventricle can impair left ventricular filling leading to the development of HF.(57, 60) These mechanisms all trigger neuro-hormonal activation of the renin-angiotensin-aldosterone system (RAAS). Over time persistently elevated hormones can cause adverse re-modelling of the atria and ventricles, leading to fibrosis and thereby increase the likelihood of developing HF. (54-57) In HF, the elevated atrial filling pressures act as a precursor for the development of AF. Raised left sided filling pressures trigger a sequela of changes within the atrium that include atrial dilatation, increased wall stress and scarring leading to fibrosis. Activation of neuro-hormonal pathways results in further adverse structural re-modelling in the atria causing fibrosis. Fibrotic tissue then acts as a driver for arrhythmias and facilitates the development and progression of AF (see **figure 1**).(54-57)

Figure 1: Shared pathophysiological mechanisms in heart failure and atrial fibrillation



Adapted from Kotecha D. et al, 2015 (55) and Anter E. et al. (2009) (61) AF=atrial fibrillation; HF= heart failure

1.3.3 Current management strategies in atrial fibrillation and heart failure

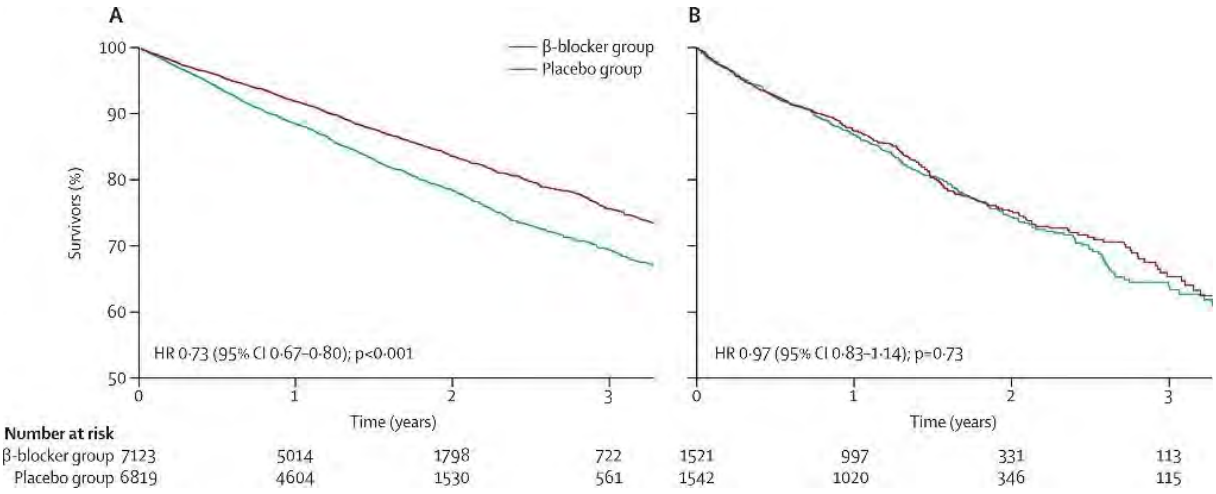
Evidence based treatment strategies for the management of patients with concomitant AF and HF is limited, patients are usually trialled on established mono disease therapies for either disease and exhibit a poorer treatment response.(55)

Management of atrial fibrillation when present with HFrEF

General suggestions for managing the AF component include treating co-morbidities such as obstructive sleep apnoea; managing cardiovascular risk factors; and lifestyle advice including weight loss. Rate control is often necessary and the main drug therapies prescribed are beta-blockers, digoxin and non-dihydropyridine calcium channel blockers. Beta blockers are an established independent treatment option in both HF and AF, and one would expect it to have a significant benefit in patients with concomitant disease.(62) An individual patient data

analysis of 11 randomised clinical trials (RCTs) in HF showed a difference in efficacy of beta blocker therapy according to baseline heart rhythm. Beta blockers conferred a significant mortality benefit in patients with sinus rhythm and HFrEF (HR 0.73; 95% CI 0.67-0.80; p<0.001), but this was not seen in individuals with AF and HFrEF (HR 0.97; 95% CI 0.83-1.14; p=0.73; see **figure 2**).(62-64) A further meta-analysis by the beta blocker collaborative group showed beta blockers reduced mortality, improved LVEF and overall prognosis in all patients in sinus rhythm with reduced LVEF. The improvement in ejection fraction was also seen in those with AF but the analysis was unable to show a clear benefit in morbidity and mortality.(65)

Figure 2: Effect of treatment in patients with HF and reduced ejection fraction from double blind randomised trials



Reproduced with permission from Kotecha D. et al, 2014. (62)
Kaplan-Meier survival curves for patients with heart failure and a reduced ejection fraction on beta blocker treatment vs. placebo in (A) sinus rhythm; (B) atrial fibrillation. HR=hazard ratio, CI=confidence interval.

With regards to the use of digoxin, a systematic review found mortality was not higher in patients with AF and HF, and confounding factors were likely to be the cause of the higher rates of death found in earlier observational studies.(64) Rate limiting calcium channel

blockers such as diltiazem or verapamil are not indicated in those with AF and HFrEF due to their negative inotropic effect.(66, 67)

If symptoms persist despite optimal rate control a rhythm control strategy can be employed using anti-arrhythmics or ablation (pulmonary vein isolation or radiofrequency ablation). Previous studies comparing rate versus rhythm control using anti-arrhythmic drug therapies in this group did not demonstrate superiority of one strategy over the other.(68) However, a meta-analysis by Chen et al comparing catheter ablation with medical therapy (n=3598) from 11 RCTs with AF and HF, found catheter ablation was associated with lower mortality (odds ratio: 0.51; p-value=0.0003), improved cardiac function and quality of life.(69) In addition, a prespecified analysis of AF and HF subgroups in the EAST trial demonstrated that early rhythm control therapy in patients diagnosed with AF alongside signs and symptoms of HF can reduce the rate of mortality, stroke, and hospital admissions for HF or acute coronary syndrome, compared to standard rate control therapy.(70) At present, a rhythm control strategy is used to in patients with AF symptoms causing morbidity and reducing QoL.(16) In patients with permanent AF and HF where catheter ablation is unsuitable or has failed, atrioventricular junction ablation can be combined with implantation of a CRT device if indicated.(57)

Management of HFrEF when present with atrial fibrillation

In HFrEF the benefits of ACEI, ARB, ARNI and beta blockers are well described. The CHARM (Candesartan in heart failure assessment of reduction in mortality and morbidity) study of candesartan versus placebo showed a similar reduction in cardiovascular death and HF hospitalisation for HFrEF patients with AF (HR 0.83; 95% CI 0.69-0.99) and without AF (HR 0.84; 95% 0.77-0.92).(71) MRA and SGLT2 inhibitors are often prescribed in

symptomatic HFrEF patients where LVEF is severely impaired, but information regarding use in patients with concomitant AF is limited.(42, 55, 72) Despite the lack of robust data on the efficacy of these known HFrEF therapies in AF, they are still recommended to reduce mortality. Kotecha et al suggested an initial plan for patients with new onset HFrEF and AF. CAN-TREAT stands for: electrical cardioversion (C) in the presence of haemodynamic instability; anticoagulation (A) for thromboprophylaxis as the risk of stroke is doubled if patients with AF and HF; diuretics to normalise (N) fluid status; rate control to a target (T) heart rate; commencing RAAS antagonist (R); Early (E) rhythm control if rate control does not alleviate symptoms and consider advanced (A) device therapies; aggressive (T) treatment of cardiovascular co-morbidities.(55)

Management of AF with HFpEF

The AF component should be managed as described above in AF with HFrEF. With regards to HFpEF, as mentioned earlier the only treatment shown to be of benefit is SGLT2 inhibitors. Interestingly, data from the EMPEROR-PRESERVED trial showed reduction in mortality and HF hospitalisation was seen in those with HFpEF and AF at baseline (HR 0.78; 95% [0.66–0.93]).(73, 74)

To summarise, the optimal treatment therapies for patients with both AF and HF remains unclear due to the lack of RCTs in this specific population. Currently we lack robust data in this cohort of patients, with limited information on underlying molecular mechanisms, diagnosis and the impact of medical and interventional therapy. Data from trial outcomes on treatment efficacy do not take into account the impact of co-morbidities, and assumes all individuals have a similar treatment response. A different approach to research is required that can help us re-define this group of patients (outside of the scope of this thesis). The use

of more patient data (big data) and advanced data analytics (machine learning) may enable us to find better ways to classify or sub-categorise AF and HF patients such that new or repurposed therapies may be found that can improve outcomes and quality of life.

1.4 Big data and cardiovascular disease

1.4.1 Definition of big data

Big data can be described as a dataset of information that is too large for storage in traditional database repositories, or too complex for processing and analysis by conventional software algorithms. Big data includes standardised data from clinical registries and electronic healthcare records (EHRs), as well as data from smartphone applications and real-time data from wearable technology.(1, 75, 76) Big data is characterised by the five "V's" these are: volume, variety, velocity, veracity and value. High volume refers to the vast amount of information that is created and stored. This has grown exponentially over the last few years mostly due to the lowering costs of storing data.(77) Variety refers to the complexity of data and its different forms which includes; structured, semi structured, and unstructured data. Structured refers to data that is easily stored and organised for example, demographics, blood values, biomarkers and disease classification using the standardised coding systems. This type of data is usually easy to process and analyse. Semi-structured tends to be less organised and can include text-based or continuously measured time series data from wearable devices. Unstructured accounts for the majority of currently available healthcare data and includes textural information such as clinical notes, discharge summaries, and different forms of imaging such as ECG, echocardiography, CT (computer tomography), MRI and X-ray.(1, 2, 77) Velocity refers to the speed at which new data are created, and the need to store and analyse in real-time. Veracity looks at the level of uncertainty within data

that occurs due to inconsistences, missingness and ambiguities. Value refers to the additional worth the data can add to what is already known.

These specific characteristics of big data have limited the use of traditional statistical techniques for analysis. At present, most research is hypothesis-based where the impact of a condition is evaluated using a set number of variables and conventional clinical end points. The construction of platforms for big data collection, storage, and the use of machine learning (ML) methods for analysis has enabled the opportunity for data driven discovery. Data driven research has the potential to identify new or previously unknown patient phenotypes, improve diagnostic capability, and deliver a more individualised approach to patient care (2, 78-80) With this in mind the design and development of big data projects should meet a minimum standards framework to ensure robust data construction and linkage, clearly defined disease and outcome definitions, appropriate ML analysis and subsequent reporting of studies for publication.

1.4.2 The need for big data in cardiovascular disease

Over the last few decades significant advances have been made in the treatment, prevention and prognostic evaluation of patients with cardiovascular disease (CVD).(81) However, it continues to have the highest rates of disability and premature deaths worldwide. A 31% global rise in mortality from CVD was seen from the year 1990 to 2010, despite a 21% reduction in age-standardised mortality rates in this time frame.(82, 83) The management of CVD is further complicated by multi-morbidity.(83, 84) The lack of progress in reducing death rates, hospitalisation, and improving patient quality of life is likely to be multi-factorial. Poor use of available data, a lack of understanding around how different patient groups behave within a disease, alongside our current research practices are all contributors to

hindering progress. The increase in our ability to collect and store data has resulted in more datasets being available for cardiovascular research however, they remain under-utilised due to lack of integration and because many were not originally collected for research purposes. The underlying causes of common cardiac conditions is a complex interplay of genetic, environmental and behavioural factors. Current definitions are not based on underlying molecular mechanisms, instead they are described as syndromes due to a collection of aetiologies and differing prognoses. This results in unpredictable treatment responses due to large inter-individual variation and presents a challenge in terms of developing new targeted therapies that produce better clinical outcomes with fewer adverse effects. (1, 2) Big data analysis using enriched datasets has the potential to help us understand the complexity of multi-morbid patient groups and can aid better disease definitions.(83)

Drug development from early target discovery through to clinical trials and post-marketing is a lengthy and expensive process. Phase I cardiovascular drugs have one of the lowest success rates in the field of medicine with only 7% reaching eventual approval. (85) Drug failures that occur in the latter phases (e.g., phase III) of drug development are often due to the lack of end point optimisation. The use of traditional end points such as cardiovascular death does not consider the knowledge we have acquired regarding the development and progression of disease. (2, 86) Big data analysis has the potential to find better trial end points and drug targets, thereby streamlining the process so that clinical trials can be planned and conducted in a more efficient and cost-effective manner.

1.4.3 Sources of data

There are a multitude of data sources available for healthcare research including: administrative databases, disease/population-based registries, clinical trials including

secondary/pooled analysis, EHR, imaging and sensor data from wearable technology. **Table 4** summarises the strengths and limitations of common data sources used for big data applications. Real-world data sets such as EHRs offer volume, whereas data collected purely for research is often on a more selected population, and so tends to be of a higher quality.(83) With regards to CVD, a diverse amount of data is collected during routine patient care. Clinical visits, inpatient episodes, blood laboratory values and imaging data all make up patient EHRs. Wearable technologies offer the opportunity for continuous real-time information in patients with CVD. Several devices are currently available these include single lead ECG devices for arrhythmia detection, wearable vests, and wrist-worn devices for heart rate and physical activity monitoring. These will be reviewed in more detail in a later section.(1, 2)

Table 4: The strengths and limitations of data sources used for big data applications

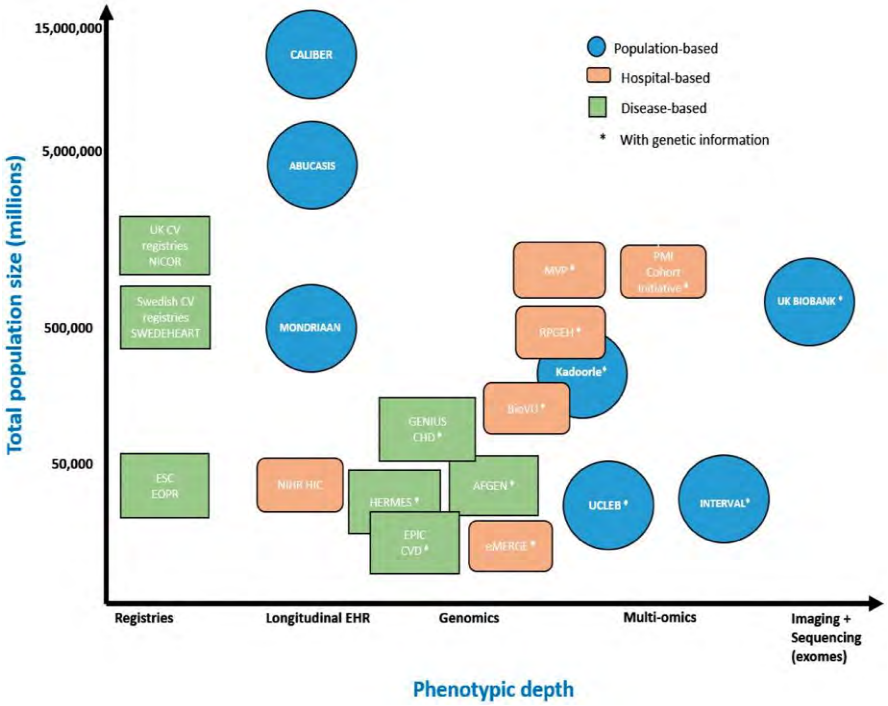
Data source and type	Strengths	Limitations
Administrative data Structured	Standardised ICD codes used which results in consistent data that is readily available	Risk of inaccurate coding Time lag (data usually completed after episode) Lacks granularity
Biomarkers (including omics) Structured	Provides additional detailed information on patient characteristics that can be used to personalise medicine	Issues surrounding false positives and trying to find valid associations between biomarkers and patient outcomes Not readily available currently
Wearable data Semi-structured	Physiological data such as heart rate and physical activity collected outside of the routine clinical care	Not widely available in clinical practice Requires ability to store and process large amounts of real-time data. Difficulty interpreting relevance due to redundancy
Clinical registry Structured	Data collected is consistent, established registries in CVD (NICOR, SWEDEHEART)	Data collection may be limited to specific procedures or diseases Time lag
Electronic health records Structured and unstructured	Diverse clinical dataset	Missing data, variable quality and data types
Medical imaging Unstructured	Opportunity for dynamic image interpretation	Difficult to correlate data from different modalities and sites

Modified from Rumsfield J. et al, 2016(1)
CVD - cardiovascular disease, ICD – international classification of diseases, NICOR – the national institute for cardiovascular outcomes research, SWEDEHEART – Swedish web-system for enhancement and development of evidence-based care in heart disease.

According to Hemingway et al the type of data available is a "trade-off between scale and depth", whereby scale denotes the number of subjects, and depth the richness or phenotypic and omics information available.(2) **Figure 3** demonstrates the volume versus detail that is provided for each type of data. This shows how registries are able to offer volume in terms of patient numbers, but the variety or depth of information provided is much less.

Conversely EHRs can provide the greatest amount of phenotypic information, as they consist of both structured and unstructured data. Additionally, molecular profiling techniques can provide high-dimensional molecular information such as genetics (genomics), specific proteins (proteomics), and metabolites (metabolomics). This information can be linked to conventional data to give an enriched dataset. The application of ML approaches to these enriched datasets can help identify genetic variations in a disease and responses to treatment (genomics), while the presence of certain proteins or metabolites can provide biomarkers for the development of certain conditions, as well as its clinical trajectory.(2, 80, 87) An example of this is the work of the HERMES (Heart failure Molecular Epidemiology for Therapeutic targets) consortium where genome wide genotyping data has been collected from participants of 51 studies (n=68,157 with heart failure; n=949 888 controls).(88) The aim of the consortium is to conduct genome wide associated studies (GWAS) of HF to identify the genetic and molecular basis of HF.(88) Shal et al performed a meta-analysis on 11 GWAS (n=47,309 with HF; 930,014 controls) and identified 12 independent variants at 11 loci that are associated with HF. Each had at least one association with common cardiovascular conditions including; CAD, AF, and reduced left ventricular function, suggesting shared genetic aetiology.(89)

Figure 3: Volume versus depth for different data types



Reproduced with permission from Hemingway H. et al, 2018.(2)

1.4.4 Big data analytics using machine learning

Artificial intelligence (AI) is a process of iterative learning by a computer algorithm that stimulates or betters human intelligence.(83) ML is a sub-category of AI that uses algorithms trained on data to learn.(90) In healthcare, computational ML algorithms are trained on large datasets to identify patterns and associations that can inform prediction and diagnosis without the need for pre-programming.

(80, 91, 92) Conventional statistical approaches such as logistic regression and multi variate analysis can inform prediction and discovering relationships between variables, but they are still limited by several factors. Firstly, traditional statistics can only look at a relationship within a small number of variables, and assumptions about the distribution of data are made. An example of this is seen in the meta-analysis of beta efficacy in HFrEF patients with renal

dysfunction (see chapter 6.2). Here regression analysis was only able to detect a weak interaction between beta-blocker efficacy and renal function.(93) Secondly, findings from this sample of data are used to accept or reject a hypothesis and this outcome is inferred to the whole population, which limits its real-world application. Thirdly, traditional statistics are unable to handle the processing and number of interacting variables found in complex datasets that contain various types of data ML approaches can overcome these limitations by handling large heterogenous real-world datasets using fewer assumptions.(80, 91) There are several types of ML approaches, and these are described next in more detail along with their limitations.

1.4.5 Types of Machine learning

Supervised learning

Supervised learning is a ML approach where a model is trained on labelled input data that is associated with a known outcome (output). Once trained this algorithm can then be applied to new data to inform prediction of future events, or classification. Types of supervised learning algorithms commonly used include ordinary least squares regression, support vector machines, and decision tress, such as Random Forests.(80, 91) Each algorithm has advantages and disadvantages and for optimal accuracy the correct choice will depend on the type of data and its characteristics, training time, number of parameters and the hypotheses in question.(75) Regression is a simplest form of ML, whereby the relationship between input variables and an outcome variable is linear or non-linear. An example of the use of regression in ML is predicting fractional flow reserve values from cardiac CT angiography images.(90) The use of classification algorithms has been used to differentiate phenotypically difficult patients with physiological hypertrophy (athlete's heart) and pathological hypertrophy (hypertrophic cardiomyopathy).(94) Support vector machines are

similar to regression algorithms except they are capable of handling more complex data and are less vulnerable to outliers Decision trees are a flow chart like hierarchical structure and random forest algorithms apply multiple decision trees to a range of variables to inform prediction or classification.(90)

Despite its ability to accurately predict the label of interest there are limitations. Supervised learning may not find previously unknown patterns and associations within data. Data labelling is time-consuming and can act as a constraint reducing the chances of finding new patterns. If the input label is incorrectly labelled due to errors in coding, then outcome predictions or classifications will also be incorrect. Supervised learning needs a large training dataset to train a model and avoid inaccurate predictions and the model should ideally be validated on new datasets.(75, 90)

Unsupervised learning

Unsupervised learning is a ML approach that learns patterns in raw unlabelled data. The ML algorithm is free to learn the relationship between variables within the dataset and to identify any hidden patterns that may lie within it. Examples of unsupervised learning algorithms include clustering (such as K means) and principal component analysis. Clustering organises data into clusters which are smaller groups containing similar characteristics. The number of clusters can be pre-determined or automated using software to find the ideal number.(90) Principle component analysis can be applied when each observation contains a high number of features. This technique reduces the dimensionality of the dataset and enables multi-dimensional data to be visualised whilst preserving most of the original data. Unsupervised learning is commonly used for exploratory analyses and to identify novel phenotypes.(80, 95)

A limitation of unsupervised learning is that results can be inaccurate and difficult to understand and verify findings as the dataset is unlabelled.(80) Certain types of data require both forms of learning to improve predictions. For example, imaging data can be noisy and requires supervised learning initially to de-noise data so that unsupervised ML algorithms can then be applied to the de-noised data to improve accuracy of prediction.(75)

Evaluating and validating model performance

Validation of the ML model is required to ensure accuracy and that the model is achieving its purpose. Different methods are available for validation with the simplest being the train-test split where prior to implementing ML algorithms, the data is usually split into 3 sets: training, validation and test data. The training dataset usually consists of the largest amount of data where ML algorithms are applied to the data to create and fit the model. The validation dataset is used to evaluate the model fit and allows for the refining of parameters. The test dataset is used to validate the final model performance either internally (on existing dataset) or externally (on new dataset). A typical split of data can be 60% for training, 20% for validation, and 20% for testing. With ML it is important to ensure the model learns sufficiently on the training dataset using the correct ML algorithms to ensure optimal model performance, and to avoid the risk of the model underfitting or overfitting the data.(75) The performance of ML algorithms is evaluated using performance statistics such as F1 score and receiver operating characteristic (ROC), these are discussed in more detail later on.

1.4.6 Advanced machine learning techniques for identifying new phenotypes

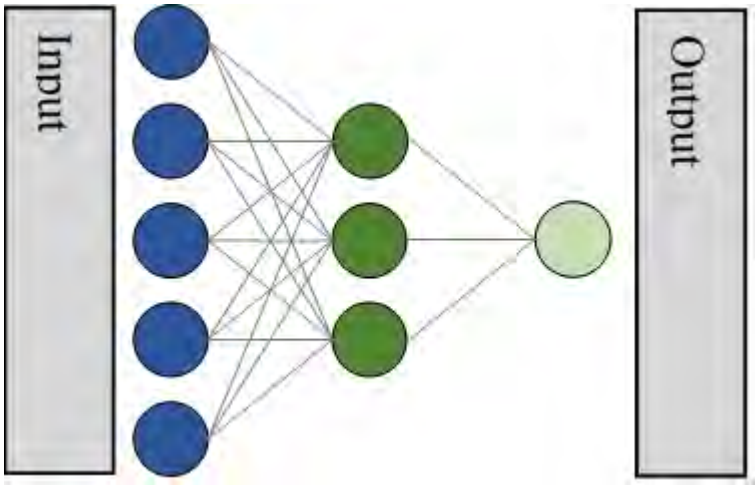
Deep learning

Deep learning is a form of is a ML that 'mimics' the concept of the human neuron using neural networks to create multiple layers that can learn rapidly across large amounts of

data.(75, 80) (75) Deep learning can be applied to supervised and unsupervised learning using models that consist of a neural network known as artificial neural network (ANN). ANN are made up of "nodes" that are interconnected, and each node performs a function. The input nodes receive information from different types of data in preparation for processing. The hidden nodes are responsible for processing this information during training and adjusting weights depending on if a decision is correct or incorrect. Lastly, the output nodes provide the final interpretation.(80, 90, 92) **Figure 4** below provides a simplified overview of a neural network. Convolutional neural networks (CNN) are a type of ANN that consists of multiple layers of hidden nodes that are used for processing large amount of information effectively. They are used in research to assess high volume time series data from wearable technology and for complex imaging analysis.(96)

Deep learning has several challenges by virtue of its analysis. Setting up the learning algorithms is time consuming and often requires costly high-performance hardware. Although deep learning has the ability to learn complex features and generalises well on unseen data, there is a risk of overfitting (model fits data too well) which reduces model performance.(97) There are several methods that can be applied to reduce this which include: using a larger dataset, data augmentation to increase sample number available for training, cross-validation to split data and assess model performance on each set. Lastly in terms of the deep learning architecture, dropping the number of layers can reduce the risk of overfitting.(75, 97)

Figure 4: Simplified overview of a neural network



Reproduced with permission from Karwath A. et al, Bunting K, Gill S. 2021.(98)
Nodes and values are represented by the blue dots; hidden layers are represented by dark green dots; and the output layer is represented by the light green dot.

1.4.7 Limitations of machine learning

AI-based approaches have the potential to have a significant positive impact on cardiovascular research, but it is not without limitations. ML models require large, good quality data for training (input). Therefore, datasets need to be large, multi-modal, adequately linked and pre-processed if AI-based analysis is to provide output predictions that are relevant clinically. If AI based approaches are to integrate successfully into healthcare, then ML algorithms need be transparent and validated so methodology can be replicated and standardised.(83, 95) Additionally, ML learning models are at risk of introducing unintentional bias which can have a significant effect on the analysis that could resonate through to clinical care. Therefore, AI based studies require a multi-disciplinary team approach to ensure each step of research is carefully planned and executed. As part of my PhD I collaborated with clinicians, health data scientists and researchers as part of the BigData@heart Innovative Medicines Initiative to develop a stepwise AI framework for healthcare professionals to understand the how and when to apply AI-based approaches to clinical research (see section 1.6).(83)

1.4.8 Use of big data and machine learning in cardiovascular disease

Big data analytics (ML) has the potential to improve patient care and reduce costs through diagnostic imaging, cardiovascular risk assessment predictive prognostic modelling, drug and medical device surveillance, as well as providing insight into disease heterogeneity, and precision medicine. (See **table 5**) (99-101)

Table 5: Opportunities of big data analytics in cardiovascular research

Description	Opportunities
Large scale data	Infrequent occurrences and rare diseases can be studied
Public health improvement	Provides cost effective methods of delivering high quality healthcare
Predictive modelling	Using machine learning algorithms to identify associations in multi-variate data
International comparison	Healthcare performance and patient outcomes can by compared more easily
Personalised medicine	Individualising the benefits and harms of treatment
Drug repurposing	Identify new uses of approved drugs outside their original indication
Genetic insights	Enriched datasets provide better disease definition at a molecular level

Adapted from Silvero A. et al, 2019. (81)

Diagnostic imaging

The application of ML to cardiac imaging such as echocardiography and MRI can automate workflows leading to faster interpretation and diagnosis.(80) An example of this is the work of Narula et al where a ML model was developed to aid echocardiographic differentiation of hypertrophic cardiomyopathy (HCM) from an athlete's heart. The model was trained on 2D speckle tracking echocardiography data from 77 athletes and 62 patients with HCM. Feature selection was used to select the most relevant variables, and an ensemble-learning predictive model made up of 3 algorithms (neural networks, random forest, and support vector

machines), *was* validated using 10-fold cross validation. The ML model was found to have an increased sensitivity (87%) and specificity (82%) when compared to standard echocardiographic measures.(94)

Wearable technology

Technological advances and increased availability have led to a rise in the use of wearable devices. Consumer wearables offer the opportunity to assess certain parameters such as physical activity, step count, and heart rate, in patients with chronic diseases outside the clinical setting. Several studies have demonstrated the potential use of wearables in cardiovascular research. A study by Khushid et al analysed accelerometer derived physical activity data from 93, 669 participants who wore a wrist worn device for 1 week, to assess for an association with incident AF. After adjustment for risk factors the study found a lower risk of AF in those who adhered to guideline recommended physical activity (HR 0.82; 95% CI 0.75–0.89).(102, 103) This suggests that wearable devices can provide objective real-time monitoring of parameters such as physical activity which can be modified (adherence to recommended activity) to reduce the risk of developing AF.(103)

Risk assessment and prognostic modelling

ML learning can be used to improve risk stratification algorithms in CVD. Current risk prediction models are developed using conventional statistical regression techniques where the outcome (prediction) is limited by data assumptions and by the number of variables that can be analysed. The use of big data and ML has the potential to identify new risk factors as well as interactions between new or pre-existing ones which could improve accuracy. An example of this was seen in the work of Alaa et al where a ML model was trained on UK biobank data from over 400,000 participants using 473 variables. The ML model was found

to be more accurate in predicting CVD risk (ROC:0.774, 95% CI: 0.768-0.780) when compared to established algorithms (ROC: 0.724, 95% CI: 0.720-0.728, $p < 0.001$).(104)

Prognostic modelling can be used to predict adverse outcomes following acute events or to identify phenotype response to therapy. As previously mentioned most cardiovascular conditions are described as syndromes and are defined by the presence of a group of symptoms as opposed to underlying molecular mechanisms. As a result, there is a wide spectrum of clinical presentation and individual treatment response. Evaluating large datasets has the potential to discover distinct disease phenotypes and to identify treatment response in each sub-group. An example of this can be seen using our clustering techniques in landmark RCTs to identify treatment responses with beta blocker therapy in patients with HFrEF (see chapter 6).

Efficiency of clinical trials

Much of what is known regarding treatment and prevention of cardiovascular conditions has been derived from well-established hypothesis-based research where a new intervention is evaluated through RCTs. They are the gold standard in research as they eliminate selection biases that occurs in routine care and provide an effective way to compare the efficacy of treatments. However, in terms of planning and conducting an RCT it can be time consuming, costly and recruiting participants for adequate power can be challenging. AI-based approaches have the potential to streamline the RCT process. ML models could be used to screen and select patients based on inclusion and exclusion criteria making recruitment to clinical trials more efficient.(105, 106) ML algorithms have the potential to generate new endpoints; standardise existing structural datapoints; identify drug efficacy and adverse reactions, and enhance statistical analysis. Additionally, the use of wearables for monitoring

could reduce the need for labour intensive in person assessments. These factors can reduce the time and cost of running a conventional RCT, as well as improving its efficiency and quality.(103, 105, 106)

1.5 Wearable technology

1.5.1 Definition of wearable technology

Wearable technology can be defined as a variety of devices that are worn in direct skin contact or loosely attached to an individual with sensors that enable continuous ambulatory monitoring. There are many types of wearable devices currently available. These include; wrist bands, chest straps, clothing (T-shirt) embedded with sensors, ECG patch, smartwatches and smart rings. To varying degrees all are capable of monitoring basic physiological data, such as heart rate, heart rhythm, blood pressure, physical activity, oxygen saturation and sleep patterns.(107-110) The use of wearable technology has seen a significant increase over the last few years, especially in observational research studies as well as non-medical consumer use. In 2017 there were 526 million connected devices worldwide, this figure has risen sharply to reach over 1.1 billion in 2022 (statistia.com).

1.5.2 Wearable sensors

Wearable devices use sensors to measure parameters such as physical activity, heart rate and rhythm. Activity sensors use an accelerometer to measure physical activity from wearable devices. Accelerometers can measure linear acceleration in each axial plane and a gyroscope to assess orientation. The amount of acceleration is usually measured electrically.(110) The accuracy of activity sensors is affected by placement, therefore location on the body is important to ensure good quality recordings. A sensor placed on the torso (e.g., embedded in wearable vests or chest straps) is the ideal location for determining posture, acceleration and

whole-body movements.(111) Global positioning system (GPS) and barometers are also included in wearables to give a more accurate assessment of physical activity and location. GPS utilizes a network of satellites to identify signal position and distance. The barometer utilizes compression changes to track change in altitude, stair count and can aid fall detection.(110) With time wrist worn wearable devices have become increasingly popular as they provide a convenient, easy to wear option that increases compliance and can provide detailed information on movement patterns including; step count, distance travelled, active minutes, calorific expenditure, and sleep quality.

Consumer accessible wearable devices are able to measure heart rate and rhythm using an ECG or AI-based technology called photoplethysmography (PPG). ECG sensors are the simplest method and remain the gold standard for measuring heart rate and rhythm. These monitors are commonly worn as chest straps or patches. However, due to their size, limited functionality and inconvenience they tend to be less appealing to consumers.(110) Single lead ECG devices offer a smaller, more convenient method of arrhythmia detection such as AF, but their accuracy is lower for other arrhythmias, and they are not able to assess physical activity.(112) Wrist worn wearables such as fitness trackers, smartwatches and smart rings all utilize PPG and light emitting diode (LED) technology. PPG signals obtained from these devices are interpreted using computational algorithms via a mobile application, to provide a heart rate or rhythm. Several patented mobile applications exist but the majority do not undergo clinical validation for accuracy. The algorithms used lack transparency and are not readily available.(113) Limited studies are available that offer a comparison between consumer wearable devices and the gold standard for heart rate and arrhythmia detection, hence the true accuracy of these devices remains unknown. As part of my thesis, I conducted a systematic review and meta-analysis of the accuracy of AF detection using smartphone PPG compared to an ECG recording (chapter 3).

1.5.3 Use of wearable technology in cardiovascular care

There are several ways in which wearable technology can be incorporated in cardiovascular medicine. **Table 6** provides a summary of the various potential clinical applications, of which a few examples are discussed in more detail next.

Risk Assessment

Current risk assessment scores for CVD calculate a 10-year risk based on the presence of clinical variables associated with the disease. At present these risk scores do not account for lifestyle habits that can have affect an individual's chance of developing CVD. Previous studies have shown that physical inactivity is an independent risk factor for developing CVD, with accelerometer measured activity of moderate to vigorous intensity being associated with a lower mortality than light or sedentary activity.(110) ESC guidelines now recommend 150 minutes of moderate physical activity each week to reduce the risk of developing CVD and to improve control of modifiable cardiovascular risk factors.(114) Wearable technology provides an opportunity to assess one's physical activity and heart rate dynamically and continuously over time. Wearable data steams can provide information on resting heart rate, heart rate recovery and heart rate variability, which is important for risk stratification as high resting heart rate, poor heart rate recovery following exercise and reduced heart rate variability have all been linked to an increased risk of developing CVD.(110) The data obtained from wearables can be analysed using ML algorithms and coupled with established risk factors to provide a more personalised and accurate assessment of an individual's risk of developing CVD. Wearable devices can also be used to monitor chronic disease such as heart rate control in AF, as well as promote positive behaviour change through motivational messages or reminders.

Screening and diagnosis

Wearable devices offer a convenient method for diagnosing arrhythmias especially when it is asymptomatic. AF is the most frequently investigated arrhythmia utilizing these types of devices in clinical studies. An example of this was seen in the mSToPS (mHealth Screening to Prevent Strokes) trial, which showed immediate ECG monitoring using a sensor patch called the Zio patch (iRhythm) picked up higher rates of new AF compared to those in the delayed monitoring group.(115) Another example is the Apple heart study where over 400,000 participants were enrolled to ascertain whether a PPG smartwatch could detect AF. If an initial tachogram (signal generated from PPG pulse interval) met the criteria for being irregular, the PPG algorithm scanned for PPG irregularities where there was minimal movement artefact. If a further 4 irregular tachograms were found, a notification of an irregular pulse was sent to the participant on their Apple watch and study application (app). The positive predictive value for AF detection was found to be 84% using PPG technology on the Apple smartwatch.(116) Wearable devices that incorporate PPG technology such as smartwatches and fitness trackers allow us to move away from using bulky, inconvenient and restrictive holter monitors for arrhythmia detection.

Table 6: Application of wearable technology in cardiovascular care

Cardiovascular use	Wearable measurement	Wearable device	Potential application	Key study outcomes
Risk assessment for developing CVD	Step count	Triaxial accelerometers (Actigraph)	Capable of measuring PA and categorising into levels which can be used to assess risk of CVD	Accelerometer measured moderate to vigorous PA associated with lower mortality than light or sedentary activity in cohort studies.(117-119)
	Heart rate and step count	Fitbit, Apple Watch	Heart rate and PA can be measured continuously	Heart rate and step count data from Fitbit, Apple watch and Wear OS, used to train ML algorithm that classified high cholesterol and hypertension. (120)
Arrhythmia detection	PPG technology or Single lead ECG	Apple Watch, AliveCor Kardia band and Kardia mobile, Zio patch	AF detection with ECG patches, or PPG; wearables used to assess AF burden and rate control	mSToPS study showed continuous ECG monitoring using an ECG patch found higher rates of AF compared to delayed monitoring. (115) Apple heart study found the PPG technology on Apple Watch had an 84% PPV for AF detection.(116) Heartline RCT trial is evaluating the use of an Apple Watch (with PPG and ECG) for improving time to AF diagnosis and clinical outcomes.(121)
HF management	Heart rate, step count and single lead ECG	PhysioMem	Can be used to guide beta blocker therapies, and assess PA and HRV for risk stratification	Previous studies have shown conflicting results for telemonitoring in HF. Further studies are required to assess use of wearables in HF management. Wearables can be used to assess HF prognosis by measuring PA via a 6MWD or measuring HRV. No studies have evaluating this at present.

Coronary artery disease	Heart rate, step count and single lead ECG	Apple Watch, Bluetooth blood pressure cuff	Wearables could be used to titrate beta blocker therapy and screen for arrhythmias	No clinical studies available
Established AF	PPG technology or single lead ECG	Apple Watch, AliveCor Kardia band and Kardia mobile, Zio patch	Wearables can be used to assess AF burden, provide rhythm guided anticoagulation, and assess rate control	No clinical studies available
Cardiac rehabilitation	Step count and heart rate	Bioharness 3, Garmin forerunner, Fitbit charge	Remote cardiac rehabilitation supported by real time wearable data could reduce costs of face-to-face programmes, may improve participation rates, whilst still ensuring active engagement in the programme	Meta-analysis of 23 RCTs in patients undergoing cardiac rehabilitation showed that remote cardiac rehabilitation was as effective as central programmes in terms of improving outcomes(122)

Adapted from Bayoumy et al, 2021(110)
PA=physical activity; CVD= cardiovascular disease; ML =machine learning; PPG= photoplethysmography; ECG= electrocardiogram; AF=atrial fibrillation; mSToPS= mHealth Screening to Prevent Strokes; HRV= heart rate variability; HF=heart failure; 6MWD= 6-minute walk distance; RCTs =randomised clinical trial

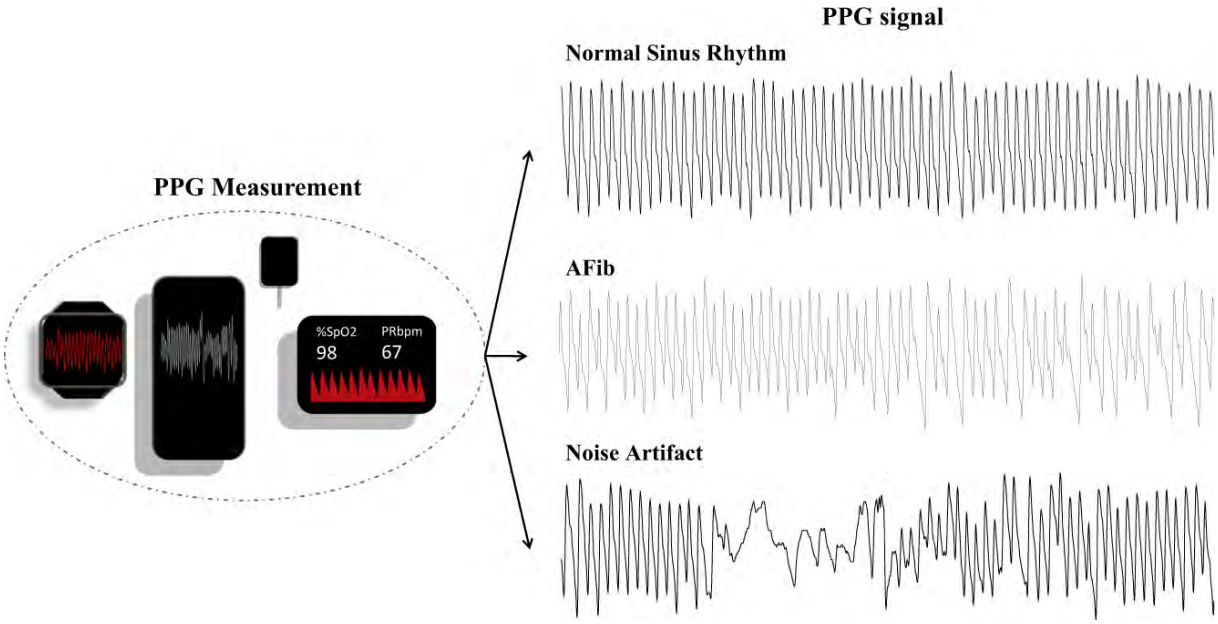
1.5.4 Photoplethysmography for arrhythmia detection

Current validated methods for detecting arrhythmias such as AF consist of an ECG or holter monitor. The issue with these conventional methods is that they are cumbersome, restrictive and do not offer additional information. PPG technology found in smartphones, and smart watches offers a low cost, non-invasive method of measuring heart rate continuously and can be used for the detection of arrhythmias.(123) PPG is a pulsatile peripheral waveform that is generated from a cardiac cycle as it descends down to the arteriolar and venous vasculature. The optically obtained PPG signal reflects changes in blood volume and respiration, and thereby provides information regarding physiological parameters such as oxygen saturations, blood pressure and pulse. PPG waveforms correspond with phases of the cardiac cycle. Ventricular systole corresponds with the positive slope of a PPG signal, following the peak the amplitude reduces to the dicrotic notch, which the beginning of diastole. Hence, beat to beat changes in blood volume are reflected in the PPG signal.(123) One of the main uses of PPG in healthcare has been for pulse oximetry to measure arterial blood oxygen saturation.

In recent years the application of PPG for the monitoring of cardiovascular parameters such as heart rate has emerged in consumer wearable fitness devices, which has led to research in using PPG as a screening tool for arrhythmias such as AF.(123) The PPG signal in AF is characterised by irregular pulse to pulse intervals and morphologies, and in sinus rhythm as evenly spaced intervals with similar morphologies (see **figure 5**). Specific features are extracted from the PPG signal and analysed using ML algorithms to assess if AF is present. However, PPG signal interpretation is limited by motion artefact due to poor contact, which is an issue as motion artefact in sinus rhythm can look similar to AF resulting in the ML algorithm incorrectly interpreting it as AF. Several techniques have been implemented to

reduce this which include identifying and removing artefact; collecting PPG signals at rest; and developing deep learning algorithms that can identify and removing artefact.(123)

Figure 5: Typical PPG waveforms seen from wearable devices



Reproduced with permission from Pereira et al, 2020(123)
(http://creativecommons.org/licenses/by/4.0/)

1.6 Stepwise workflow for the application of artificial intelligence techniques

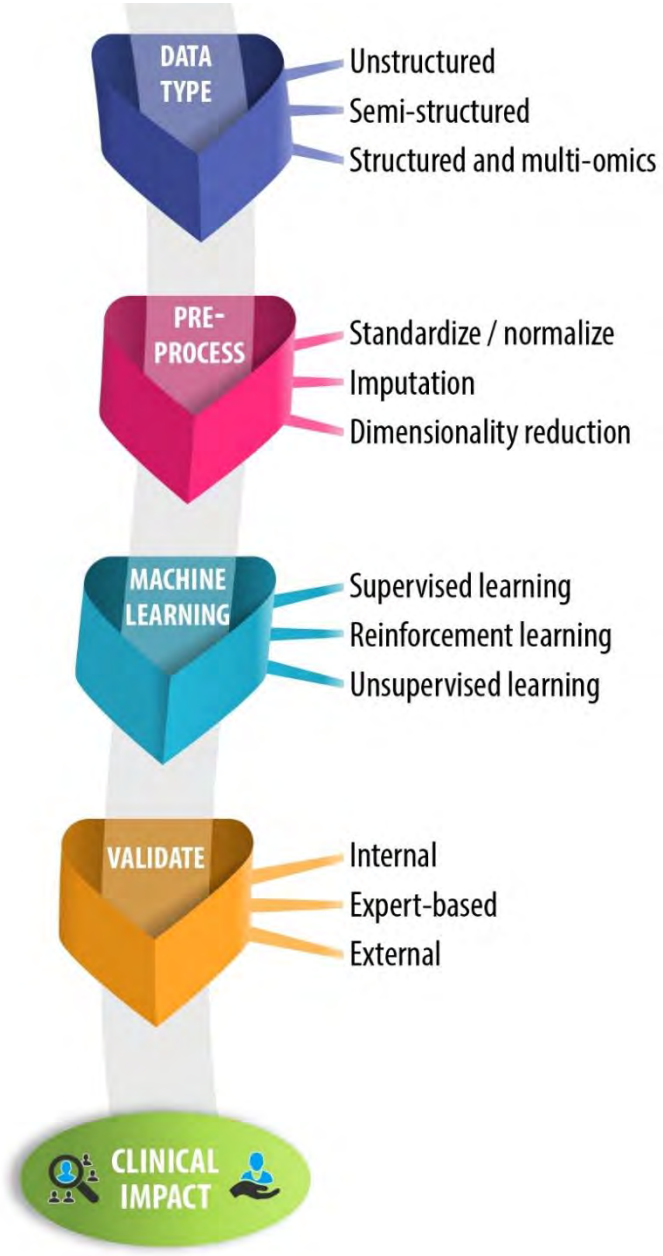
As part of my PhD I chose to address the lack of the knowledge amongst healthcare professionals surrounding the use of AI-based approaches in healthcare. I collaborated with health data scientists and researchers, as part of the BigData@heart Innovative Medicines Initiative to create a step wise framework. This was created for healthcare professionals to

understand the how and when to apply AI-based approaches to clinical research. It was designed to:

- Encourage stronger collaborations between clinicians and heath data scientists
- Provide a better understanding and educate healthcare professionals about what AI techniques to use and why
- Show how AI can improve study design and add value by providing case studies that have produced findings which are relevant to clinical practice.

The workflow has been published and below each of the four main steps are described along with the context of why certain aspects were included (see **figure 6** below).

Figure 6: Artificial intelligence framework for clinical application



Reproduced from Gill S et al, 2023
Overview of the artificial intelligence framework

1.6.1 Step 1: Data type

Following the formation of a clear hypothesis, the reader is directed to the first step in the workflow which relates to data type. This advises the reader to think about the type of data modality (structured, semi-structured and unstructured), which determines what data is

collected (see **table 7**), and the AI-based algorithms needed for analysis. The issue of potential bias is introduced, where just like in conventional research AI models are prone to bias and this begins with where the data came from and why it was collected. This question helps the reader to determine if underlying bias (for example, older age or a certain ethnicity group) in data is already present.

Table 7: Type of data modality and data collected

Data modality	Types of data collected
Structured information	Demographics Blood biomarkers Electronic healthcare records
Semi-structured information	Text-based data Continuous time series data
Unstructured information	Any type of imaging data

Structured data obtained from clinical trials is commonly used as it is easily interpreted to healthcare (see the clustering work described in section 6.3), whereas unstructured data such as sensor data often requires more processing so that latent, hidden or otherwise unknown relationships can be identified (see chapter 5 for wearables analysis). Data can also be combined from other anonymised sources in a secure way that maintains data protection and privacy. With regards to sample size, this is dependent on the type of data collected, for example in the wearables sub-study (chapter 5) there were only 53 patients recruited but due to the continuous data collection process millions of data points were obtained per patient over the study period. If only a small number of observations are obtained, then AI techniques do not provide any added value.

1.6.2 Step 2: Pre-processing of data

The second step in the workflow is pre-processing and this is a critical step in any AI-based approach as the output of the model is only as good as the data fed into it (input). The reader is reminded about the importance of exploring any underlying bias and transforming data into a standardised format suitable for ML analysis. In terms of transforming data numerical variables are usually converted to a normal distribution (standardisation) or a pre-defined range (normalisation). Categorical data is commonly converted into a binary format. Other types of data are pre-processed differently, for example, time-series data is normalized by the capture period (see chapter 5), and imaging data can be contrast-adjusted, scaled or segmented.

Another important factor discussed in pre-processing is the need to investigate missingness of data, which can occur for three main reasons. Missing data may have occurred during due to a completely random process such as when data collection is paused, for example due to a brief connection issue as seen during the wearables-sub-study (chapter 5). Secondly, it can occur at random such as when a patient is lost to follow up or withdraws from a study. Finally, it can occur due to a non-random process where the presence of missing data cannot be explained. The are several ways to deal with missing data which include; imputation or deletion. Another important consideration mentioned in pre-processing is dimensionality reduction, where datasets with a large number of variables are reduced in dimension (see section 6.3.2), by linear or non-linear approaches such as autoencoders.

1.6.3 Step 3: Choosing the right machine learning approach

The third step in this framework involves choosing the correct ML algorithms to answer the research question posed. Supervised ML uses labelled data to perform a prediction such as disease prognosis. Unsupervised ML are exploratory and learn patterns from unlabelled data, so they can identify unknown patterns between variables, as seen with the clustering technique (see section 6.3) described earlier.

There are a number of machine learning algorithms available which have been used in healthcare, these include:

- Decision Trees which are formulated by iteratively splitting training data until it can no longer be split. Commonly used for classification and regression
- Random Forests which build models based on multiple decision trees, with each tree trained on a random sub-set of data
- Neural networks consisting of multiple layers of computation, which can learn across a dataset. Deep neural networks use additional layers to model complex non-linear relationships (see chapter 5)
- Autoencoders use dimensionality reduction to learn a compressed version of original data (see section 6.3)

Each of these AI-based approaches and more have been used in healthcare research and with time their use in clinical studies will continue to expand.

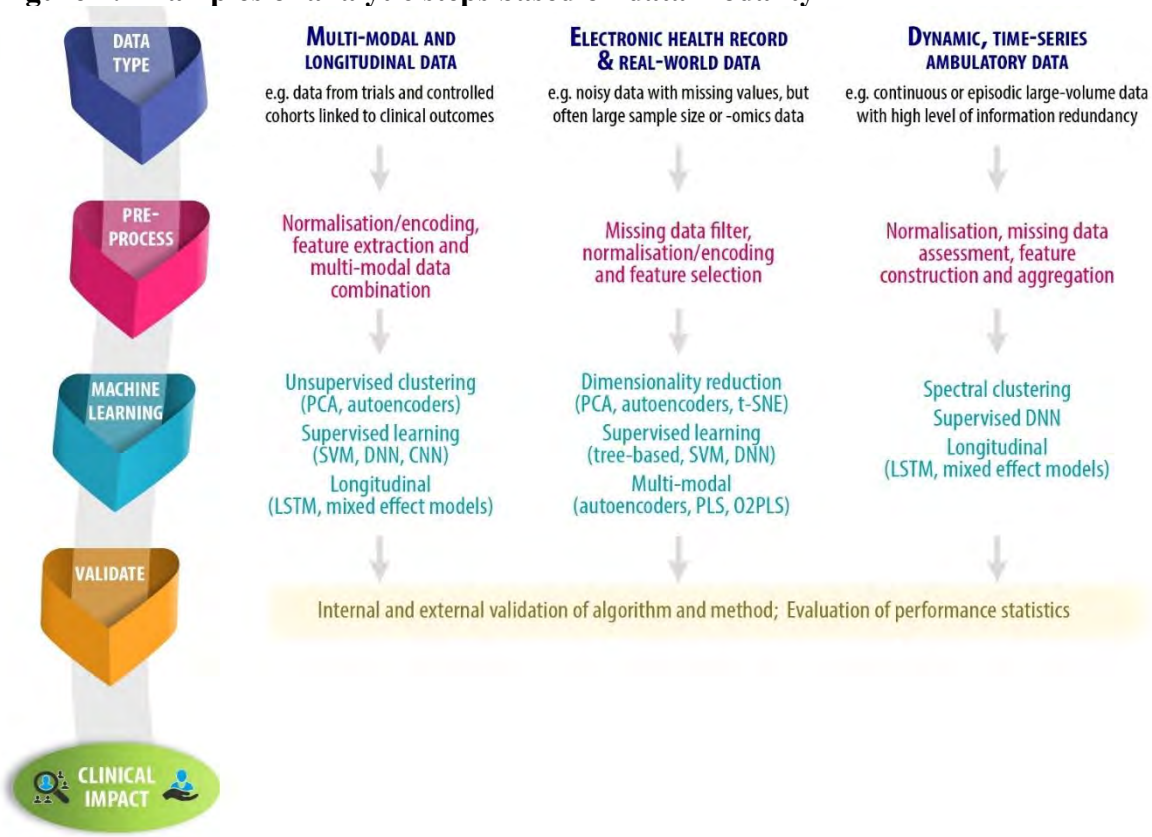
1.6.4 Step 4: Validating and evaluating methods and results

The final step in the workflow directs the reader to the importance of evaluating and validating the performance of a ML model, to ensure the findings are applicable to the real-world. Evaluating ML output is not dissimilar to evaluating conventional predictive outcomes in that challenges arise when generalizing from one data set to another, and then again to real-world practice. Validation methods are described and can either be internal or external, external validation is the gold standard as it evaluates model performance on previously unseen data from a different cohort or study, ideally from a different location and time period. Internal validation such as the train-test split or cross validation (discussed in chapter 1) can be used to ensure the model is working correctly and is not overfitting (fits training data too well) data. Sole reliance on this method is not suggested, although may sometimes be necessary where models cannot be shared.

External validation on unseen data is the suggested default option, but this is not always possible and internal validation is accepted. There are different forms of validation which include K-fold cross validation, which divides the data into distinct K-subsets (folds); each subset is used once for testing and the remaining (K-1) subsets for training a model. Each fold will have a performance metric that can be combined to estimate overall performance of the final model based on unseen data. An alternative is bootstrapping which is a method that randomly splits the data into training and test sections randomly multiple times with replacement so that model performance on unseen data can be estimated and confidence intervals created.(83)

Evaluation of machine learning algorithms requires performance measures these include sensitivity, specificity, positive or negative predictive values, F1 score and ROC. The F1 score is a measure of the model's accuracy (how many times the model made a correct prediction compared to the whole dataset) and a combination of the precision and recall scores of the model. Another measure is the ROC analysis where the area under the curve is indicative of model performance irrespective of class distribution. Both performance statistics have a value range from 0 to 1, with higher values showing improved model performance. **Figure 7** below provides examples of each analytical step based on different data modalities.

Figure 7: Examples of analytic steps based on data modality



Reproduced by Gill S. et al, 2023
CNN = Convolutional neural network; DNN = Deep neural network; LSTM = Long short-term memory recurrent network; PCA = Principal component analysis; PLS = Partial least squares regression; O2PLS = orthogonal 2-way PLS; SVM = Support vector machine.

1.6.5 Patient and public involvement

With rapid expansion of AI technology in healthcare, patient and public involvement is essential to the development of AI based studies, public acceptance and dissemination of the results. Following consultation with the card*AI*c patient and public involvement and engagement team, **table 8** provides an insight into the patient and public perspective regarding the use of AI in healthcare.

The AI stepwise workflow was applied to the wearables sub-study (chapter 5) and cluster analysis (chapter 6.3).

Table 8: Patient and public perspective on artificial intelligence

Thoughts about artificial intelligence (AI)?	**What are the challenges of implementing AI in healthcare?**
An exciting new approach that offers huge potential for better healthcare. Could improve and speed up diagnosis, treatment and referral processes. Becoming commonplace in daily life but not as visible for healthcare applications.	Lack of understanding and training on AI amongst healthcare staff, and how it could support their clinical work. Cost of undertaking AI, with rapid technical advances resulting in equipment becoming obsolete quickly. Large energy requirement for computer processing and lack of environmental awareness in line with developing public attitudes. Need for transparent processes to ensure and demonstrate accuracy of data and reliability of algorithms.
Where are the opportunities for using AI in the healthcare setting?	**What do patients and the public want and need before more widespread use of AI in healthcare?**
Better monitoring, awareness and interest in healthcare issues, for example using consumer wearable devices. Potential to reduce the need for patients to attend primary or secondary healthcare facilities. Can leverage existing datasets without further direct patient involvement and ability to augment clinical decision-making. Can help with unknowns by revealing information not immediately obvious to clinicians, which could improve treatments for patients. To facilitate more efficient mass screening, reduce false positives and focus therapy on those at higher risk of poor outcomes.	Better communication with the public is required to improve understanding of AI in healthcare and ensure future engagement. Reassurance regarding data security, in particular ownership, privacy and potential sale of data to third parties and companies. Technological developments must not accentuate existing health inequalities but work to reduce them.

1.6.6 Limitations of the AI framework

Although this AI-based framework attempts to provide guidance and standards on the use of AI approaches in healthcare, caution should be exercised regarding the underlying assumptions and limitations of these techniques. An important limitation is the black box

effect seen in deep learning where the entire computational process is not transparent.(83) Furthermore, there is a lack of knowledge amongst healthcare professionals and the general public regarding what AI technology is, how and when it can be used in healthcare, and what the pros and cons of are. Coupled with the rapidly expanding availability and application of AI the void in our knowledge is increasing (124). This published article was designed to tackle some of these issues by providing a framework for how to use AI based approaches in clinical research, with real-life case studies and the involvement of the PPI team. However, a deeper understanding about the use of AI in both clinical research and daily life is required across all stakeholders which can only be achieved through further education.(83)

Current published AI studies appears to lack step 4 in the framework, evaluation and validation of the AI technique, which is a significant concern. A systematic review of 82 studies evaluating the diagnostic accuracy of healthcare professionals versus deep learning algorithms found only 25 performed an external validation.(125) Considering this the CONSORT-AI (consolidated standards of reporting trials-artificial intelligence) extension was derived which is a reporting guideline that should be used for AI-based clinical trials to ensure greater transparency and completeness (126, 127). With regards to AI based or hybrid studies researchers need to clear about the methods used and should follow the FAIR principles (Findability, Accessibility, Interoperability and Reusability)(128), so that other research groups can access data, provide independent validation and additional clinical value.(83)

AI-based approaches required large datasets, going forward with this in mind specific consideration needs to be given to data protection. Data privacy issues, consent (opt-in and opt-out), data access justification, data sharing and dissemination, were some of the major

issues highlighted by an international stakeholder group for the use of electronic healthcare data in research.(129) Strong patient and public engagement during study design and development process can be helpful to ensure these factors are addressed early on.(130)

2. Research aims and hypotheses

2.1 Aims

The first aim of this PhD was to evaluate the current state of the art approaches for improving disease phenotyping using big data and ML in CVD. This highlighted the potential opportunities wearable technology could provide in terms of screening and monitoring. From a screening perspective, it was clear that PPG technology found in smartphones and smartwatches offered a novel screening opportunity to identify common cardiovascular conditions such as AF. I was particularly interested in this technology as it has the potential to provide a low cost and non-invasive screening method for AF detection. However, the published studies available that utilized this technology were often limited by size and bias. Therefore, I chose to conduct a systematic review and meta-analysis of smartphone PPG technology for AF detection compared to standard ECG recordings.

The use of non-medical consumer wearable devices by the public continues to rise. Their ability to provide continuous data on physiological parameters in patients with chronic diseases such as AF and HF makes their use in clinical research appealing. Currently, there is limited robust RCT data available regarding the use of these wearable devices. Incorporated into an RCT, the next part my PhD looked at the use of wearable devices for monitoring PPG heart rate and physical activity, and whether they could provide useful information in clinical research. The second aim of my PhD was to set up and conduct a wearables sub-study within

the RATE-AF RCT. This would evaluate the effect of beta blockers and low-dose digoxin on heart rate control and physical activity levels using sensor data derived from the wearables.

The final part of my PhD focused on evaluating the use of conventional statistical methods and ML techniques. The aim was to evaluate beta-blocker efficacy in multi-morbid HFrEF and improve phenotyping based on drug efficacy. Through this, it was clear that the use of AI in clinical studies lacked clarity and standardisation. Hence, the final aim of my PhD was to develop a stepwise AI framework to aid healthcare professionals with an understanding of how and when to apply different types of AI-based approaches in healthcare research.

2.2 Hypotheses

Hypotheses 1: On systematic review and meta-analysis the sensitivity, specificity and diagnostic accuracy of smartphone PPG applications for AF detection is comparable to a 1, 3 or 12-lead ECG.

Hypotheses 2: Within the RATE-AF trial, the effectiveness of beta blockers and low-dose digoxin on PPG heart rate (measured from the Fitbit wearable device), is comparable for long term AF rate control.

Hypotheses 3: Within the RATE-AF trial, the effectiveness of beta blockers and low-dose digoxin on total step count (measured from the Fitbit wearable device), as a measure of physical activity is comparable.

Hypotheses 4: Within the RATE-AF trial, neural networks trained on wearable sensor data are comparable to conventional measurements of ECG heart rate and 6-minute walk test for prediction NHYA class at the end of the trial.

Hypothesis 5: In multi morbid HFrEF patients, a machine learning cluster approach is superior to conventional statistical analysis for identifying sub-groups with similar beta-blocker efficacy.

3. A systematic review and meta-analysis of smartphone detection of atrial fibrillation using photoplethysmography

3.1 Introduction

3.1.1 Atrial fibrillation and need for early detection

With increased life expectancy and an ageing population, the incidence and prevalence of AF has increased globally by 3-fold over the last 50 years.(131, 132) This rise is partly due to improved awareness, detection and survival from cardiovascular conditions that can lead to its development.(132) Despite research efforts to improve our understanding of the underlying disease pathophysiology and develop novel therapies, the treatment and prevention of AF remains challenging and poses a strong public health concern due to its associated high risk of stroke, morbidity and mortality, alongside rising healthcare costs.(132-135)

One of the most significant and debilitating complications of AF is thromboembolic stroke where the risk is present regardless of the type (persistent, paroxysmal or permanent) of AF present. AF related strokes make up a third of all ischaemic strokes and are more likely to be fatal and disabling when compared to strokes from other causes.(132) According to Schnabel et al between 11.5% and 24% of those who present with a stroke or transient ischaemic attack (TIA) are found to have a new diagnosis of AF on their ECG at time of presentation, or during a monitoring period following the event.(132) AF related strokes are preventable if anticoagulation is commenced promptly in those found to be at risk (CHA_2DS_2-VASc score) of ischaemic stroke. (136, 137) Early identification is key especially in those with silent AF, to reduce the impact AF-related stroke can have on patient quality of life.(138) Additionally, an early diagnosis of AF can be beneficial in terms of risk factor modification, commencing

rate control therapy or offering definitive treatment (ablation) for symptomatic individuals, and identifying and treating co-morbidities such as heart failure.(16, 139-141)

3.1.2 Screening for AF

The ESC guidelines advocate opportunistic screening of AF in patients aged over 65 years. This is done by a one-off pulse recording; single time point ECG or ambulatory ECG monitoring over a two-week period.(16) The risk with one-off screening is that it does not detect the brief and sporadic paroxysmal AF episodes that are equally likely to cause a stroke in those pre-disposed. A systematic review and meta-analysis of clinical trials that evaluated a one-off ECG or pulse palpation for AF detection (n=122, 571), found in those aged over 65 (n=18,189), single timepoint screening identified 1.4% with undiagnosed AF.(142) Longer intermittent hand-held ECG recordings of up to two weeks identified new AF in up to 7.4% of the general population aged between 75-76 years.(143, 144) In the STROKESTOP study (systematic ECG screening for AF among 75-year-old subjects), single time point ECG recordings had a sensitivity rate for new AF detection of 0.5%, which was much lower than the group with twice daily home ECG monitoring with higher sensitivity rates of 3%.(136, 142, 145) From this we can see that frequent screening for AF is more likely to detect those who are asymptomatic. Advances in technology have led to the development of novel, non-intrusive wearable screening devices such as smartphones, smartbands or smartwatches which are small, portable, easy to use and cost effective. These wearable devices offer the opportunity for patient-activated, self-detection of arrhythmias such as AF, which can increase screening coverage and improve early detection of asymptomatic paroxysmal AF.(123, 141, 146) There are a number of different types of AF screening tools available. These include:

- Intermittent pulse palpation
- Continuous ambulatory ECG holter monitoring ranging from short term (24 to 48 hours) to longer term (up to 2 weeks)
- Implantable cardiac monitors (loop recorder device)
- Intermittent single lead rhythm strip using a hand-held device connected to smartphone
- Intermittent patient-initiated PPG on smartphone
- Intermittent PPG on smartwatch or wearable

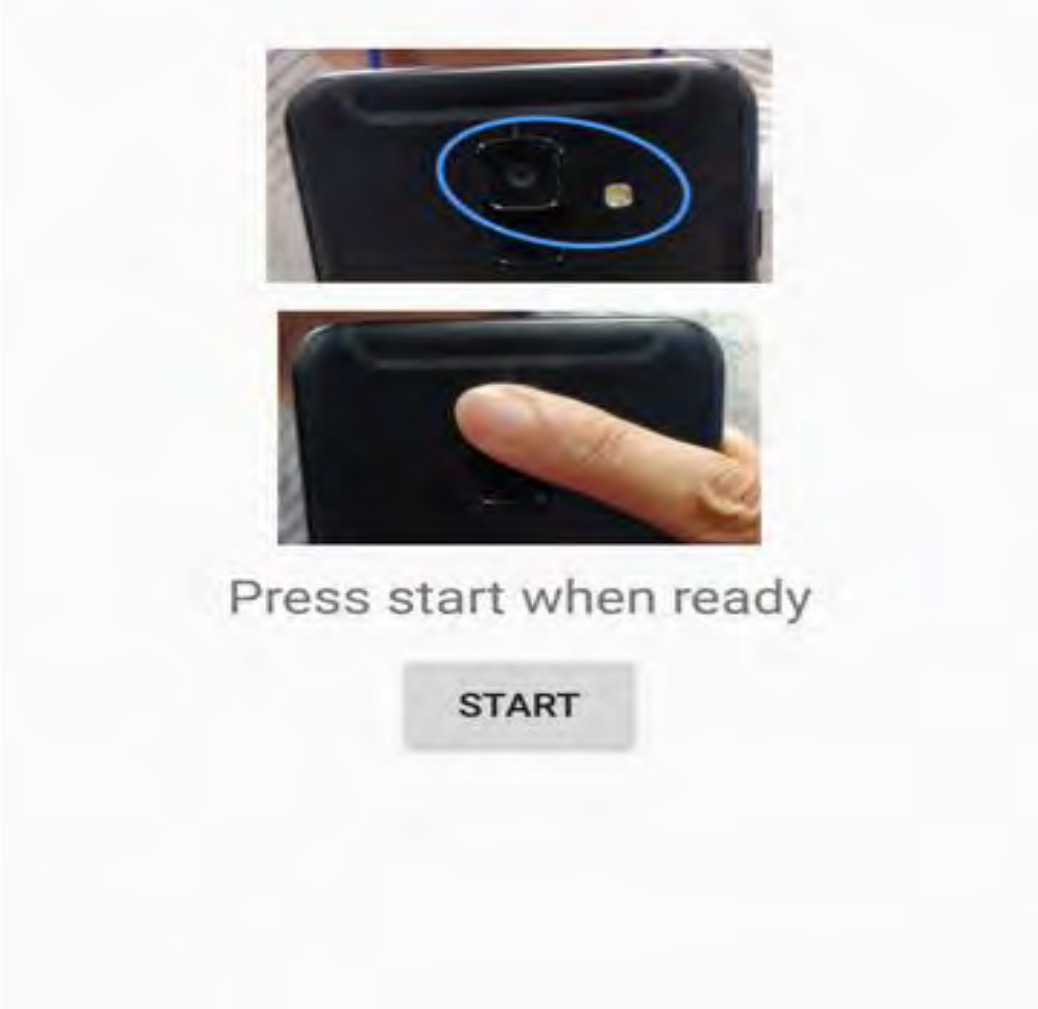
3.1.3 Photoplethysmography

Smartphone applications and smartwatches utilize PPG technology that is found in smartphone cameras for AF screening by patients (see **figure 8**). **Figure 9** below shows simultaneous PPG (blue), blood pressure (red) and ECG (green) signals. The premature ventricular complexes can be seen in the PPG signal as well as the blood pressure and ECG recording. A PPG device requires two components; a light source to emit light to a tissue, and a photodetector to measure light that is reflected from the tissue which is synchronous with changes in blood volume.(147) A downloadable application can analyse the PPG signals obtained from screening devices using sophisticated machine learning algorithms that assess for variations in beat-to-beat intervals and pulse wave morphology.(141)

Recognising PPG technology as a way to improve the identification, phenotyping and management of patients with AF, I chose to systematically review novel AF detection methods. At the start of my PhD although several types of wearable screening devices were used in research, the use of smartwatches and smartbands to detect AF were not readily

available to consumers. Therefore, I chose to look at the accuracy of currently available smartphone PPG applications for AF detection. I completed a systematic review assessing the diagnostic accuracy of AF detection using smartphone PPG applications compared to ECG recordings, with the aim of understanding their efficacy against a gold standard measurement and to provide clinicians with guidance on their value, potential use in clinical practice and limitations.

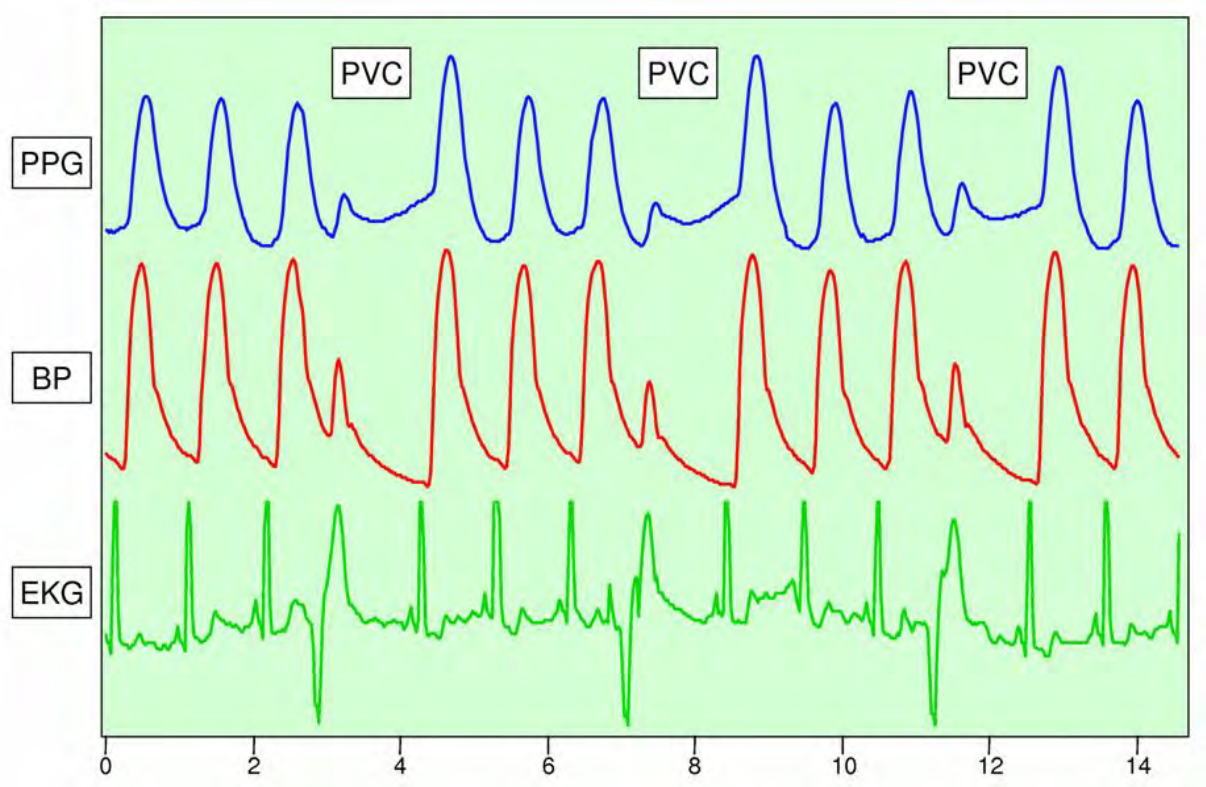
Figure 8: Smartphone camera fingertip photoplethysmography



Picture produced by S. Gill using Samsung A6 smartphone

.

Figure 9: Comparing PPG, blood pressure and ECG signals



(Image reproduced from
https://en.wikipedia.org/wiki/Photoplethysmogram#Sites_for_measuring_PPG)

3.2 Methods

3.2.1 Study design

For this review I lead the team who performed the search and extraction process. This Systematic review and meta-analyses were conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and was prospectively registered with the PROSPERO database of systematic review. (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019109455). Initial screening and extraction for systematic review and meta-analysis were performed by two independent operators (KB and SG). Full text publications were short listed based on inclusion and exclusion criteria set out below (see **table 9**). Inclusion criteria was kept quite broad to ensure all publications included conference abstracts could be evaluated. All forms of AF were included, and a reference ECG was required for comparison of accuracy (AF detection) to smartphone camera PPG signal. Examples of publications that were excluded as they did not meet the inclusion criteria included: PPG signals generated from wrist worn devices (due to the requirement for additional hardware); hand-held ECG devices (such as Alivecor) as these did not use PPG technology, and studies that lacked a reference ECG or did not detect AF.(145)

Table 9: Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Persistent, paroxysmal, and permanent AF evaluated	Editorials excluded
Original research studies and conference abstracts included	Reviews excluded
Participants needed to be aged 18 and over	Studies generating PPG signal from other wearable devices
Comparison of PPG to a 1, 3 or 12-lead ECG	Studies that lacked a reference ECG
Smartphone camera PPG signal from fingertip or face used for AF detection	Studies that did not meet the inclusion criteria

3.2.2 Search strategy

A systematic review of MEDLINE (1950 to December 1st 2020), EMBASE (1980 to December 1st 2020), and the Cochrane Library (until 1st December 2020), were performed without language restriction. The search terms used in Embase, Medline and Cochrane databases can be found in tabular form in **appendix 1 and 2**. Reference lists of relevant studies were also manually searched for any further available literature. This study was conducted with the principals of the World Medical Association's Declaration of Helsinki. In this case, separate ethical approval was not required as the study is a meta-analysis of previously published tabular information from relevant studies.

3.2.3 Outcomes

In this systematic review the primary outcome was validation of AF detection using a smartphone camera PPG signal against a gold standard ECG. This was done by calculating sensitivity, specificity, positive predictive value and negative predictive value. In addition, the new test (smartphone camera PPG signal) was assessed for overall accuracy against the gold standard test for AF detection (standard 1, 3 or 12-lead ECG). **Table 10** summarises the outcomes that were assessed.

Table 10: Calculating sensitivity, specificity positive predictive value and negative predictive value for AF detection using PPG signals from smartphone camera

	AF present on ECG	**AF absent on ECG**
PPG smartphone test positive	True positive (a)	False positive (b)
PPG smartphone test negative	False negative (c)	True negative (d)
	Sensitivity a/(a+c)	Specificity d/(b+d)
	Positive predictive value a/(a+b)	Negative predictive value d/(c+d)

PPG=photoplethysmography; AF=atrial fibrillation; ECG=electrocardiogram

Sensitivity

Sensitivity is defined as the percentage of individuals who are true positives for the disease and the test correctly picks this up. For this review it will be the test's ability to pick up individuals with AF who test positive .

Specificity

Specificity is defined as the percentage of individuals who do not have the disease and the test correctly classifies them as not having it. For the review it will be the number of individuals who do not have AF and correctly test negative using the test.

Positive predictive value

Positive predictive value (PPV) is the percentage of individuals who have a positive test and are positive for the disease. For this review it will be the percentage of individuals with a positive test who actually have AF. Higher values indicate the test is as good as the gold standard.

Negative predictive value

Negative predictive value (NPV) is the percentage of individuals who have a negative test and are negative for the disease. For this review it will be the percentage of individuals with a negative test who do not have AF.

3.2.4 Data extraction

The publications obtained from the literature searches were transferred to a more readable tabular format in Microsoft Excel. Two reviewers independently screened the title and abstract to assess if they were eligible for full text review. Following this both lists were reviewed by myself, and any discrepancies were highlighted and discussed between us. If there was still uncertainty the original publication was reviewed, and a decision was made. A final list of publications was created.

In the next stage two reviewers assessed the full text of each article or abstract using four evidence-based hierarchy criteria :

(1) original research reporting findings for all outcomes considered

(2) original research reporting findings for some, but not all, outcomes

(3) conference abstract reporting findings for all outcomes;

(4) conference abstract reporting findings for some, but not all, outcomes.

Any of the full-text studies that met one of the criteria above was selected for extraction of data and further information. A table in Microsoft Excel was created to collect relevant information. The following key points were collected:

- Type of study (article or abstract)
- Total number of participants
- Study setting (primary or secondary care)

- Study design
- Prevalence of AF
- Type of smartphone used
- Type of PPF used (finger or facial)
- Type of Smartphone application used
- Recording time for PPG
- Number of PPG measurements made each recording
- ML algorithm used for PPG analysis
- Outcomes measured (sensitivity, specificity, PPV, NPV, and accuracy)
- Reference test used and how analysis was obtained
- Demographic and medical history information

If missing parameters were found, the lead authors from the publication in question were contacted to obtain additional data, but none was received. Industry involvement in the study was noted if appropriate. Examples of this included funding, authors who work in industry, and the provision of devices or technology such as the smartphone mobile application used in the study. A summary of included full-text publications and conference abstracts can be found in **tables 11 and 12** respectively).

Table 11: Summary of full text publications

Study	Study design and key enrolment criteria	Setting and sample size	Population characteristics	Technology for AF detection	Reference test
Brasier 2019(148)	Prospective, multi-centre Age >18 years, capable of written consent Supported by industry	Secondary care N=672 AF prevalence 42%	Age 78 years (median); Female 45%; hypertension 72%; diabetes 31%, heart failure 36%; stroke 16%, OAC 49%	iPhone 4S; Preventicus app; 300 seconds recording; Data quality check performed prior to rhythm analysis that used beat-to-beat changes of pulse wave time intervals and morphology	Blinded interpretation of single lead ECG by 2 cardiologists with group consensus; Three study comparisons with PPG signal analysed at (a) 60 seconds; (b) 120 seconds; and (c) 300 seconds.
Chan 2016(149)	Prospective, single centre Age ≥ 65, history of hypertension, diabetes Supported by industry	Primary care, N=1013 AF prevalence 3%	Age 68 years (mean); Female 53%; hypertension 90%; diabetes 37%; heart failure 4%; stroke 11%	iPhone 4S; CRMA app; 3 X 17 second recordings, baseline wander and noise filtered. AF detection based on a lack of repeating patterns in the PPG waveform, using SVM. Labelled AF if 2 of 3 recordings irregular	Blinded interpretation of single lead ECG by 2 cardiologists with group consensus.
Fan 2019(150)	Prospective, single centre Age >18 years Excluded if unable to use smartphone or had memory impairment Supported by industry	Secondary care N=108 AF prevalence 48%	Female 42%; diabetes 30%, heart failure 13%; stroke 12%; OAC 46%	Huawei Mate 9, Huawei Honor 7X; Preventicus app; 180 second recording analysed	12-lead ECG interpreted by 2 cardiologists with group consensus.
McManus 2013(151)	Prospective single centre AF for DCCV	Secondary care N=76 AF prevalence 100%	Age 65 years (mean); Female 35%; hypertension 71%; diabetes 28%; heart failure 21%; stroke 12%	iPhone 4S; unknown app; 120 second recording, analysed using 2 statistical techniques (RMSSD and ShE)	12-lead ECG interpreted by trained physicians with group consensus.
McManus 2016(152)	Prospective single centre AF for DCCV & premature beats	Secondary care, N=121 AF prevalence 81%	Age 66 years (mean); Female 18%	iPhone 4S; PULSESMART app; 120 second recording analysed using 3 statistical techniques (RMSSD, ShE, Poincare plot)	12 or 3 lead ECG, interpreted by trained physicians with group consensus.

Mutke 2020(153)	Prospective, multi-centre; data from 2 trials (WATCH AF and DETECT PRO) Supported by industry	Secondary care N =1330 AF prevalence 47%		iPhone 4S; Preventicus app; 60 second recording analysed using beat-to-beat variations via a nonlinear rhythm analysis, signal quality check not performed	Single lead ECG. Interpretation by 2 cardiologists with group consensus.
Poh 2018(154)	Retrospective analysis with DCNN for AF detection Supported by industry	Validation data from primary care N= 1013 AF prevalence 3%	Age 68 years (mean); Female 53%; hypertension 90%; diabetes 37%; Stroke 11%; heart failure 4%	iPhone4S; unknown app; 3 x 17 second recordings analysed using six AF detection algorithms (CoV,5 CoSEn, nRMSSD + ShE, RMSSD + SD1/SD2, Poincaré plot and SVM)	Blinded interpretation of single lead ECG by 2 cardiologists with group consensus.
Proesmans. 2019(155)	Prospective multi-centre Age ≥ 65 years, paroxysmal or persistent AF Supported by industry	Primary care N=223 AF prevalence 46%	Age 77 years (mean); Female 53%; diabetes 20%; heart failure 29%; stroke 22%; OAC 56%	iPhone 5S; Fibricheck app; 3 x 60 second recordings; signal quality evaluated using RR-interval variability analysis; AF detection based on recurrent neural network algorithm	Blinded 12-lead ECG interpretation by 2 cardiologists with group consensus.
Rozen 2018(156)	Prospective single centre Age >18 years, AF for DCCV Supported by industry	Secondary care N=97 AF prevalence 90%	Age 68 years (mean); Female 25%	iPhone; CRMA app; 3 x 20 second recordings analysed using SVM to classify PPG waveforms; feature extraction used to determine self-similarity of waveform; labelled AF if at least 2 of the 3 recordings irregular	Blinded 12-lead ECG interpretation by 2 cardiologists with group consensus.
Yan 2018(157)	Prospective single centre Supported by industry	Secondary care; N=233 AF prevalence 35%	Age 70 years (mean); Female 30%; hypertension 60%; diabetes 35%; heart failure 32%; stroke 19%	iPhone 6S; CRMA app; 3 x 20 second recordings, baseline wander and noise filtered; AF detection using SVM (based on lack of repeating patterns); AF if irregular in ≥1, or 3 consecutive uninterpretable measurements	Blinded interpretation by cardiologist of 12 lead ECG; two study comparisons of (a) facial PPG and (b) finger PPG.

Abbreviations: AF=atrial fibrillation; ECG=electrocardiogram; Cov=Coefficient of variation; CoSEn=Coefficient of sample entropy; CRMA=Cardiio Rhythm smartphone application; DCCV=direct current cardioversion; DCNN=Deep convolutional neural network; OAC=oral anticoagulation; PPG=photoplethysmography; RMSSD=Root mean square of successive RR differences; ShE=Shannon entropy; SVM=Support vector machine.

Table 12: Summary of characteristics for conference abstracts

Study	Study design and key enrolment criteria	Setting and sample size	Population characteristics	Technology for AF detection	Reference test
Grieten 2018-1(143)	Prospective multi-centre Age>65 years	Primary care N=242 AF prevalence 38%	Age 77 years (mean); Female 57%; hypertension 84%; diabetes 22%; stroke 23%; heart failure 29%; OAC 55%	Unspecified smartphone; Fibricheck PPG app and algorithm	Single lead ECG
Grieten 2018-2 (158)	Prospective, multi-centre	Primary care N=1056 AF prevalence 0.8%	Age 59 years (mean); Female 59%	Unspecified smartphone; Fibricheck PPG app and algorithm	Single lead ECG
Karim 2017(159)	Prospective, single centre Sinus arrhythmia and ectopics excluded	Secondary care N=140 AF prevalence 50%		Unspecified smartphone; PPG Preventicus app and algorithm	12-lead ECG
Kuan 2018(160)	Prospective, single centre	Secondary care N=194 AF prevalence 35%	Age 70 years (mean); Female 27%	iPhone; CRMA app 3 X 20 second recordings, AF labelled if pulse irregularity found in ≥1 PPG readings or 3 uninterpretable readings	12 lead ECG
Maitas 2012-1(161)	Prospective, single centre AF for DCCV	Secondary care N=52 AF prevalence 100%		iPhone 4S; unknown PPG app, 120 second recordings, analyzed using 2 statistical techniques (RMSSD, Shannon entropy) to examine heartbeat intervals and beat-to-beat variability	12-lead ECG
Maitas 2012-2(162)	Prospective, single centre AF for DCCV	Secondary care N=33 AF prevalence 100%		iPhone 4S; unknown PPG app, 120-300 second recordings, analyzed using 3 statistical techniques (RMSSD, ShE and Sample entropy)	12-lead ECG

Mortelmans 2017(163)	Prospective, single centre Majority with history of AF	Primary care N=242 AF prevalence 66%		Unspecified smartphone; Fibricheck PPG app, 60 second recording	Single lead ECG
Mutke 2019(164)	Retrospective analysis of data from 2 prospective multi-centre validation trials	Unclear N=1096 AF prevalence 44%		Unspecified smartphone and PPG app; 60 second recording automated signal quality check and categorized into AF or SR using unspecified algorithm	Blinded interpretation of single lead ECG by cardiologists
Napolitano 2015(165)	Retrospective analysis of pulse recordings from patients who underwent DCCV	Data form secondary care N=121 AF prevalence 81%		iPhone 4S; PULSESMART app, 120 second recording analysed using 3 statistical techniques RMSSD; ShE; Poincare plot	12 or 3 lead ECG
Proesmans 2019-2(166)	Prospective, single centre AF for DCCV and in-patients with AF	Secondary care N=164 AF prevalence 37%	Age 64 years (mean); Females 42%	7 iOS and 7 android smartphones; unknown PPG app	12 lead ECG
Proesmans. 2018(167)	Prospective, 2 centre Age over 40 years	Primary care N=1095 AF prevalence 0.5%		Unknown smartphone and PPG app; automatic algorithm and visual interpretation	Single lead ECG
Rozen 2017(168)	Prospective, single centre AF undergoing DCCV	Secondary care N=113 AF prevalence 100%	68 years (mean); Female 23%	iPhone; CRMA PPG app, 6 recordings analysed by algorithm 2 out of 3 irregular recordings needed to label AF	Blinded 12 lead ECG or single lead ECG (rhythm strip) interpretation by 2 cardiologists
Siu 2016(169)	Prospective, single centre Either >65 years, DM, or hypertension	Primary care N= 1027 AF prevalence 3%	Age 68 years (mean)	Unspecified smartphone; CRMA PPG app	Single lead ECG interpretation by 2 cardiologists

Smeets 2019(170)	Prospective single centre AF scheduled for DCCV/ablation & those on continuous monitoring	Secondary care N=150 AF prevalence 37%	64 years (mean); Females 42% (AF)	6 Android and 2 iOS smartphones; unknown PPG app;1 recording per device	12-lead ECG
Vaid 2015(171)	Prospective single centre AF undergoing DCCV	Secondary care N= 37 AF prevalence 78%	69 years (median); Female 19%	iPhone; CRMA PPG, 6 recordings analyzed by CRMA algorithm 2 out of 3 irregular recordings needed to label AF	Blinded 12 lead ECG or single lead ECG (rhythm strip) interpretation by 2 cardiologists
Vandenberk 2018-1(172)	Prospective single centre History of AF	Unclear N=344 AF prevalence 50%		Unspecified smartphone, Fibricheck PPG app, analyzed by 2 blinded cardiologists	Blinded 12 lead ECG analysis by 2 cardiologists
Vandenberk 2018-2(173)	Prospective single centre History of AF	Unclear N= 322 AF prevalence 55%		Unspecified smartphone and PPG app, 60 second recording analyzed by clinicians	Blinded single lead ECG interpretation by 2 cardiologists
Yan 2016(174)	Prospective single centre	Secondary N=85 AF prevalence 29%	Age 72 years (mean); Female 33%	iPhone 6s, CRMA PPG app, 3 X 20 second recordings, irregularity in 1 or more recordings needed to label AF	12 lead ECG

Abbreviations: AF=atrial fibrillation; ECG=electrocardiogram; CRMA=Cardiio Rhythm smartphone application; DCCV=direct current cardioversion; OAC=oral anticoagulation; PPG=photoplethysmography; RMSSD=Root mean square of successive RR differences; ShE=Shannon entropy

3.2.5 Risk of bias

Diagnostic accuracy studies often vary in their results due to how the study was designed and conducted. In view of this each study that underwent the extraction process was assessed for quality and risk of bias by two independent reviewers using the (QUADAS-2) tool.(175) This assesses four key domains: (1) patient selection (2) index test used (smartphone application for AF detection via fingertip or facial PPG signal); (3) reference standard used (1, 3, or 12-lead ECG); and (4) flow and timing. The QUADAS tool is completed in four phases: 1) state the question for review; 2) develop specific guidance that will aid answering this question; 3) Construct a flow diagram for the study; 4) Assess bias and applicability. The first three domains are assessed for risk of bias using signalling questions and concerns regarding applicability questions. **Table 13** below summarises the questions asked for each domain, a score of high, low or unclear was given for reach study for risk of bias and applicability. The review question was to compare PPG based algorithm against an ECG measure for AF detection.

Table 13: QUADAS tool assessment questions

Quality assessment	Review questions
Domain 1: Patient selection	
Risk of bias	Describe methods of patient selection Was a consecutive or random sample of patients enrolled? Was a case control design avoided? Did the study avoid inappropriate exclusions?
Applicability	Describe included patients, prior testing, intended use of index test and setting Any concerns that included patients and setting does not match the review question?
Domain 2: Index test	
Risk of bias	Describe the index test and how it was conducted and interpreted Were index test results interpreted without knowledge of the results of the reference standard? If a threshold was used, was it pre-specified?
Applicability	Describe the index test and how it was conducted and interpreted Is there concern that the index test, its conduct or interpretation differ from the review question?
Domain 3: Reference standard	
Risk of bias	Describe the reference standard and how conducted and interpreted Does the reference test correctly target the condition? Were reference test results interpreted without knowledge of the results of index test?
Applicability	Concerns that the target condition as defined by reference standard does not match review question?
Domain 4: Flow and timing	
Risk of bias	Describe any patients who did not received index test and or reference standard, or who were excluded Describe interval and any interventions between index and reference standard Was there an appropriate interval between index test and reference standard? Did all patients receive a reference standard? Did all patients receive the same reference standard? Were all patients included in analysis?

Two independent reviewers reviewed each study for bias and applicability independently using the above table, following this any discrepancies were reviewed and discussed.

3.2.6 Statistical analysis

Where available weighted means were calculated according to sample size from baseline demographic data. Forest plots were used to visualise sensitivity, specificity, PPV, NPV and diagnostic accuracy for smartphone AF detection. If not directly reported in the study, 2 x 2 contingency tables were constructed from reported values for sensitivity, specificity and corresponding confidence intervals.(145) This was not possible for one study as the number of AF patients was uncertain.(168) Publication bias was assessed using a funnel plot of linear regression of the log odds ratios on the inverse root of the sample size, with asymmetry and/or a non-zero slope coefficient with $p<0.1$ indicative of small study bias. Those studies with sensitivity and specificity alongside confidence intervals were meta-analysed using bivariate mixed-effects regression modelling. A summary receiver operating characteristic (SROC) plot was then created to show the overall diagnostic accuracy of PPG smartphone cameras for AF detection with 95% prediction regions. Heterogeneity was assessed using the I^2 statistic, and visually a bivariate box plot approach was used.(145) Statistical analysis was performed using the statistical program Stata Version 14.2 (StataCorp LP, Texas) and the MIDAS program.(176)

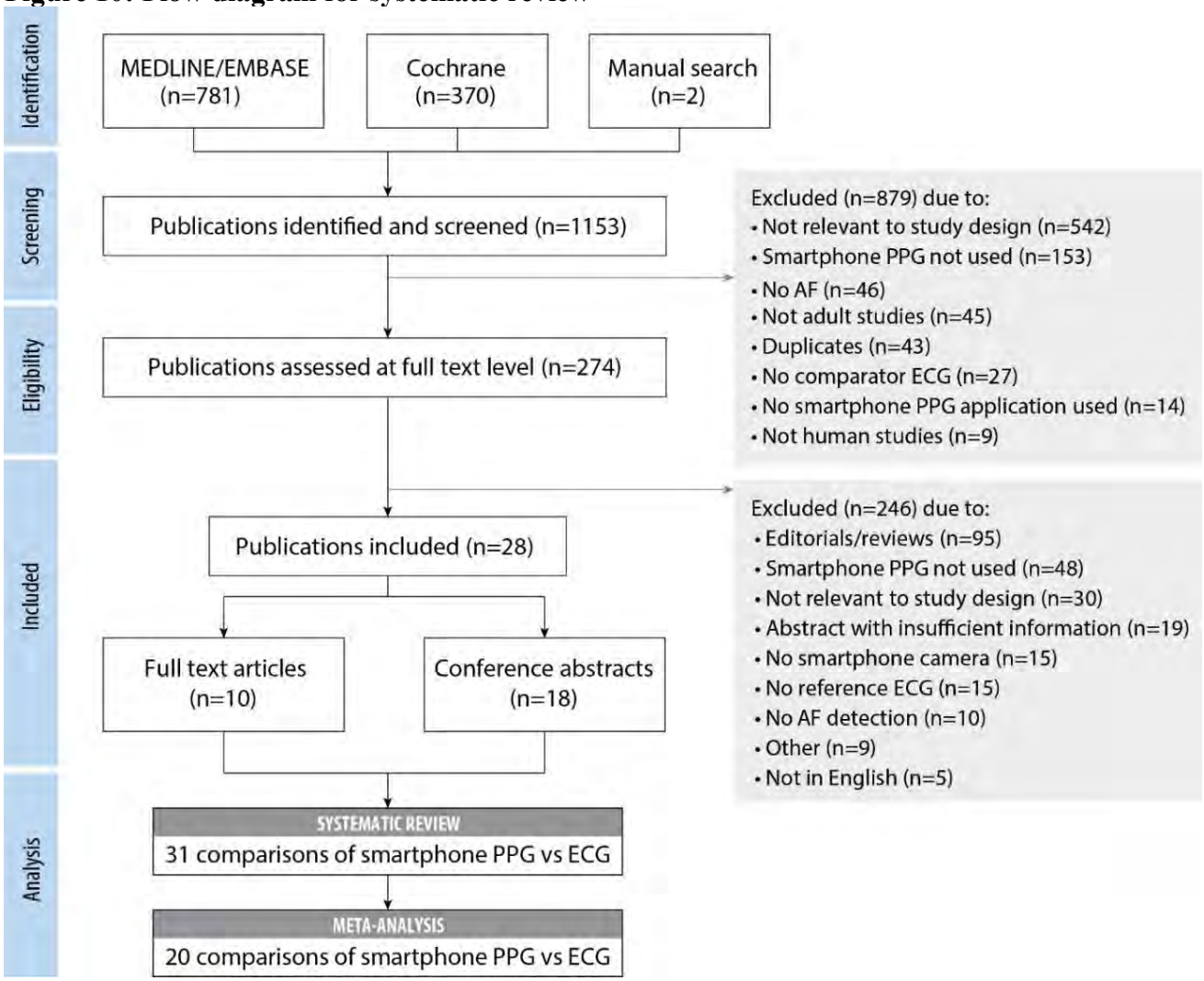
3.3 Results

3.3.1 Design, devices and population

Overall, the search strategy identified 1,153 publications in total, of which 879 were mostly excluded either due to the study design not being relevant (n=542) or smartphone PPG was not used for AF detection (n=153). At full text level, 274 studies were screened and a further 246 were excluded mainly due to many being editorials or reviews (n= 95). A total of 28 studies were included in the systematic review and meta-analysis, of which 10 were full text original research publications and 18 were conference abstracts. **Figure 10** below provides a flow diagram to show the number of studies included and excluded during each stage of the screening and review process.

As mentioned above a total of 28 studies were included in the systematic review. Twenty-five of these studies were prospective with 19 conducted in a single centre (150-152, 156, 157, 159-163, 166-174) and 6 in two or more centres.(143, 148, 153, 155, 158, 167) A total of 16 of the 28 studies were conducted in secondary care(148, 150-153, 155-157, 159-162, 168, 170, 171, 174) and seven in primary care(143, 149, 155, 158, 163, 167, 169), 3 studies did not specify where they took place(164, 172, 173) and 3 used retrospective databases.(154, 164, 165) Of the studies conducted in secondary care, nine included those with a known diagnosis of AF who were due to undergo cardioversion or catheter ablation.(151, 152, 156, 161, 162, 166, 168, 170, 171)

Figure 10: Flow diagram for systematic review



Reproduced by Gill S. et al. 2022

Abbreviations: AF = atrial fibrillation; ECG = electrocardiography; PPG = photoplethysmography

With regards to smartphone devices the commonest one used was the Apple iOS device, which was used in 16 studies (148, 149, 151-157, 160-162, 165, 168, 171, 174), 1 study used an Android smartphone(150), 2 studies used a combination of different devices(166, 170), and 9 did not specify the type of smartphone used.(143, 158, 159, 163, 164, 167, 169, 172, 173) In terms of smartphone PPG applications 8 studies used the Cardiio Rhythm mobile application (CRMA), which was the commonest. The Fibricheck application was the next most frequent application and was used in 4 studies. The Preventicus application was used in 5 studies, and two studies used the PULSESMART application. Nine studies did specify the smartphone application used.(151, 153, 154, 161, 162, 166, 167, 170) In terms of a reference standard, 13 studies used a 12-lead ECG, whilst 11 studies used a single-lead ECG. Four studies used a combination of 12-lead, 3-lead or single lead ECG. Three studies performed PPG and ECG recordings simultaneously.(150, 164, 173)

The systematic review contained a total of 11,404 participants, of these 2,422 (21.2%) had a confirmed ECG diagnosis of AF. The prevalence of AF in each study varied significant ranging from 0.5% to 100%. Where stated the average age ranged from 59 to 77 years, with a weighted mean of 67 years (SD 4.9), the number of women represented in each study ranged from 18% to 59%, with a weighted average of 48.2% (see **table 14** below for baseline demographics of participants).

Table 14: Baseline demographics of participants included in systematic review

Characteristic	Number of studies providing data	Weighted mean (SD) or n (%)	Minimum	Maximum
Age, years	15	67 (SD 4.9)	59	77
Women	18	709 (48%)	18%	59%
Prevalence of AF	26	2422 (31%)	0.5%	100%
Hypertension	7	527 (83%)	59%	90%
Diabetes	8	239 (33%)	20%	37%
Stroke	8	125 (14%)	11%	23%
Heart failure	8	170 (16%)	4%	38%

Abbreviation: AF = atrial fibrillation; SD = standard deviation

3.3.2 Risk of bias and Publication bias

The overall quality of the studies included was found to be low with no study receiving a low-risk score in all four domains of bias and applicability. Scores across all full text publications and conference abstracts were inconsistent despite the latter having limited information. **Table 15 and 16** below provide a detailed description of the QUADAS-2 assessment performed for each study. The most significant concern regarding high risk of bias related to the domains of patient selection and conduct of study. In patient selection, bias was introduced when non-randomised patients were included, as well as those with an increased likelihood of having AF based on their co-morbidities. With regards to conduct of study, data excluded from final analysis and ambiguity around the timing of reference and index tests introduced bias.(145) **Figure 11a** below shows the proportion of studies with low, high and unclear bias scores.

A Deek's funnel plot was used to assess for any evidence of publication bias. Regression analysis was used to show the relationship of study sample size on diagnostic accuracy. This showed a non-zero slope coefficient across all the comparisons and an asymmetrical funnel plot (see **figure 11b)**. Once the largest study was excluded, the p-value for the slope coefficient was found to be 0.06, with a p value <0.10 which suggested that publication bias from small studies with bigger effect sizes was present. A bivariate box plot was used to demonstrate variability, amongst the studies and showed a number of the studies were outside the fence area (see **figure 11c** and **table 17** for numbers linking to each study).

Table 15: Risk of bias and applicability concerns for full text studies

Study author and year	QUADAS 2-Risk of Bias				QUADAS 2-Applicability concerns			
	Patient selection	Index test: PPG	Reference standard: ECG	Flow and timing	Patient Selection	Index test: PPG	Reference standard: ECG	Domains with high risk
Brasier 2019 (148)	High risk	Unclear	Low risk	High risk	Low risk	Low risk	Low risk	2/7
Chan 2016 (149)	Unclear	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	1/7
Fan 2019 (150)	High risk	Unclear	Low risk	High risk	Low risk	Low risk	Low risk	2/7
McManus 2016 (152)	High risk	Low risk	Unclear	High risk	Low risk	Low risk	Low risk	2/7
McManus 2013 (151)	High risk	High risk	Unclear	Low risk	Low risk	Low risk	Low risk	2/7
Mutke 2020 (153)	High risk	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	2/7
Poh 2018 (154)	Unclear	Low risk	High risk	High risk	Unclear	Low risk	Low risk	2/7
Proesmans 2019-1 (155)	High risk	Unclear	Low risk	High risk	Low risk	Low risk	Low risk	2/7
Rozen 2018 (156)	High risk	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	2/7
Yan 2018 (157)	High risk	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	2/7

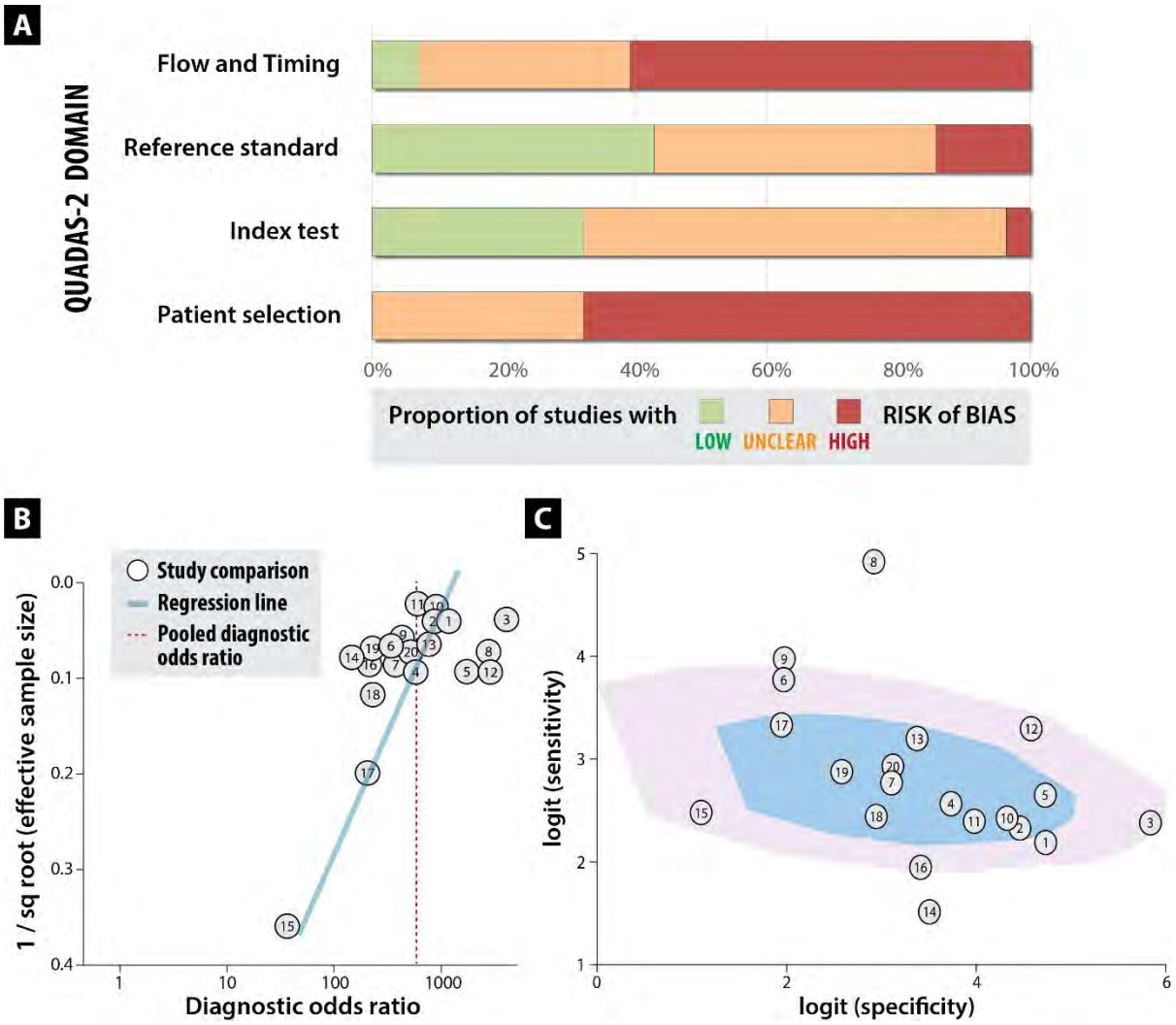
Abbreviations: PPG = Photoplethysmography; ECG = electrocardiography

Table 16: Risk of bias and applicability concerns for conference abstracts

Study author and year	QUADAS 2-Risk of Bias				QUADAS 2-Applicability concerns			
	Patient selection	Index test: PPG	Reference standard: ECG	Flow and timing	Patient Selection	Index test: PPG	Reference standard: ECG	Domains with high risk
Grieten 2018-1 (143)	High risk	Unclear	Low risk	High risk	Low risk	Low risk	Low risk	2/7
Grieten 2018-2 (158)	High risk	Unclear	High risk	High risk	Low risk	Low risk	Low risk	3/7
Karim 2017 (159)	High risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk	1/7
Kuan 2018 (160)	Unclear	Unclear	Unclear	Unclear	Unclear	Low risk	Low risk	0/7
Maitas 2012-1 (161)	High risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk	1/7
Maitas 2012-2 (162)	High risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk	1/7
Mortelmans 2017 (163)	High risk	Unclear	High risk	High risk	Low risk	Low risk	Low risk	3/7
Mutke 2019 (164)	Unclear	Unclear	Low risk	Low risk	Low risk	Low risk	Low risk	0/7
Napolitano 2015 (165)	High risk	Unclear	Unclear	High risk	Low risk	Low risk	Low risk	2/7
Proesmans 2019-2 (166)	Unclear	Unclear	High risk	Unclear	Unclear	Low risk	Low risk	1/7
Proesmans 2018 (167)	High risk	Unclear	Unclear	High risk	Low risk	Low risk	Low risk	2/7
Rozen 2017 (168)	High risk	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	2/7
Siu 2016 (169)	Unclear	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk	0/7
Smeets. 2019 (170)	Unclear	Unclear	Unclear	High risk	Low risk	Low risk	Low risk	1/7
Vaid 2015 (171)	High risk	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	2/7
Vandenberk 2018-2 (173)	Unclear	Unclear	Low risk	Unclear	Low risk	Low risk	Low risk	0/7
Yan 2016 (174)	Unclear	Low risk	Unclear	Unclear	Low risk	Low risk	Low risk	0/7

Abbreviations: PPG = Photoplethysmography; ECG = electrocardiography

Figure 11: Assessment of bias, weighted funnel plot and bivariate box plot



Reproduced from Gill S. et al. 2022

11A: Overall risk of bias based on QUADAS-2 criteria; 11B: weighted funnel plot with the vertical axis displaying the inverse of the square root of the effect sample size (1/root (ESS), and the horizontal axis displaying the diagnostic odds ratio. 11C bivariate box plot with the inner shaded area representing the median distribution for sensitivity and specificity, and the outer area representing the 95% confidence bound. ESS=effective sample size

Table 17: Summary of sensitivity, specificity and accuracy of PPG

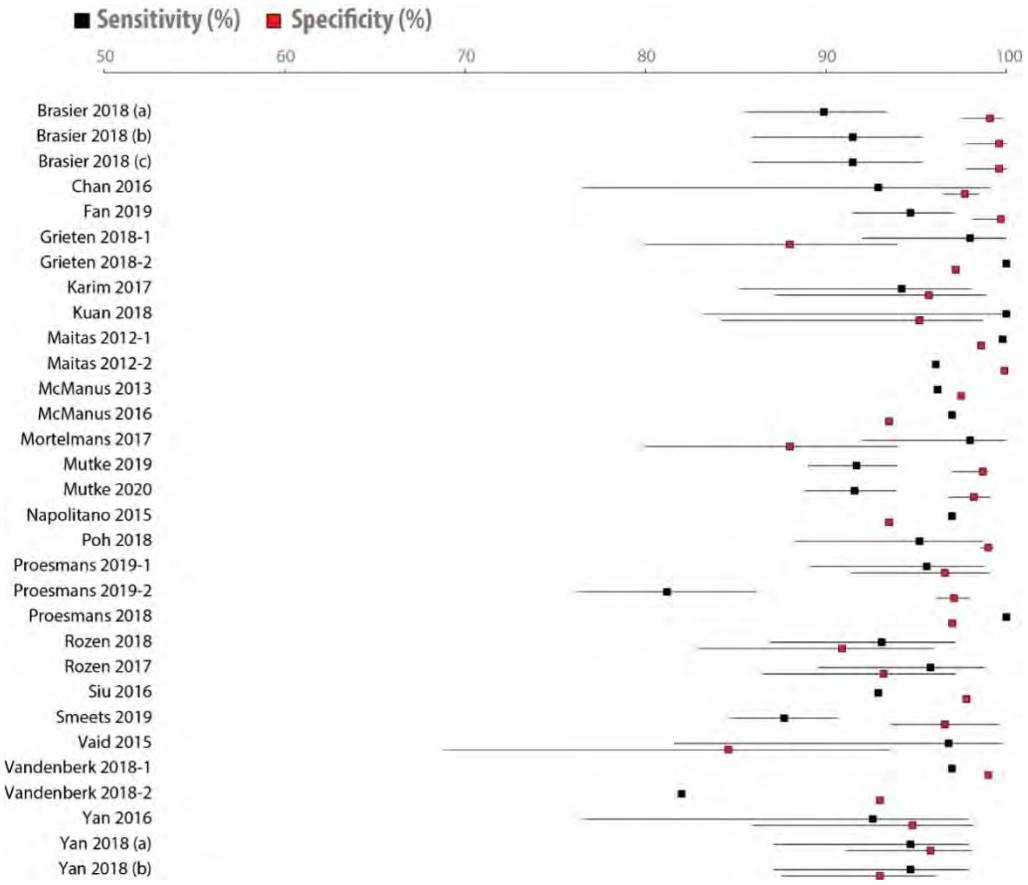
Study	Meta-analysis comparison number	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Accuracy % (95% CI)
Brasier 2019a (148)	1	90 (86-93)	99 (98-100)			89
Brasier 2019b (148)	2	91 (87-95)	99 (97-100)			78
Brasier 2019c (148)	3	92 (86-95)	100 (98-100)			61
Chan 2016 (149)	4	93 (77-99)	98 (97-99)	54 (38-68)	100 (99-100)	
Fan 2019 (150)	5	95 (92-97)	100 (98-100)	100 (98-100	96 (93-98)	98 (96-99)
Grieten 2018-1 (143)	6	98 (92-100)	88 (80-94)	88 (82-93)	98 (92-99)	
Grieten 2018-2 (158)		100	97			
Karim 2017 (159)	7	94 (85-98)	96 (87-99)			
Kuan 2018 (160)*	8	100 (83-100)	95 (84-99)			
Maitas 2012-1 (161)		100	99			99
Maitas 2012-2 (162)		96	100			98
McManus 2013 (151)		96	98			97
McManus 2016 (152)		97	94			95
Mortelmans 2017 (163)	9	98 (92-100)	88 (80-94)			93 (89-96)
Mukte 2019 (164)	10	92 (89-94)	99 (97-99)	98 (92-96)	94 (92-96)	
Mutke 2020 (153)	11	92 (89-94)	98 (97-99)			95
Napolitano 2015 (165)		97	94			95
Poh 2018 (154)	12	95 (88-99)	99 (99-99)	73 (65-79)	100 (100-100)	
Proesmans 2019-1 (155)	13	96 (89-99)	97 (91-99)	63 (61-65)	100 (100-100)	
Proesmans 2019-2 (166)	14	81 (76-86)	97 (96-98)	95 (94-96)	89 (87-91)	91 (89-93)
Proesmans 2018 (167)		100	97			97
Rozen 2018 (156)	15	93 (87-97)	91 (83– 96)	92 (86-96)	92	92
Rozen 2017 (168)		96 (90-99)	93 (87-97)	93 (86-97)	96 (90-99)	96
Siu 2016 (169)		93	98			
Smeets 2019 (170)	16	88 (85-91)	97 (94-100)	98 (97-99)	77 (72-82)	90 (88-92)
Vaid 2015 (171)	17	97 (82-100)	85 (69-94)	83 (67-93)	97 (83-100)	
Vandenberk 2018-1 (172)		97	99			
Vandenberk 2018-2 (173)		82	93	92	84	
Yan 2016 (174)	18	93 (77-98)	95 (86-98)			
Yan 2018a (157)	19	95 (87-98)	96 (91-98)	92 (84-96)	97 (93-99)	95
Yan 2018b (157)	20	95 (87-98)	93 (88-96)	88 (80-93)	97 (93-99)	94

Abbreviations: PPV = positive predictive value; NPV = negative predictive value; CI = confidence interval

3.3.3 Detection of AF

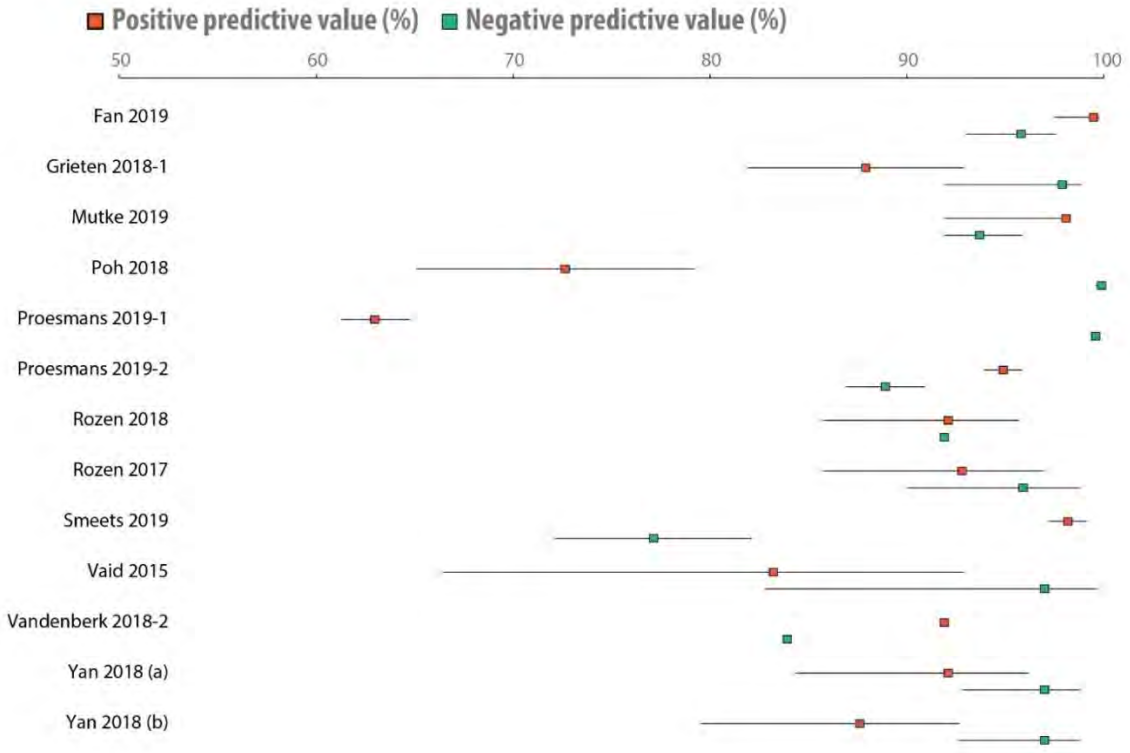
A summary of diagnostic accuracy measures (sensitivity, specificity, PPV and NPV, overall accuracy) are provided in **table 17** above. Of the 28 studies included in this review, 31 comparisons for AF detection using the PPG smartphone applications compared to a conventional ECG were provided. From these comparisons, 29 used fingertip smartphone PPG and 2 facial smartphone PPG. The values for diagnostic accuracy varied greatly and included the following ranges: sensitivity from 81% to 100%; specificity from 85% to 100%; PPV 54% to 100%; and NPV 77% to 100%. Forest plots were constructed in Microsoft Excel to compare sensitivity, specificity, PPV and NPV with corresponding confidence intervals (where available) for each of the 28 studies (see **figures 12 and 13 below**). Eighteen comparisons for accuracy were reported and this ranged from 61% to 99%.

Figure 12: Forest plot of the sensitivity and specificity of smartphone PPG vs ECG



Reproduced from Gill S. et al. 2022 (145)
Horizontal lines represent the 95% confidence interval (where available) for each comparison. Sensitivity is represented by the black square and specificity by the red square.

Figure 13: Forest plot of positive and negative prediction of smartphone PPG vs ECG

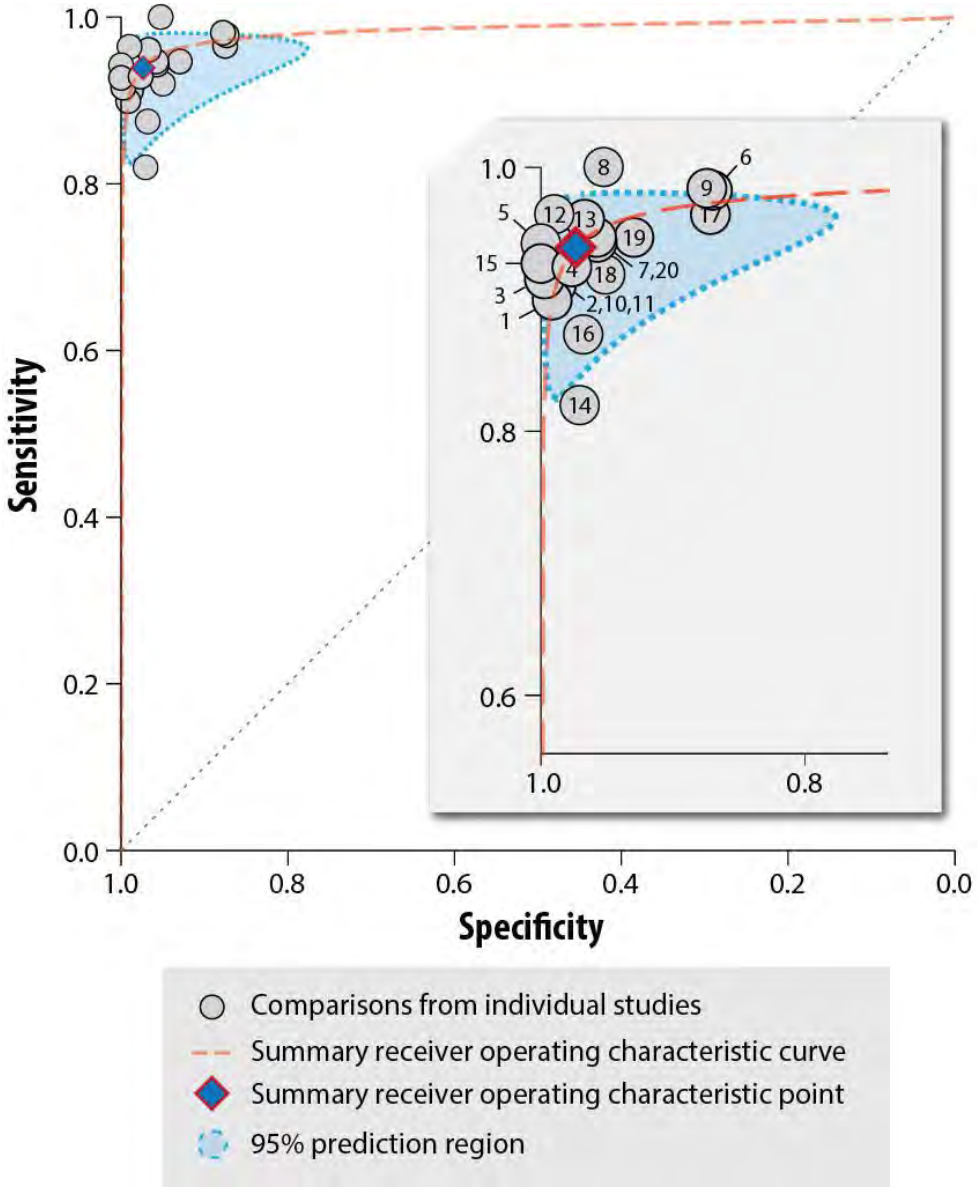


Reproduced from Gill S. et al. 2022 (145)
Horizontal lines represent the 95% confidence interval (where available) for each comparison. Positive predictive value is represented by the red square and negative predictive value by the green square.

3.3.4 Meta-analysis

A meta-analysis was conducted of 20 comparisons of AF detection from a total of 17 studies, giving a population (n) of 5,561, of which 1,674 had AF. The pooled sensitivity was calculated to be 94% (95% CI 92-95%), and demonstrated significant heterogeneity with an I^2 statistic of 49.6% and a p-value of 0.01. The pooled specificity was calculated to be 97% (95% CI 96-98%), again this showed heterogeneity with an I^2 statistic of 85.3%; and a $p<0.01$, indicating significance. Overall, the area under the ROC curve was calculated as 0.98 (95% 0.97-0.99), and again demonstrated substantial heterogeneity across the studies with an I^2 of 98% and a significant p-value of <0.0001 (See **figure 14 below**).

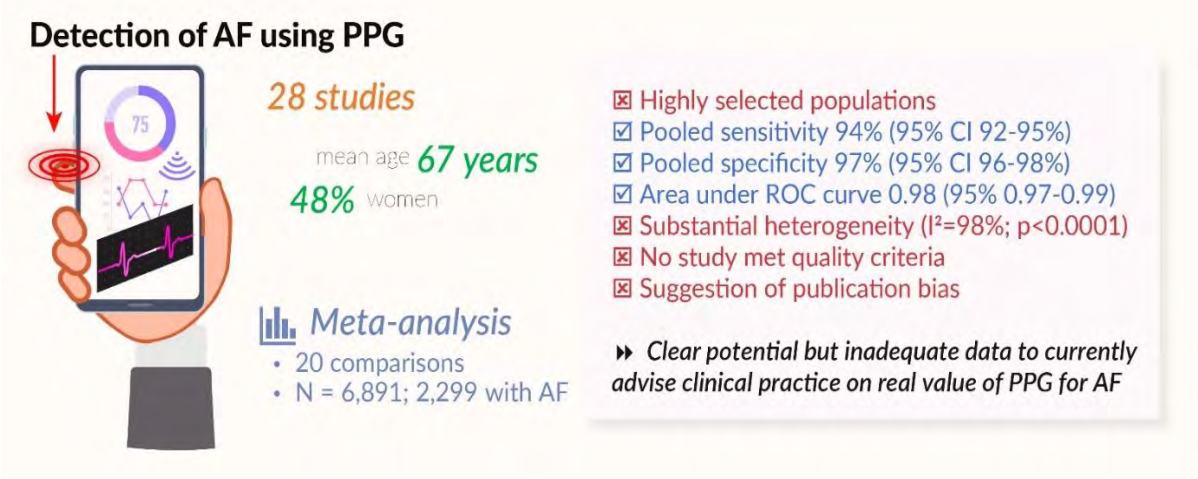
Figure 14: Summary receiver operator characteristic plot



Reproduced from Gill S. et al. 2022 (145)
This ROC (receiver operating curve) includes all comparisons in the meta-analysis (see Table 15 for numbers linking to each study comparison) with summary receiver operator characteristics.

The findings of this systematic review and meta-analysis were published in the *Heart* journal. A graphical summary of the main findings was constructed and can be seen in **figure 15** below.

Figure 15: A graphical summary of the main findings within this systematic review and meta-analysis



Reproduced from Gill S. et al. 2022 (145)

3.4 Discussion

3.4.1 General findings

This review identified that smartphone-based PPG technology is capable of AF detection, but at present there is a lack of evidence available to demonstrate its true accuracy when compared to a gold standard test (ECG). Individually most of the smartphone-based PPG applications yielded high diagnostic accuracy for AF detection, with a sensitivity and specificity for AF detection at meta-analysis of 94% (95% CI 92-95%) and 97 % (95% CI 96-98%) respectively. However, these appear to be unrealistically high as the studies were small, mostly single-centre and subject to bias. An in-depth assessment of each study for risk of bias identified that patient selection, study design, flow and timings introduced a high risk of bias. This was due to selecting patients more likely to have AF detected; and a lack of information regarding study design, especially with regards to the timings of index and reference test. There was also evidence for the presence of publication bias with asymmetry suggesting studies with negative findings were unlikely to undergo publication.(145)

In general, commercially available smartphone applications were used to detect AF using PPG technology. Each patented algorithm used for AF detection acted as a source of heterogeneity.(177) The majority of studies lacked algorithm transparency and information regarding how data is assessed for quality and noisy signals filtered out, as well as the precise algorithms used to determine AF from other arrhythmias such as atrial flutter, tachycardia or ectopics was missing making replication and validation of these smartphone PPG applications difficult. To summarise, these findings suggest larger well-designed studies free from commercial influence with clear algorithms comparing PPG with a gold standard

reference are required to assess the true diagnostic accuracy of smartphone PPG technology for AF detection.

3.4.2 The need for AF detection

As identified in my introduction, AF is a heterogenous condition and clinically it can manifest with symptoms or it can be asymptomatic and is therefore only identified incidentally during routine medical review.(178) Those individuals with undiagnosed, usually silent AF can present with an ischemic stroke which may be the first manifestation of AF, and is only found incidentally from a routine admission ECG or during subsequent monitoring.(179) With its growing prevalence, there is a clear healthcare priority to increase our ability to offer an effective community AF screening programme (145, 180). This is particularly important as undiagnosed AF is associated with significant complications such as ischaemic stroke which can have a substantial impact on patient quality of life. Effective screening offers the possibility of providing early detection and is vital in those at risk of stroke where timely commencement of oral anticoagulation can prevent these sequelae, and reduce mortality and morbidity associated with stroke.(177)

AF is sporadic in nature, hence there is a genuine requirement for non-invasive, patient activated screening techniques that can be used to detect AF over a long period of time in a scalable manner.(145) Currently AF screening mostly consists of opportunistic pulse palpation or prolonged monitoring using an ambulatory ECG or medical devices, which have a finite monitoring period.(16, 145, 181) Therefore, the development of new technology that is easy to use, and promotes patient participation in monitoring their own heart rhythm, would provide more opportunity to detect AF.(145) Over recent years the use of smartphones has grown rapidly especially in those aged over 65 years. In 2012 smartphone

ownership in this cohort was only 3%, this has rose substantially to 65% in 2020 (source:statista.com).(145) With smartphones now becoming omnipresent, in most homes across the world, the potential to increase non-invasive community AF screening worldwide seems plausible, especially in the older generation who have the most to gain.

3.4.3 Other forms of photoplethysmography

In this review I specifically looked at the value of fingertip and facial PPG, but other forms of PPG AF detection are also available. An example is the Apple smartwatch which applies a proprietary algorithm on recorded tachograms to determine AF(141). The Apple Heart study recruited 419,297 participants and used smartwatch-based PPG for intermittent monitoring of AF. Of those recruited 2161 (0.52%) received an alert for an irregular pulse. Out of 450 participants who wore a 7-day ECG patch, 34% (n=153) received a new diagnosis of AF, yielding a PPV for AF detection of 84%.(116, 145, 178, 182) The Huawei heart study recruited 187,912 participants to wear smartwatches and smart bands. Of these, 0.23% (n=424) received a notification indicating an irregular pulse was detected and follow up was required. The PPV for AF detection in this study was found to be higher at 92%.(183) The more recent Fitbit study, a prospective remote clinical trial evaluated the effectiveness of a PPG algorithm for identifying undiagnosed AF during simultaneous single lead ECG monitoring. A total of 455,699 participants were recruited and used either Fitbit fitness trackers or smartwatches that incorporated this algorithm. An irregular heart rhythm detection notification was received in 4728 participants (1%), of which 1057 had single lead ECG patch monitoring and 340 (32%) had AF. During ECG monitoring 225 participants had another irregular heart rhythm notification and AF was confirmed on ECG in 221, giving a PPV of 98.2% (95% CI, 95.5%–99.5%). (184) These studies highlight that PPG technology has the potential to provide a non-invasive and long-term method for AF detection. Of note,

smartwatch participants were more likely to be younger, as seen in the Apple Heart study where 52% of the participants who wore the smartwatch were below the age of 40, this dropped to just 6% in those aged over 65 years.(145, 182) AF screening in a younger population can lead to low detection rates, and there is a risk of false positive results which may lead to further unnecessary tests and use of healthcare resources. In addition, the clinical impact of identifying silent AF in younger individuals without any risk factors for thromboembolism is currently unknown and may cause unnecessary anxiety. Despite the growing number of smartphone users, utilizing PPG technology for the purposes of screening still requires the consumer to invest in expensive technology and this may disadvantage those of low socioeconomic status.(145)

In summary this review has highlighted the need for well-designed real-world studies, with a low risk of bias to establish the true diagnostic accuracy of smartphone PPG, as sensitivity and specificity are not likely to be as high as those found in this systematic review and meta-analysis in an unselected population. RCTs should be independent of device or algorithm manufacturers and sufficiently powered for endpoints that include stroke and cost-effectiveness if clinical efficacy is to be determined.(145) Additionally, commercial providers of smartphone applications, need to provide greater transparency regarding what algorithms are used for AF detection. With regards to its place in the diagnostic pathway, further work is required to evaluate the role smartphone PPG technology will have alongside conventional AF screening methods.

3.4.4 Limitations

Although PPG technology has the potential to be used as a patient initiated screening tool for AF detection there are limitations that were identified during this review. Firstly, PPG

technology (hardware and software) is rapidly evolving, therefore data from both published articles and conference abstracts were included. Unlike published articles these abstracts have not undergone a vigorous peer review process. Study authors were contacted to obtain further information however no response was obtained. As a result, the full information regarding baseline demographics, detailed methodology, smartphone device and application and statistical analysis was often missing. (145) This lack of information made ascertaining study quality and bias particularly challenging for abstracts.

Secondly, this review only evaluated smartphone fingertip and facial PPG and excluded PPG technology in smartwatches which has since been evaluated in the Apple Heart study. The inclusion of smartwatches and fitness trackers which support an AF detection algorithm may have provided a better reflection of the true accuracy of PPG for AF detection. Thirdly, in terms of study design retrospective, prospective and case control studies were included in the analysis. The majority of these studies were based in a population that was likely to have a higher proportion of individuals with undiagnosed AF. These findings may provide an explanation as to why diagnostic accuracy was generally overestimated. At the time of writing there were no randomised clinical trials that evaluated smartphone PPG technology in an unselected population, emphasizing the need further trial-based research. With regards to PPG signal interpretation, most studies used commercially available algorithms which were not transparent about the AI algorithms used to identify AF, making validation of results difficult. Study heterogeneity was present alongside evidence of publication bias, as many of these studies were supported by commercial companies that may be less likely to publish non-positive studies to preserve intellectual property. Overall study quality was poor, with selection bias having a significant impact on the generalisability of our findings.(145)

3.5 Conclusion

AF is one of the commonest cardiac arrhythmias encountered by healthcare professionals, however due to its paroxysmal nature and heterogenous presentation in many patients, the detection of AF can be challenging using conventional methods. With the number of smartphone users growing, PPG technology can provide a non–invasive, patient-led screening tool for AF detection. PPG technology has the potential to offer large scale AF screening with high sensitivity and specificity. However, at present the current evidence-base is inadequate consisting of small, low-quality studies with significant bias and therefore cannot provide clinicians with guidance on their use for AF detection. The integration of smartphone PPG technology as an AF screening tool in its infancy and continues to evolve. Alongside robust large-scale clinical trials comparing PPG AF detection with reference standards, further development in the hardware, software applications and algorithms are required to ensure this technology can be validated and can provide clinicians with reliable information.

Recognising the potential smartphone PPG technology has to improve AF detection, this review provided the impetus to incorporate PPG technology within my research. As such an independent academically-developed PPG smartphone application was chosen and incorporated into the RATE-AF wearables sub-study (see chapter 4 for methods and 5 for results).

4 Design and rationale of the RATE trial and wearables substudy

4.1 Rationale for RAte control therapy evaluation in permanent Atrial Fibrillation (RATE-AF) trial

Individuals who develop AF and HF suffer a poorer quality of life, increased hospitalisations and death.(55, 131, 185, 186) Whilst nearly all patients require rate control to slow down their heart rate and improve their AF associated symptoms, some will also need restoration of sinus rhythm (rhythm control) to improve their symptoms. Rhythm control is often employed in patients with paroxysmal or persistent AF. This treatment strategy has been evaluated in numerous trials, and a meta-analysis of 25 RCTs showed that rhythm control strategies were associated with a higher rate of adverse events, but there was no difference in mortality when compared to rate control.(187) A drawback of the rhythm control strategy is that recurrence of AF is high, and these patients often still require rate control to treat the symptoms arising from AF paroxysms. In addition, 40-50% of AF patients are unsuitable for rhythm control as they have permanent AF and require long term rate control.(185, 188)

The commonest drugs therapies used for rate control are beta-blockers, rate-limiting calcium channel antagonists, digoxin and amiodarone.(62, 64, 189, 190) In the presence of HFrEF, calcium channel antagonists are contraindicated. Due to its side effect profile, amiodarone is only used in patients where rate control with combination therapy has failed.(16) This leaves beta blockers and digoxin as the main pharmacological agents commonly prescribed for rate control. Current guidance from both ESC and National Institute for Health and Care Excellence (NICE) recommends beta blockers as a first line agent due to their rapid onset of action, as well as the prognostic benefit (reduction in all-cause mortality) seen in those with HFErF and sinus rhythm.(16, 62, 191) Previous observational studies have found an increased mortality rate in AF patients on digoxin.(192-194) However, these are likely to

have been affected by prescription and selection bias, as in clinical practice digoxin is prescribed to sicker patients with multiple co-morbidities. This was confirmed by a systematic review that showed digoxin did not have a positive or negative effect on mortality.(64, 191, 195)

In summary rate control is used universally in almost all patients with any form of AF. Despite its widespread use there is a lack of evidence available to guide clinicians and health care professionals on the choice of appropriate therapy resulting in variations in prescribing practice. Current evidence is limited to small observational studies with limited insight into the effects of treatment on mortality and patient quality of life.(191)

The RATE-AF trial was designed as the first RCT directly comparing the use of digoxin and beta-blocker therapy for rate control in patients with permanent AF (deemed not suitable for rhythm control) and symptoms of heart failure. The aim was to assess long term benefits of each treatment in terms of patient reported quality of life, functional capacity, cardiac function in the form of systolic and diastolic function and biomarkers of treatment response.(191)

4.1.1 RATE AF trial design

The RATE-AF trial was a prospective, randomised open label blinded endpoint (PROBE) clinical trial evaluating the effectiveness of digoxin and beta blockers as initial heart rate control therapy. A PPI (patient and public involvement) team were involved in the design of the trial with the aim of improving patient quality of life. . The primary outcome at 6 months was patient-reported quality of life with an emphasis on the physical component summary score of the Short form (SF)-36 health survey.(191) Secondary outcomes at six and twelve months included:

- Additional generic and AF-specific patient-reported quality of life scores such as, the SF-36 global and domain-specific scores, AF Effect on Quality of Life (AFEQT) overall score and the European Quality of Life Five Dimension (EQ-5D-5L) summary index and visual analogue scale.
- Changes in left ventricular ejection fraction (LVEF) and diastolic function (E/e')
- Functional assessment in the form of change in EHRA classification, 6-minute walking (6MW) test and heart rate
- Change in NT-pro-BNP levels as a surrogate marker for cardiac strain at six months(191)

4.1.2 Trial criteria

Patients were recruited from primary and secondary care; Recruitment began in December 2016 and ended in October 2018. The trial inclusion and exclusion criteria are described below in **tables 18** and **19**.

Table 18: Inclusion criteria for RATE AF trial

Inclusion criteria
Adult patients aged 60 years or older
Permanent AF, characterised (at time of randomisation) as a physician decision for rate-control with no plans for cardioversion, anti-arrhythmic medication, or ablation therapy
Symptoms of breathlessness (NHYA Class II or more)
Able to provide written informed consent

NHYA = New York Heart classification; AF = Atrial fibrillation

The exclusion criteria were minimised to reflect usual clinical practice. The list of exclusion criteria can be found in **table 19** below.

Table 19: Exclusion criteria for RATE AF Trial

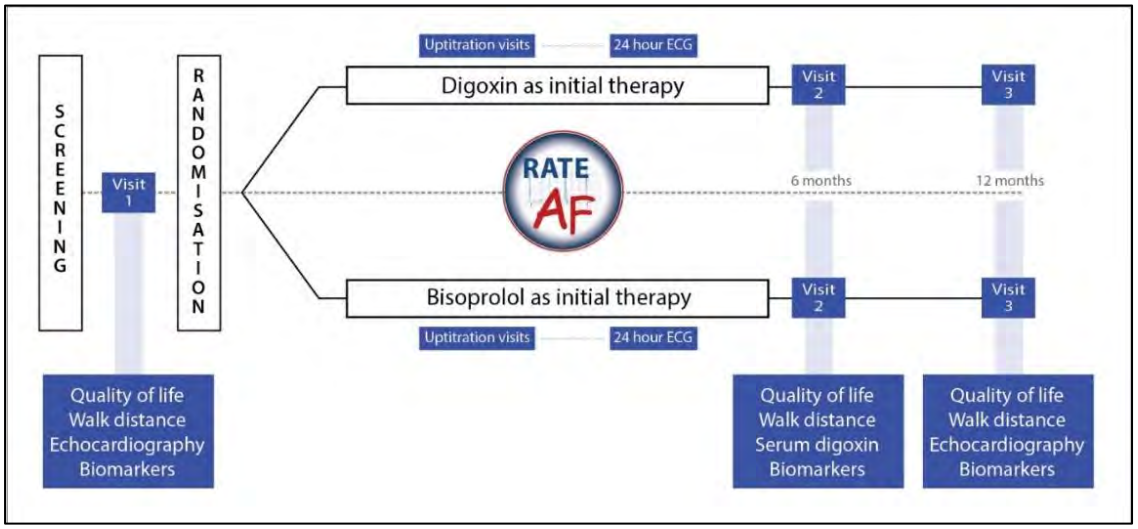
Exclusion criteria
Established indication for beta-blocker therapy (e.g. myocardial infarction in last 6 months)
Known contraindications for therapy with beta blocker or digoxin e.g. a history of severe bronchospasm that would preclude use of beta-blockers, or known intolerance to these medications.
Baseline heart rate less than 60 beats per minute.
History of second- or third-degree block.
Supraventricular arrhythmias associated with accessory conducting pathways (e.g. Wolff-Parkinson-White syndrome) or a history of ventricular tachycardia or fibrillation.
Planned pacemaker implantation (including cardiac resynchronisation therapy), pacemaker-dependent rhythm or history of atrioventricular node ablation.
Decompensated heart failure (evidenced by need for intravenous inotropes, vasodilators or diuretics) within 14 days prior to randomisation.
A current diagnosis of obstructive hypertrophic cardiomyopathy, myocarditis or constrictive pericarditis
Received or on waiting list for heart transplantation.
Receiving renal replacement therapy.
Major surgery, including thoracic or cardiac surgery, within 3 months of randomisation.
Severe, concomitant non-cardiovascular disease (including malignancy) that is expected to reduce life expectancy

4.1.3 Enrolment and Randomisation

Figure 16 (trial schema) provides an overview of the RATE AF recruitment process together with the investigations carried out at each visit.

Participants were randomised in a 1:1 ratio to either Digoxin 62.5- 250mcg or bisoprolol 1.25-15mg once daily. . Randomisation was provided using a computer-generated programme at the Birmingham Clinical trials Unit (BCTU). A minimisation algorithm was used to ensure balance between both arms with regards to important clinical variables stratifying for baseline EHRA (class 1/2A and 2b/3/4) and gender. Trial arm allocation was concealed until the patient was successfully recruited.

Figure 16: RATE-AF trial schema



Adapted from Kotecha D. et al, 2017 (191)
ECG=electrocardiogram.

4.1.4 Trial treatment

The pharmacological agents Bisoprolol and Digoxin were the investigational medicinal products for the trial. At randomisation, participants were allocated to open label treatment with either drug. Digoxin is a cardiac glycoside that originated from the foxglove plant. Its primary mechanism of action is through its direct inhibition of the sodium potassium adenosine triphosphatase (Na+/K+/ATPase) within the myocardium. This results in an increase in the amount of intracellular calcium available for contractile proteins to bind with, which in turn increases the contractility of the heart. It also reduces the heart rate by

stimulating vagal nerve activity which slows down conduction through the atrioventricular and sinus node. Participants were initially started on 62.5-125mcg daily, and were up-titrated to 125-250mcg daily (maximum trial dose of 250mcg daily). A single loading dose of 250-500mcg was prescribed to digoxin naïve participants.(191, 196)

Bisoprolol is a beta 1 selective adrenoreceptor blocker that competitively inhibits norepinephrine and epinephrine from binding to the beta 1 adrenoreceptor found in cardiac conduction tissue. This results in a reduction in both resting and exercise heart rate, as well as an initial drop in resting and exercise cardiac output with little change in stroke volume. Trial starting doses were 1.25-5mg, and doses were up-titrated to 10mg daily (maximum trial dose was 15mg daily).(196)

4.1.4 Schedule of assessments

The trial was embedded into usual care within the National Health Service, with participants attending formal visits at baseline, 6-months and 12-month. Endpoints acquired at these visits were patient-reported quality of life, NTproBNP, symptoms and functional capacity using mEHRA and NYHA class, 6MW test, heart rate and adverse events. At their baseline visit following the consent and randomisation process, patients were asked questions regarding their medical history and a physical examination was conducted. A series of patient reported quality of life questionnaires (SF-36, AFEQT and EQ-5D-5L) as described above were completed. This was followed by an ECG, 6MW test recording distance walked and peak heart rate, echocardiogram and biomarker assessment. Following their baseline visit, patients attended for up-titration visits so rate control medication could be optimised to obtain a heart rate of ≤100 beats per minute. At their 6-month follow up participants were asked to attend for a review. At this appointment all quality of life questionnaires were

completed, compliance to trial medication assessed, and an ECG, 6MW test and serum biomarkers were conducted. At their final 12 month visit the same tests were performed alongside a final echocardiogram (see **figure 16**).(191, 196) A summary of the schedule for assessments can be found in **table 20**. As part of the research team, I was involved in carrying out the medical history, physical examination and prescribing the allocated medication to participants once they arrived for their baseline, up-titration, 6-month and 12-month visits.

Table 20: Schedule of assessments for RATE AF trial

Procedures	Baseline visit	Up-titration visit	6-month visit	12-month visit
Assessment of eligibility	X			
Informed consent	X			
Medical history review	X			
Medication review	X	X	X	X
Physical examination	X	X	X	X
Quality of life questionnaires	X		X	X
Functional and cognitive assessments	X		X	X
Echocardiogram	X			X
12 lead ECG	X		X	X
6-minute walk test	X		X	X

ECG=electrocardiogram

4.2 RATE-AF main trial results

A total of 160 patients successfully completed the RATE AF trial, each arm of the trial had 80 patients. .(196) The average age was 76 years (standard deviation (SD), 8 years), and 46% were female. At baseline most patients had moderate (modified EHRA Class 2b) or severe (modified EHRA class 3) symptoms from their AF which affected their daily activity. From a heart failure perspective, the NHYA mean was 2.4 (SD, 0.6) and the median NTproBNP was 1057 pg/ml (IQR 744-1522pg/ml), 19% of patients had an LVEF of <50% on echocardiogram.(196) In terms of the primary outcome there was no significant difference in the SF-36 quality of life physical component score (adjusted mean difference, 1.4: 95% CI -1.1 to 3.8; p-value =0.28).(196) With regards to the 17 secondary outcomes, 16 including heart rate control and change in LVEF showed no significant difference. At 6 months there was a statistically significant difference in modified EHRA class with 53% of patients in the digoxin group reporting a two-class improvement, compared to just 9% in the bisoprolol group (adjusted odds ratio (OR) 10.3, 95% CI 4-26.6; P<0.001).(196) At 12 months the NTproBNP level in the digoxin group was significantly lower than the beta-blocker group (ratio of geometric means 0.77, 95% CI 0.64-0.92; P=0.005). Digoxin appeared to be better tolerated with fewer patients reporting adverse events compared to beta-blockers (29 events vs.142 events; P < 0.001).(196, 197)

4.3 Design and rationale of the RATE-AF wearables substudy

4.3.1 Introduction

One of the secondary objectives of the RATE-AF trial was to assess the effect of bisoprolol and digoxin on functional capacity in terms of physical activity and heart rate control. Conventional measures for these parameters usually consist of periodic measurements taken during clinical visits. For heart rate these are: manual pulse palpation or apex beat rate measured for 1 minute, or ECG derived heart rate. Traditional physical activity tests include 6-minute walk distance test and International Physical Activity Questionnaire (IPAQ) score. The issue with these types of measurements is that they can only be taken during clinical trial visits or at fixed time points during the trial and therefore only provide data for a snapshot of when they were taken. They are unable to explore the variations in heart rate and physical activity that can occur in a patients' normal environment during their day-to-day activities. This is important as it can provide more information about patient symptomatology and impact on quality of life, as well as rate control. When compared to conventional trial outcomes, consumer wearable devices offer the potential to collect physical activity and heart rate data continuously, in a patients' day-to-day environment, often with minimal input.

A summation of patient physical activity and continuous heart rate monitoring are both endpoints where wearable devices can provide substantially more information than conventional periodic review of physical capacity or heart rate. (109, 198, 199) Wearable devices have the potential to provide superior information on heart rate variance and any interaction with physical activity over longer periods of time. This approach would have advantages in patients with concomitant medical disorders such as AF and HF as they produce symptoms that have a significant impact on patient quality of life and general well-being. Hence, they are ideally suited to monitoring of physiology in real-world settings.

Despite its potential, there is very limited robust data to know whether consumer wearable devices can have a place in the medical management of patients with CVD.

The primary goal of the sub-study was to explore whether physical activity and ambulatory PPG heart rate assessment from consumer wearable devices (sensor data) could ascertain if there was difference in heart rate control and physical activity in the 2 arms of the RATE-AF trial. Previous studies such as the Apple Heart study and the Fitbit Heart study have demonstrated that PPG heart rate from wearable devices can be used in clinical trials.(116, 184)

Another objective was to use an AI approach to ascertain if sensor data can prediction clinical progression with the same accuracy as current clinical assessments for heart rate (ECG) and physical activity (6-minute walk test), thus aid clinical management. To obtain more detailed information, RATE-AF trial participants were invited to join an optional sub-study where wrist worn wearable devices connected to a smartphone were provided to assess ambulatory heart rate and physical activity. I designed and set up a sub-study integrating the use of wearable technology to assess ambulatory heart rate and physical activity in patients already enrolled within the RATE-AF trial. To obtain this detailed information RATE-AF study participants were given the opportunity to wear a commercially available wrist–worn device connected to a smartphone that was capable of continuous ambulatory monitoring of heart rate and physical activity. As mentioned earlier, an independent PPG application developed by academics at the University Medical Center, Utrecht, the Netherlands for AF detection was incorporated into this sub-study. Patients were asked to engage in two smartphone applications; one to assess their heart rhythm for the presence of AF; another interactive application which provided an educational resource on AF and allowed patients to record their AF symptoms and quality of life score.

4.3.2 Ethics

The technology for this research was supported by funding from the European Union Innovative Medicines Initiative BigData@Heart programme, 116074 (https://www.bigdata-heart.eu/).(200) The work was a collaboration between the University of Birmingham cardAIc team at Health Data Research UK Midlands, University College London, Utrecht Medical Centre (the Netherlands), and the Hyve, an IT support company based in the Netherlands. An amendment was made to the RATE-AF trial protocol and subsequently approved by the Research Ethics committee.

4.3.3 Wearables sub-study aims

The overall aim of this sub-study was exploratory, I set out to investigate the following:

- Feasibility of integrating wearable technology into an on-going RCT in the older population.
- Whether continuous heart rate PPG data from the Fitbit wearable can be used to evaluate the effectiveness of beta-blockers and low dose digoxin for long term rate control.
- Whether beta-blockers or low dose digoxin have an effect on physical activity measured as total step count on the Fitbit device.

Whether heart rate and physical activity measured from the Fitbit device are comparable to conventional measurements of ECG heart rate and 6-minute walk test in predicting changes in patient outcomes such as NHYA class.

4.3.4 Device selection and data type

With clear aims and hypothesis, the next step in the design of the wearables sub study was to consider the type of data that needs to be collected and the how this would be done. Consideration was given to the RATE AF trial cohort, which consisted of older patients (mean age 76 years) with multiple co-morbidities. Wearable technology has the potential to encourage older patients to actively engage in their own healthcare. They can also aid clinicians in the management of chronic conditions and response to treatment with active monitoring of symptoms and health parameters such as heart rate, blood oxygen saturations, physical activity, and blood pressure.(201) At the same time, a number of factors must be considered that may limit older patients in the trial from engaging in wearable technology these include:

- Reduced manual dexterity
- Slower processing and impaired memory
- Reduced visual acuity
- Previous technology experience
- Functional ability of individual
- Motivation for use
- Device aesthetics (not bulky), and ease of use (simplicity)

When selecting the most appropriate wearable technology the above factors were considered, alongside the need to have a device that was capable of collecting heart rate and physical activity data continuously. In general, older patients are more likely to engage with wearable technology if it is not a medical device.(201) Therefore, for the purposes of this sub-study we opted for an easy-to-use wrist worn commercially available wearable device as it is non-

invasive, painless, and extremely low risk. They are similar in size and dimension to a watch, manufactured primarily in rubber, and are marked as lifestyle devices, therefore not denoting participation in a research study, reducing stigmatisation and eliminating the possibility of identifying an individual as a research participant. The commonest commercially available brand of wrist worn device that is capable of monitoring physical activity and heart rate is a Fitbit. It is also the most used consumer device in clinical research with 952 publications at the time of writing.(202, 203) The Fitbit charge 2 model was selected for this sub-study due to its comfortable strap and large screen. Each patient was supplied with a new device that was set up in advance and a charger. There were no exclusions related to age or prior proficiency with information technology. The Fitbit charge 2 device could measure the following parameters:

Physical activity

- Movement in the form of step count calculated from a three-axis accelerometer
- Physical exercise using heart rate data
- Floors climbed using changes in barometric pressure combined with steps
- Distance covered using stride length

Sleep

- Estimated using a combination of heart rate, heart rate viability and movement
- Sleep stages (light, deep and rapid eye movement)
- Duration
- Wake episodes

Heart rate

'Purepulse' is a patented PPG technology used in the Fitbit charge 2 device that uses LEDs on the skin surface to monitor changes in blood volume and thereby continuously estimate heart rate. This device is a fitness tracker and is not capable of AF detection. (204) In terms of the accuracy of heart rate monitoring, a study by Benedetto et al evaluated the heart rate accuracy of the Fitbit charge 2 simultaneously against ECG criterion.(204) The study showed that individual heart rate assessments could be underestimated by up to 30 beats per minute. Higher intensity exercises appeared to produce great inaccuracies in heart rate. The actual Fitbit algorithm for PPG HR is not available in the public domain. Hence, the cause of these inaccuracies remains unknown. The study was limited to 15 participants and so the results are interpreted with caution.(204)

4.3.5 Smartphone selection

A smartphone capable of supporting the Fitbit charge 2 device and the PPG application was required so participants could utilise the full functionally of their wearable device and engage in the apps provided. At the time of the sub-study Apple smartphones were not compatible with the data collection platform that was required. Therefore, an android smartphone was chosen as it supported passive data collection through the Remote Assessment of Disease and

Relapses (RADAR) base platform, as well as the specific PPG application smartphone camera requirements.

Two Android smartphone models were selected that both supported the RADAR base platform. The LG K8 and Samsung galaxy A6 smartphone. The camera application was tested on both smartphones to assess if they could support the PPG application and the findings are detailed in the **table 21** below.

Table 21: Android smartphone Photoplethysmography performance

Smartphone model	Finger positioning over camera	Camera colour bit expression	Performance of camera
LG K8	Good	Satisfactory	Poor (17-27 fps)
Samsung galaxy A6	Good	Good	Very good (30-46fps)

*Frames per second

For the PPG app to provide a satisfactory signal that can be analysed using ML algorithms, at least 30 frames per second (fps) were required. During testing the LG K8 smartphone range was only 17fps under moderate stressful conditions and 27fps under perfect conditions, below the 30fps required for a good PPG signal to be produced. The Samsung Galaxy A6 smartphone camera performance was very well underly stressful conditions producing a minimum of 30fps, which rose to 46fps under optimal conditions. With these findings in mind the Samsung galaxy A6 smartphone was selected for the sub-study.

Each participant was provided with a Samsung galaxy A6 smartphone (with a pre-paid mobile contract) that was already set up with the appropriate software and could be used as a regular mobile phone. Participants were able to make calls to all UK landlines and mobiles and access the internet. Software on the phone connected to the wearable device via the Fitbit application and the RADAR app so that data could be passively collected regarding

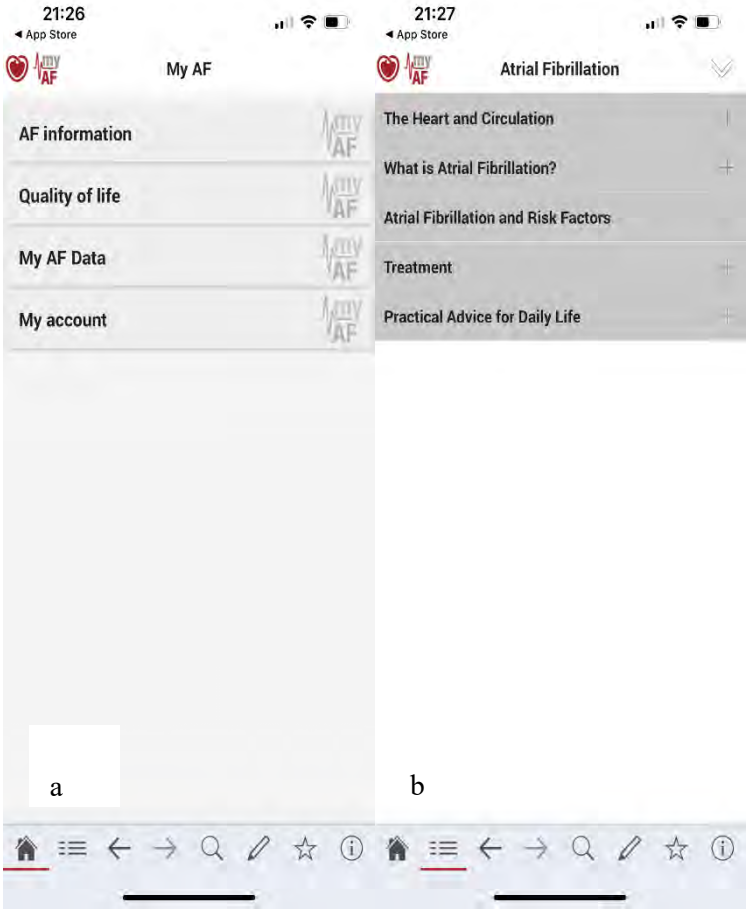
heart rate, physical activity in the form of steps and phone usage. The global positioning system denoting a participant's location was anonymised in a way that does not reveal actual geographic location. Similarly, the content of text messages, or phone numbers that have been contacted were not collected.

4.3.6 Software applications

Three software applications were pre-installed onto the Galaxy A6 smartphone prior to each participant's first appointment. The RADAR application was downloaded to enable passive remote data collection from the Fitbit device and smartphone. It also incorporated the PPG application where participants were asked to place their fingertip over the camera to record a PPG signal. The Fitbit application was installed to activate the device, allow for temporary data storage on the smartphone and to enable users to engage with the interface. The dashboard provided participants with a summary of their activity, and the ability to sync their device with the smartphone provided. Prior to the sub-study start date an application was made to Fitbit for intraday access. This granted access to participant data in a fine granular form for all parameters in a 24-hour period. Of particular interest was the continuous heart rate data at 1-second intervals, and physical activity in the form of steps at 1-minute intervals.

A third application called the 'my AF' application was installed to provide an educational resource for patients. The app was collaboratively designed by the CATCH ME Consortium (Characterizing Atrial fibrillation by Translating its Causes into Health Modifiers in the Elderly), the ESC Task Force on AF and the European Heart Rhythm Association. The app provided patients with information regarding AF pathology, symptoms, and management strategies (see **figure 17** below).

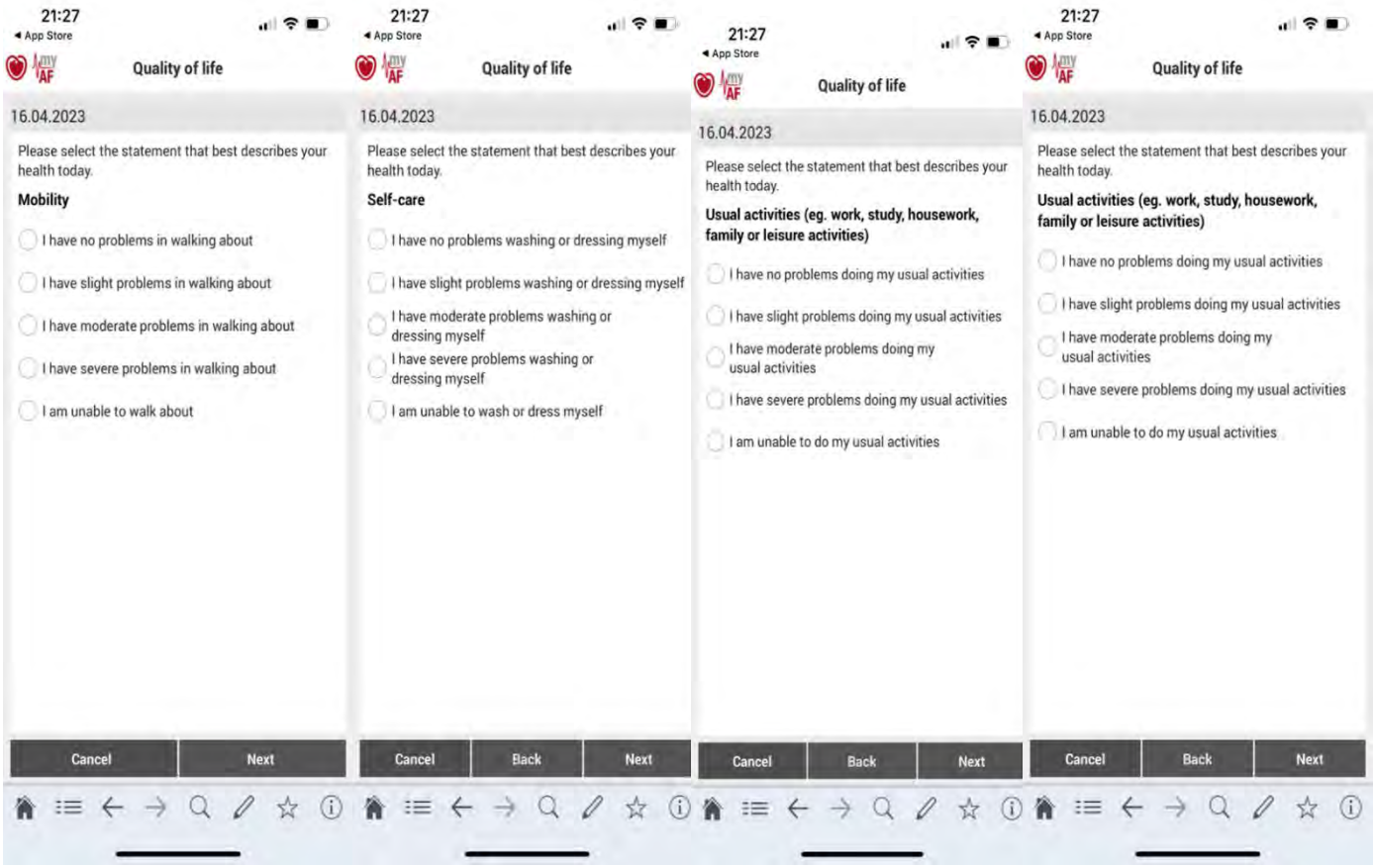
Figure 17: The My AF application



Screenshot taken directly from 'My AF' online application

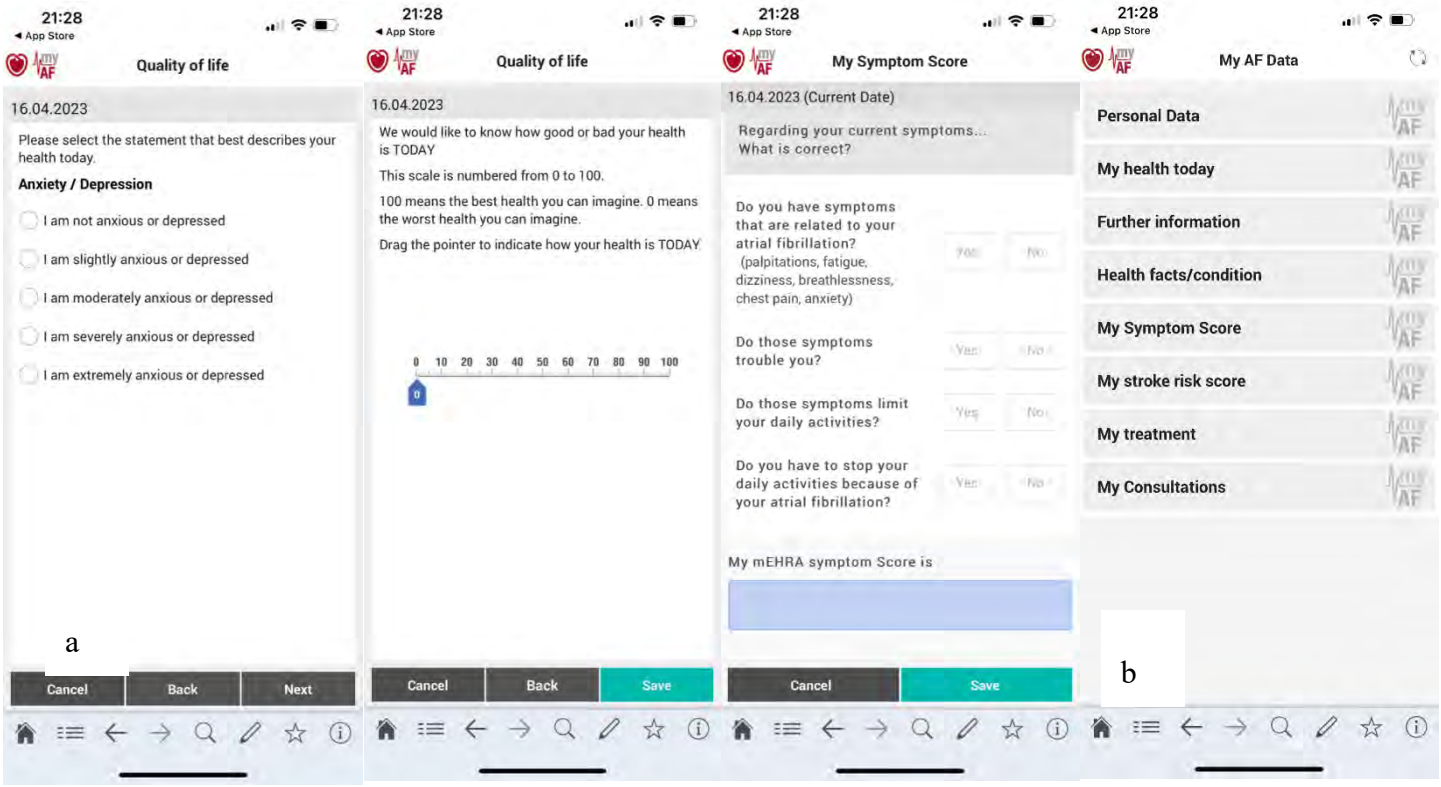
The My AF application home screen is shown in **figure 17a**. Once the application has been downloaded from the app store of the smartphone, an account for each participant was made and baseline data was completed by me. Participants where then shown the features of the application starting with the AF information screen (**figure 17b**). This was followed by demonstrating how to record a quality-of-life score (**figure 18** and figure **19a**), and an EHRA AF symptom score (figure **19b**)

Figure 18: Quality of life scoring on My AF app



Screenshot taken directly from 'My AF' online application

Figure 19: Quality of life and symptom scoring on My AF app



Screenshot taken directly from 'My AF' online application

4.3.7 Remote Assessment of Disease and Relapses (RADAR) base platform

The RADAR base platform originated from the Innovative Medicines Initiative project called Remote Assessment of Disease and Relapse – Central Nervous System (RADAR-CNS). This was a collaborative project with University College London and its academic and industry partners who developed a novel way of collecting sensor data from wearable devices and smartphones. Physiological data was collected passively and actively in the form of questionnaires and tests. The focus was to evaluate if sensor data could be used to predict relapse in those suffering from epilepsy, multiple sclerosis or depression. Since then, the RADAR base platform has gone on to become an open-source platform that can integrate, collecting and storing continuous real-time sensor data from several data streams. The

platform is able to remotely collect data passively and actively. Passive data in the form of physiological parameters, movement from accelerometer sensors, location, audio, calls and texts are collected using the passive remote monitoring technologies (pRMT) application. Real time data from smartphone and device sensors is sent directly through the pRMT app via cloud storage to the RADAR base. Active data collection involves an interaction with participants in the form of questionnaires and is collected using the active remote monitoring technologies (aRMT) application. To facilitate the wider use of the RADAR base for remote monitoring, the platform was released under open-source Apache licence in January 2018.

For this sub-study integration of wearable and smartphone data streams, real-time remote passive data collection, secure data transmission, storage and access for retrospective analysis was required. (See **figure 20**). A RADAR base platform and management portal was deployed and maintained by the Hyve. The Hyve used a RADAR based stack to deploy the platform in a cloud-based service hosted by Amazon Web Services (AWS) in the AWS Europe (London) region. The Hyve also created a management portal to manage users, participants and data streams.

Figure 20: RADAR platform deployment in the wearables sub-study

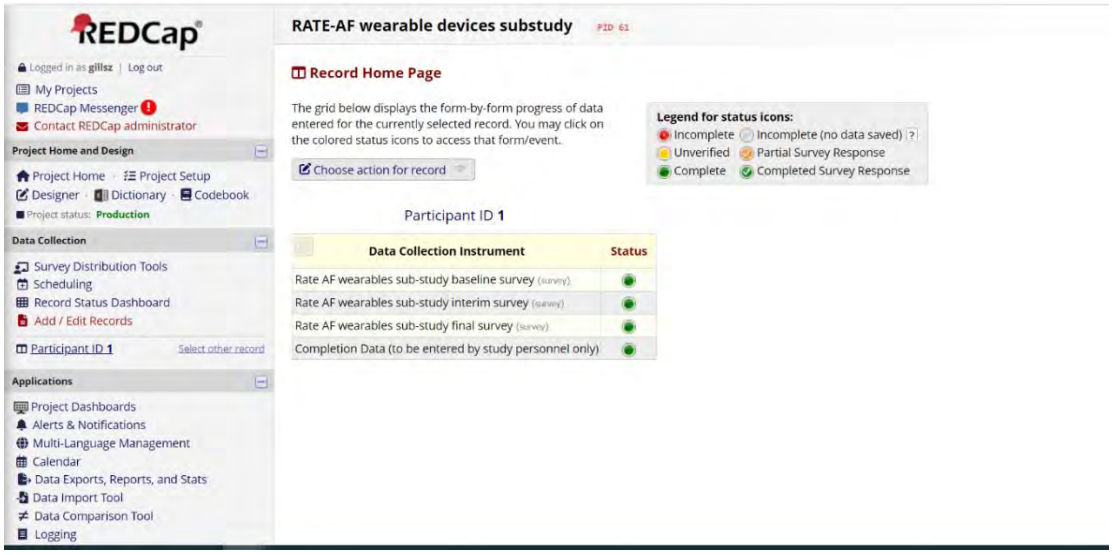
4.3.8 Data handling

All the information collected was pseudo-anonymised and kept confidential. Data collected via the wrist-worn wearable device and smartphone was encrypted, uploaded to a secure server and was not identifiable by patient name. Data from the Fitbit could be temporarily cached on the smartphone until an appropriate Wi-Fi or mobile data connection was available and was then automatically deleted from phone memory. The Hyve assisted with the hosting of the RADAR platform and made periodic checks to ensure sensor data was being received from registered participants. Data collected form the devices was securely stored in the established RADAR base platform and then downloaded by the University of Birmingham data science team to a local secure server.

4.3.9 Research Electronic Data Capture (REDCap)

Clinical data from each sub-study visit (baseline, interim and final) was entered on to a secure web-based application called Research Electronic Data Capture (REDCap). REDCap was developed by Vanderbilt university and allows researchers to create online project surveys and databases for data collection and storage. It has an easy-to-use interface that runs on most operating systems and does not require any software installation. Project access is restricted to individual users which provides an audit trail. The project creator has full access to the project created and can add, amend or modify survey questions and grant other users access as required for that particular project. There is also built in security options such as the de-identification of data for export using Microsoft excel, pdf and common statistical packages such as Stata.(205) For this sub-study I was responsible for setting up the sub-study project, creating electronic data collection forms using the REDCap system hosted by the University of Birmingham (see **appendix 3, 4, and 5**), and for data entry and management for all participants. **Figure 21** below shows the REDCAP interface for a participant in the sub-study.

Figure 21: The REDCAP interface



Screenshot taken directly from REDCAP web-based application

4.3.10 Patient recruitment

Patients were enrolled directly from the RATE AF trial and so met the inclusion criteria outlined earlier. At the time of the sub-study starting in December 2018, all patients had completed their baseline assessments (as a minimum) and recruitment to the trial was complete. A database with all RATE AF trial patients was reviewed by me to assess eligibility to the sub-study. Those who had completed the trial were excluded and those who had less than 4 weeks remaining were deemed unsuitable due to insufficient time for adequate data collection. I contacted all suitable patients by telephone and provided an overview of the sub-study and what it entailed. In brief I explained the RATE AF trial was assessing how medications affect quality of life, heart rate and physical activity. The wearables sub-study was an optional extra part of the trial which would allow us to obtain further information regarding heart rate and physical activity. Patients were informed that they would be asked to wear a wrist-worn device that would collect information regarding heart rate, and movement. They would also be provided with a smartphone that would

connect to the device and could be used for personal use for the duration of the study. Those patients who were interested or who agreed to participate were given a new 45-minute appointment slot at the BCTU, Queen Elizabeth hospital, Birmingham, if no foreseeable trial appointments were already in place. If patients had upcoming up-titration or six-month review appointments, then these were used instead. The first appointment gave patients the opportunity to read the specific PIL (see **appendix 6**) written by the RATE-AF PPI Team, ask questions and sign the RATE-AF optional consent form (see **appendix 7**), if enrolment was agreed.

4.3.11 Appointment schedule

Below is an outline of the appointment schedule

Pre-appointment

- Fitbit Charge 2 device and Samsung Galaxy A6 smartphone fully charged
- Email account set up on smartphone and RADAR pRMT, Fitbit and ESC My AF applications downloaded

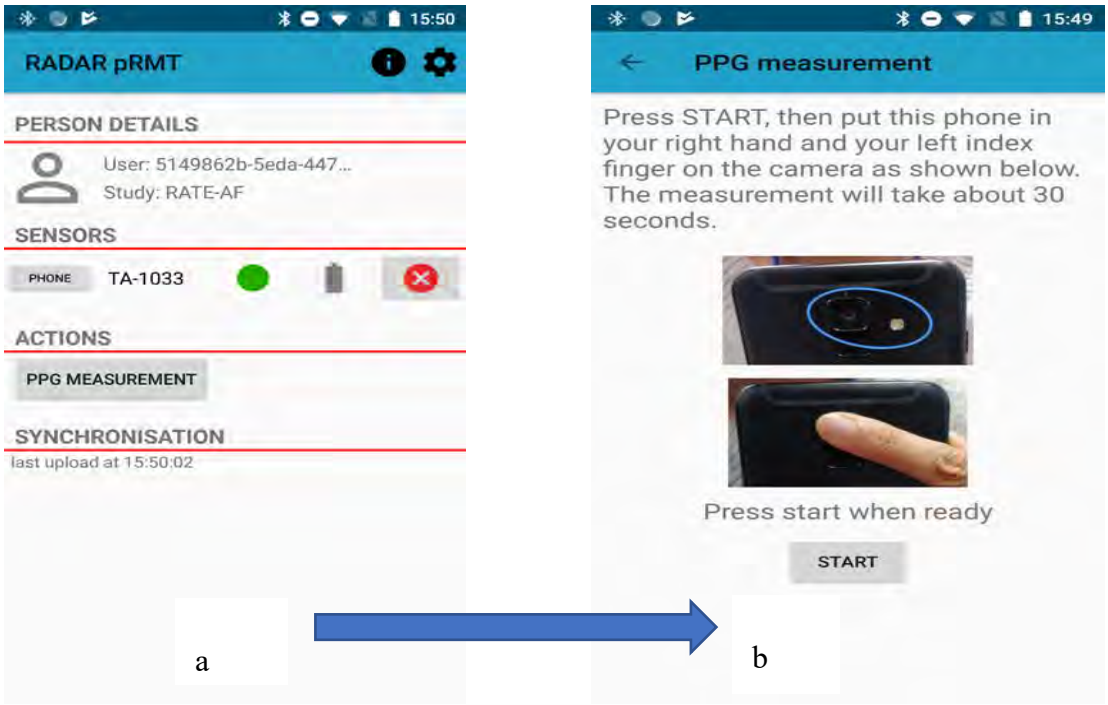
First appointment

- Patients given the opportunity to read the PIL explaining the optional sub-study
- Sub-study consent form signed
- 12-lead ECG recording made
- NTproBNP serum blood sample taken
- Patients were asked to complete the three Quality of life questionnaires (SF36, AFEQT and EQ 5D-5L summary index and visual analogue scale

- Information regarding height, weight, EHRA classification, NYHA class, hand dominance, wrist Fitbit worn, previous smartphone ownership and access to WIFI were recorded on the REDCap baseline survey.
- The Fitbit Charge 2 device was paired to the Samsung Galaxy A6 smartphone by entering email account created specifically for this sub-study. Participant details including name, height, weight and gender were entered by myself and the device set up was customised to the participant's requirements (larger font for reduced visual acuity).
- At this point participants were shown how to use simple functions on the smartphone such as making phone calls and sending messages, as well as a demonstration on how to charge it. They were asked to carry the phone throughout the day so that data from the Fitbit could periodically sync with the smartphone Fitbit application. Participants were asked to keep the phone switched on throughout the trial period and to charge it every night.
- Participants were then shown how to charge the Fitbit device (a fully charged battery would usually last 5 days), and were advised to wear the Fitbit device continuously, only removing it for: showering, bathing, swimming and charging.
- The next step was to create an anonymised RADAR identification on the management portal, which generated a QR code that was scanned by the RADAR pRMT app to begin passive monitoring of sensor data. **Figure 22a** below shows the screen that is seen once this process has been completed and a RADAR account has successfully been created.
- Under the actions tab of the RADAR pRMT (**figure 22a**) screen is the PPG application. Patients were asked to click on PPG measurement which then produces the screen seen in **Figure 22b**, where they can make a pulse recording using PPG via

the smartphone camera. Here the patient presses START and then places their left index finger over the camera as demonstrated on the images shown in **Figure 22b**. A 30 second recording is made, and this process is repeated twice each week.

Figure 22: RADAR Application screen



22a & 22b screenshot taken from Samsung A6 smartphone of RADAR application. Figure 22b that was incorporated into the application was taken by me

With consent the ESC My AF application was set up and the participant's demographics and medication history completed. The application was also connected to the AF manager application. Participants were shown how to record their quality-of-life score using the EQ-5D-5L questionnaire which measures their general health on that day, and their EHRA symptom score for AF. Participants were asked to complete this twice weekly.

- At the end of the first appointment a follow up telephone appointment was arranged in the next 7-10 days

End of first appointment

- Quality of life question responses, NTproBNP value were recorded on the REDcap baseline survey and a pdf of the ECG was uploaded.

Second appointment

- Telephone follow up was provided to assess the following:
    - Able to charge Fitbit and wear continuously
    - Able to charge smartphone each night and carrying throughout the day
    - Connected to internet if available
    - 4G data symbol present if not connected to WIFI
    - Bluetooth switched on
    - Fitbit syncing with smartphone
    - Assess if patient able to do PPG recordings and quality of life and EHRA scores twice weekly
- Any concerns or difficulties were navigated and an attempt to solve them over the telephone was made.
- If this failed an appointment (clinic or home) was made to troubleshoot and re-educate

Third appointment

- At 1 month a face to face follow up review was conducted
- The points listed above were addressed and recorded on the REDCap interim survey.

Final appointment

- The final sub-study appointment coincided with the end of the RATE AF trial
- An ECG was recorded and an NTproBNP serum sample taken
- Participants were asked the following questions:
    - if they carried their smartphone regularly in the last 7 days
    - if they had engaged in the smartphone functions
    - if they felt the smartphone was easy to use
    - If they were recording their PPG measurements twice weekly
    - If they were recording their AF symptoms and quality of life score twice weekly
    - If use of my AF app improved their understanding of AF
    - If use of my AF app improved compliance to AF medication
    - How many days Fitbit device was worn for in last 7 days
    - How many hours Fitbit device worn for each day
    - If the Fitbit device improved motivation in terms of physical activity
    - If the Fitbit device increased their physical activity
    - This information alongside end of trial quality of life questionnaires was recorded in the RedCap final survey
- All equipment was returned at the end of the trial

4.3.12 Assessing data collection and pre-processing

Contact was made with participants by me throughout the course of the sub-study. As detailed above each patient received a follow up telephone call after the introductory training provided at the first appointment to ensure compliance and correct use. The most challenging

part of the sub-study was ensuring adequate data collection. Several steps were taken to ensure data collection was optimised:

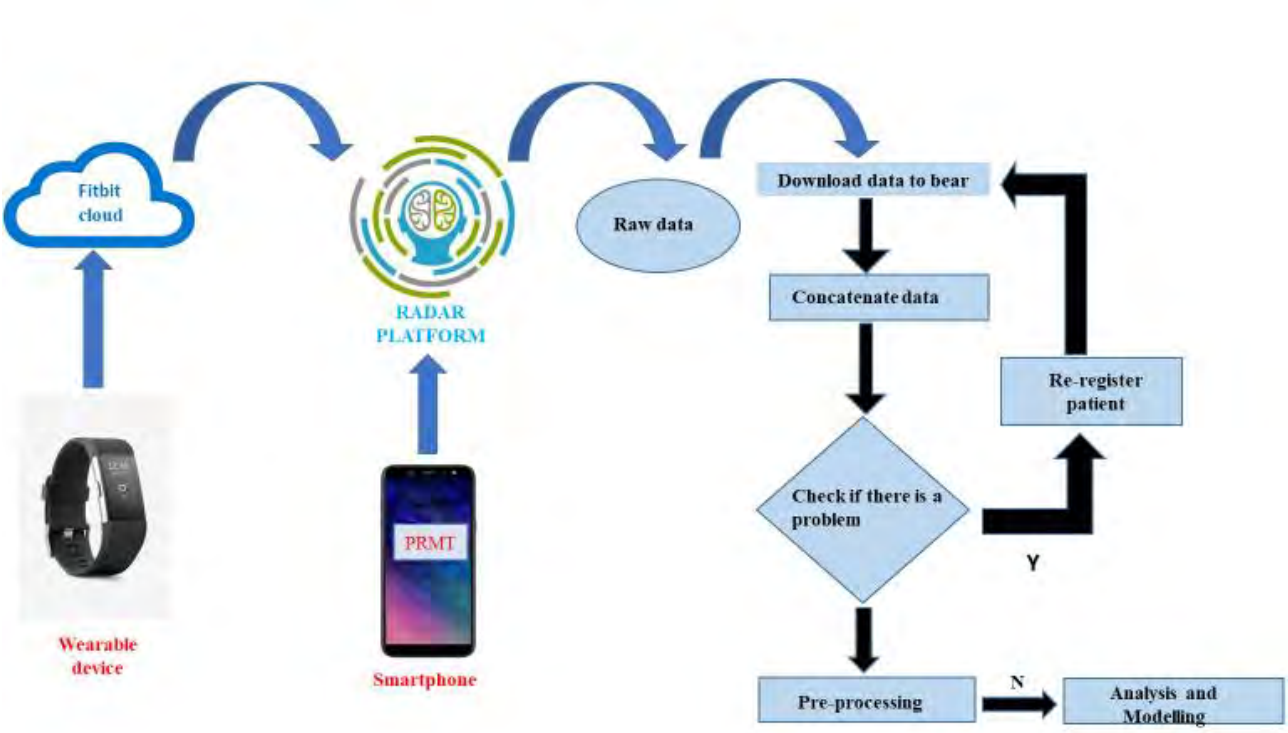
1. A monthly review of mobile data by each participant was provided by the University of Birmingham's telecommunications department. Each participant's data usage was reviewed and if mobile data usage was approaching 4G, a home visit or follow up appointment was made to have their SIM switched to 8G.

2. Weekly checks were made by the Hyve to ensure data for each participant was coming through and the platform was storing data without any issues. During this sub-study there were occasions where data upload would pause due to: maintenance of the cloud server, contractual issues, and docker problems. All were easily resolved once identified and data could be retrospectively stored once the platform was back up and running.

3. Periodic review of data capture was conducted with a data scientist responsible for ensuring data was successfully downloaded from the RADAR platform to a local server. **Figure 23** below provides a schematic of the data review process that was carried out. Data capture for each patient was reviewed and if data was not being received the patient was contacted to see if their problem could be resolved over the telephone. If resolution could not be provided remotely then a follow up appointment was made to review the device issue at the BCTU. In patients that had difficulty travelling to the clinical trials unit, home visits were arranged. The commonest technical concerns usually resulted from accidently switching off mobile data, and a temporary hitch in the RADAR platform which sometimes required smartphones to be re-registered with the platform to avoid loss of data streaming from the device. Participants received further appointments as needed to answer any questions, maintain engagement and address any technical concerns RADAR platform to the University

of Birmingham server it was concatenated and as described above data capture for each patient was reviewed. Following this, the heart rate and step count data was processed into 1-minute intervals to form time-series data. The zero-property of the data was preserved where patients were stationary and missing data was not imputed.

Figure 23: Schematic of data review process



4.3.13 Limitations for sub-study design

Throughout the design of this sub-study there were several limitations noted. The first was that the sub-study started with only 10 months remaining on the RATE AF trial. At this point the majority of patients were deemed unsuitable for enrolment as they had either finished the trial or, had inadequate time remaining. In addition, as the trial was well underway, I was unable to provide written instructions to participants on the twice weekly monitoring (My AF and PPG application) that was required to ensure they were fully engaged in the sub-study. This proved quite tricky for some participants who were less technically minded and required

regular, sometimes weekly follow up to improve understanding and compliance. Secondly, at the time of conducting this study the RADAR and PPG application were only compatible with Android smartphones which limited the type of smartphones that could be used. This was further limited by the camera requirements which had to be of a minimum standard to provide an adequate signal for fingertip PPG analysis using smartphone camera.

The University of Birmingham already had a contract in place with EE, a mobile network provider for 4G SIM cards for the first 3 months of the sub-study, hence there was a limited number of 8G SIM cards available. My initial design incorporated 8G SIM cards to allow for ample data capture. Although the 4G cards were satisfactory for most patients there were six patients who had reached maximum data capacity within a given month and required an upgrade to 8G. In March 2019 all remaining participants had their mobile data usage increased to 8G. Data collection limitations have been outlined above.

4.4 Statistical analysis plan

A statistical analysis plan was completed prior to data analysis. The planned analysis was predominantly exploratory in nature and used trial endpoints collected during RATE-AF follow-up. No blinding was performed of wearable data or outcomes. Data pertaining to the randomised allocation (digoxin versus beta-blockers) was analysed in an intention-to-treat approach. The analysis was split into 2 sections; conventional statistical methods and advanced AI-based approaches.

4.4.1 Conventional statistical methods

Prior to conducting any analysis, the data underwent pre-processing to standardise clinical and sensor data. Continuous measurements of HR and step count were pooled at 1-minute

intervals to form time-series panelled data. Primary analysis using mean heart rate and total step count was carried out over a pre-specified 20-week period (to factor for missing data) The results were summarised and presented as a number, percentage, mean and standard deviation, median with interquartile range (IQR) or standard error of mean with variance.

Analysis was performed using random effects modelling with Xtreg and the generalised least square regression (GLS) method. Time series data was analysed for a 20-week period to reduce the number of missing data points. Mean heart rate and standard error of mean were assessed over a twenty-week period for heart rate and total step count. The Kruskal-Wallis non-parametric test was used to determine any statistically significant difference between the two treatment arms. To account for multiple repeated measurements over time in individual participants, generalised estimating equation models were generated using a random-effects estimator. Statistical analyses were performed using Stata version 17 (StataCorp LP, Texas), with a p-value <0.05 denoting statistical significance.

4.4.2 Artificial intelligence-based analysis using machine learning algorithms and validation

Neural networks were used to assess if continuous sensor data was comparable to conventional periodic trial measurements, taken at the closest RATE-AF trial appointment (designed and implemented by colleague, Dr Barksy). The approach utilized the AI framework that was developed and published to guide clinicians and researchers on the use of AI-approaches in clinical studies.(83) Unlabelled wearable sensor data from staggered 4-hour time periods were used to develop a self-supervising CNN, as seen in **figure 24**. Sampling overlapping subsets of data significantly multiplied the amount of sensor data available for model training. Heart-rate data was normalised to have mean 0 and variance 1,

while step-count data was condensed into the interval [0,1], preserving the zero-property of time points where patients were sedentary. Data was also augmented with a third channel denoting *missing-ness*, which is 0 for time points when data was missing, and 1 where data was present. This avoided missing data imputation and allowed the model to learn the significance of missing data instead of making assumptions.

The clinical model's predictor variables included ECG heart rate and 6MW results, taken at the start of the trial. The wearable sensor model's predictor variables were latent embeddings derived from the corresponding patients' wearable sensor data. Since the purpose of this model was to predict the patient's future NYHA class, the model was validated by embedding data from each patient's first week and using that embedding to predict their NYHA class at the end of trial. The self-supervised model was trained using all data for each patient other than in this first week, whilst also holding out a subset of patients as a validation set to monitor under/overfitting during model training. This hold-out set comprised 20% of the patient group, repeated across 5 iterations with k-fold cross-validation. Two logistic regression models were compared for prediction of NYHA class at the end of the trial (5 months later): (1) a conventional trial model including ECG heart rate and 6MW test results (distance travelled, time taken and participant speed); and (2) the wearable sensor model fitted with different predictor variables.

The ML model was evaluated using a performance statistic known as F1 score, which is an evaluation metric combining precision and recall. Like the area under the receiver operator curve, the F1 score ranges from 0 to 1, with 1 being the best score for classification. Machine learning analyses were performed using Python (Python Software Foundation, Delaware) with scikit-learn, and TensorFlow (Google Brain, California).

Figure 24: Neural network architecture

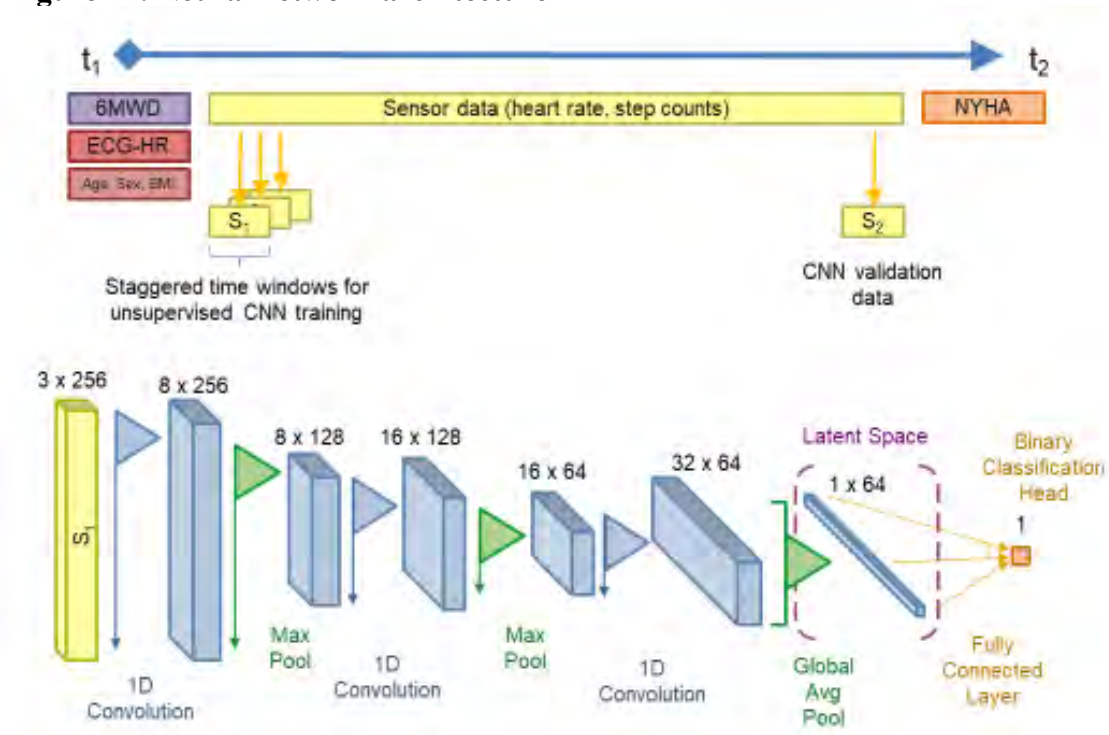


Figure provided by Dr Andrey Barsky (Health Data Research, Midlands, UK)
Top panel: At t1, 6-minute walk test, ECG heart rate, age, sex and body mass index are taken at the closest trial appointment to the start of the wearables sub-study.
Lower panel: The convolutional neural network (CNN) architecture

Chapter 5: The use of wearable technology to evaluate dynamic heart rate control in the RATE-AF randomised trial

As mentioned previously clinical trials use periodic, single timepoint assessments to evaluate the efficacy and safety of therapeutic interventions, as well as to provide information on physiological or functional parameters. In the case of cardiology studies, this often requires multiple clinical visits to obtain tests such as ECG for assessment of heart rate, and 6MW test for assessment of physical capacity. These in-hospital measurements are time consuming and costly from a resource perspective. They are also not dynamic, providing a limited snapshot of that person in an environment dissimilar to their home which may not be an accurate reflection of their functional status.(206, 207)

Wearable technology provides an opportunity for continuous measurement of physiological and functional parameters, with non-medical grade consumer devices being increasingly used by the public. Wearables offer a potentially non-invasive and affordable way to collect real-time data on common parameters such as heart rate and physical activity in a patients' day-to-day environment, often with minimal input. They can provide substantially more personalised information than conventional periodic review.(109, 198, 199, 208) Despite the large volume of data acquired from these devices, they are not without their shortcomings. When compared to standard medical devices drawbacks include lower quality data collected with frequent missing values. In addition, conventional statistical methods are not able to adequately handle the volume and complexity of data obtained from wearable technology and so often necessitates the use of machine learning approaches for data analysis.(83)

Whilst there is considerable potential for consumer wearable devices to contribute to cardiovascular disease management, there are few robust studies independent from manufacturers that can highlight opportunities and limitations.(145)

During the start of my PhD the RATE-AF randomised controlled trial was already underway and I took this opportunity to embed a wearable study within it which look at two biomarkers from the Fitbit wearable device; PPG heart rate and physical activity in the form of total step count.

5.1 Hypotheses tested

The following hypotheses were tested:

1. The effectiveness of beta blockers and low-dose digoxin on PPG heart rate (measured from the Fitbit wearable device), is comparable for long term AF rate control.
2. The effectiveness of beta blockers and low-dose digoxin on total step count (measured from the Fitbit wearable device), as a measure of physical activity is comparable.
3. Neural networks trained on wearable sensor data are comparable to conventional measurements of ECG heart rate and 6-minute walk test for prediction NHYA class at the end of the trial.

5.1.1 Data not analysed

Before detailing the results, it is important to mention that not all of the data generated was analysed in this thesis. Firstly, the feasibility of integrating wearable technology in an ongoing RCT with a mean age of 76 was not fully addressed. This was because PPG data on activity from the smartphone was not analysed, due to limited collection. Participants often forgot to carry their phones daily result in significant missing activity data. The smartphone

PPG application that was installed to assess for AF could not analysed and neither was the data collected on applications regarding Patient Reported Outcome Measures (PROMs). The comparison between Fitbit heart rate and ECG heart rate was not performed.

5.2.1 Overall baseline characteristics

One hundred and sixty patients participated in the RATE-AF trial of which 80 were randomised to each treatment arm and received either low dose digoxin or beta-blocker therapy. A total of 72 of these participants (36 from each treatment arm) were eligible to participate in the wearables sub-study. The remaining participants had completed the main trial or did not have sufficient time left for data collection (**Figure 25**). Of those eligible, a total of 53 enrolled onto the sub study. A total of 8 declined to participate in the digoxin group and 11 in the beta-blocker group, principally as they did not want to use a wearable device or attend further trial appointments. The characteristics of those that declined were similar to the participating cohort. Overall, both groups had similar co-morbidities, symptomology, ECG heart rate as well as a mean LVEF categorised as normal. The differences included a larger proportion (63%) of women who declined to participate; as did those with reduced physical capacity based on their 6MW distance (6MWD) with a median distance covered of 266 metres (IQR 44-330 metres), in those who declined versus 384 metres (IQR 207-437 metres) in those who participated (**table 22)**.

Figure 25: RATE-AF wearables sub-study enrolment flowchart

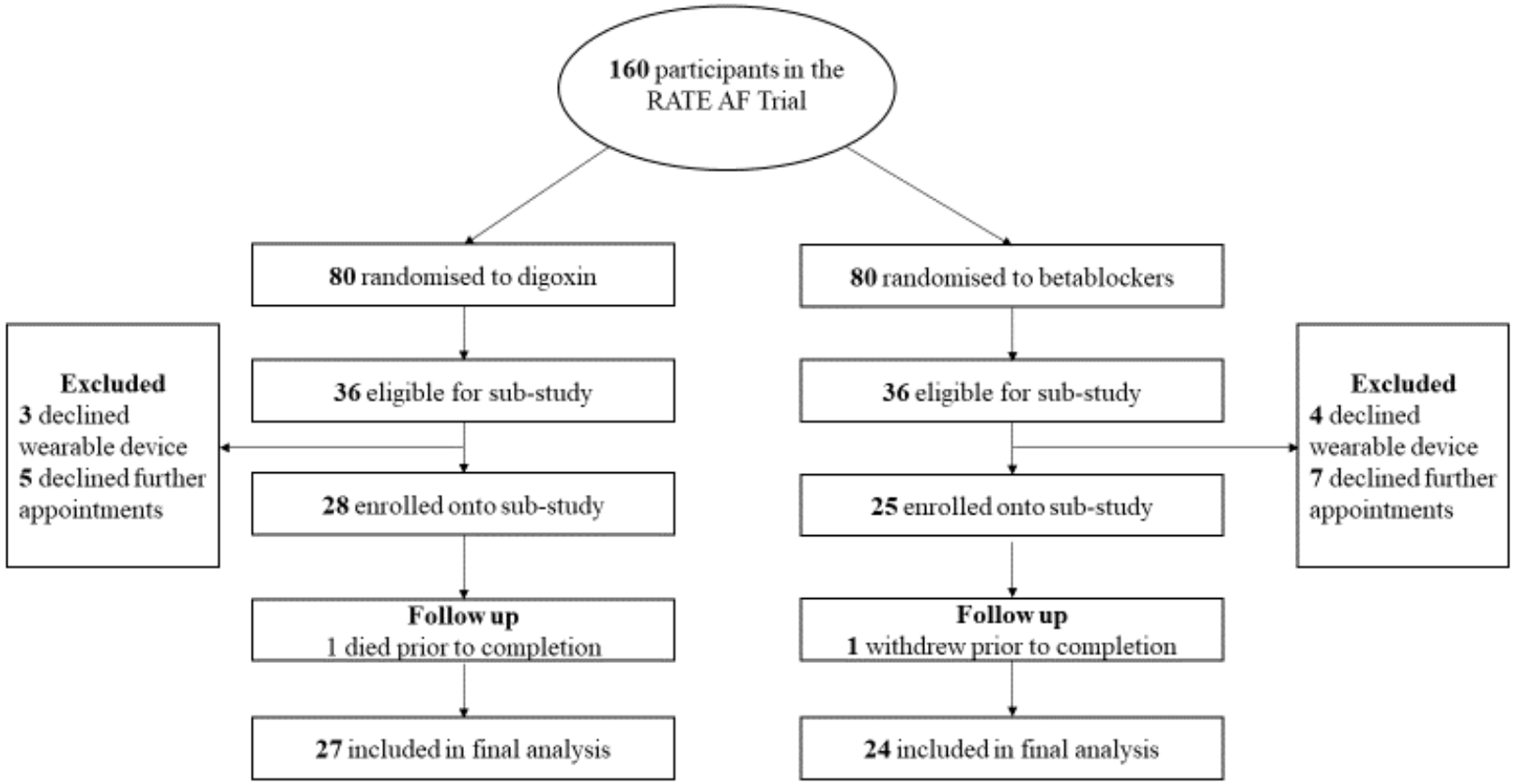


Table 22 : Baseline characteristics of patients included or declined participation

Characteristic		Included in the wearables sub-study (n=53)	Declined the wearables sub-study (n=19)
Age at randomisation, years	Mean (SD)	75.6 (8.4)	76.4 (7.1)
	Range	61-90	61-87
Gender, women n (%)		21 (39.6%)	12 (63.2%)
Hypertension, n (%)		39 (73.6%)	16 (84.2%)
Diabetes mellitus, n (%)		11 (20.8%)	6 (31.6%)
Previous stroke or transient ischemic attack, n (%)		8 (15.1%)	5 (26.3%)
Treatment with inhalers for COPD or asthma, n (%)		14 (26.4%)	10 (52.6%)
Diagnosed with heart failure, n (%)		24 (45.3%)	12 (63.2%)
LVEF on echocardiogram, mean % (SD)		55.6 (8.8)	63.2 (5.8)
LVEF on echocardiogram <50%, n (%)		18 (34.0%)	0 (0%)
NT-proBNP, median pg/mL (IQR)		1099 (770-1725)	913 (584-1691)
NYHA class, n (%)	I	0 (0%)	0 (0%)
	II	38 (71.7%)	9 (47.4%)
	III	14 (26.4%)	10 (52.6%)
	IV	1 (1.9%)	0 (0%)
mEHRA class, n (%)	1	0 (0%)	1 (5.3%)
	2a	3(5.7%)	7 (36.8%)
	2b	26 (49.1%)	8 (42.1%)
	3	22 (41.5%)	3 (15.8%)
	4	2 (3.8%)	0 (0%)
ECG heart rate, mean beats/min (SD)		97.7 (20.4)	100.8 (18.0)
Systolic blood pressure, mean mmHg (SD)		137.9 (17.2)	136.1 (19.0)
6-minute walk distance, median metres (IQR)		384 (207-437)	266 (44-330)

SD=standard deviation; COPD=chronic obstructive pulmonary disease; LVEF=left ventricular ejection fraction; IQR= interquartile range; NHYA= New York Heart Association; mEHRA=modified European Heart Rhythm Association; ECG=electrocardiogram

Fifty-three participants were enrolled in the sub-study, with a mean age at randomisation of 75.6 years (SD) 8.4; range 61 to 94 years) and 40% women. Of the 53 participants 28 (53%) had been randomised to low-dose digoxin and 25 (47%) to beta-blockers a mean of 30 weeks

prior to their entry into the wearables sub-study (SD 8 weeks; range 12 to 46 weeks). Both groups were well balanced with respect to demographics and clinical measurements. The commonest co-morbidities present were hypertension (74%) and heart failure (45%). There were more patients with a formal diagnosis of heart failure in the digoxin group, but mean LVEF measured in a blinded fashion at the start of the trial was similar for both groups: 55.7% [SD 8.8%] for digoxin versus 55.4% [SD 9.0%] in the those randomised to beta-blockers. Overall, the majority of patients reported mild to moderate symptoms of AF and HF (NHYA class II; 72% and EHRA class 2b; 49%) that impacted their daily physical activity. Only 3 participants had symptoms of such severity that daily physical activity was discontinued (NHYA and EHRA class 4). The median NT-pro-BNP, baseline ECG heart rate and capacity for physical activity was similar in both groups, with median 6MWD of 372 metres for digoxin (IQR 150-435) and 384 metres for beta-blockers (IQR 229-438). Heart rate at trial baseline was 98 beats per minute (SD 20.4), with 1 patient in each group presenting in sinus rhythm. At the trial mid-point, the mean dose of digoxin used was 157µg (SD 10µg) with serum digoxin level 0.79 ng/mL (SD 0.29), and in the beta-blocker group 3.2mg of bisoprolol (SD 1.9).(196) The duration of the wearable sub-study was a median of 23 weeks (IQR 4-38). 51 participants were included in the final analysis (see **table 23**).

Table 23: Baseline characteristics by randomised group

Characteristic		Randomised to digoxin N = 28	Randomised to beta-blocker N = 25	Overall N = 53
Demographics				
Age at randomisation, mean years (SD) range (years)		74.2 (8.4) (61-90)	77.2. (8.3) (64-88)	75.6 (8.4) (61-90)
Gender, women n (%)		11 (39.3%)	10 (40%)	21 (39.6%)
BMI (Kg/m^2) (SD)		32.3 (6.5)	29.2 (7.1)	30.8 (6.9)
Baseline comorbidities				
Hypertension, n (%)		22 (78.6%)	17 (68%)	39 (73.6%)
Diabetes mellitus, n (%)		5 (17.9%)	6 (24%)	11 (20.8%)
Previous stroke or transient ischemic attack, n (%)		5 (17.9%)	3 (12%)	8 (15.1%)
Treatment with inhalers for COPD or asthma, n (%)		8 (28.6%)	6 (24%)	14 (26.4%)
Diagnosed with heart failure, n (%)		16 (57.1%)	8 (32%)	24 (45.3%)
LVEF on echocardiogram, mean (SD), %		55.7 (8.8)	55.4 (9.0)	55.6 (8.8)
Echocardiogram LVEF < 50%, n (%)		11 (39.3%)	7 (28%)	18 (34.0%)
Clinical measurements at baseline trial visit				
NT-proBNP, median pg/mL, (IQR)		1112 (766-1831)	1057 (829-1717)	1099 (770-1725)
NHYA class, n (%)	I II III IV	0 22 (78.6) 5 (17.9) 1 (3.6)	0 16 (64) 9 (36) 0	0 38 (71.7) 14 (26.4) 1 (1.9)
mEHRA class, n (%)	1 2a 2b 3 4	0 1(3.6) 16 (57.1) 11 (39.3) 0	0 2 (8) 10 (40) 11 (44) 2 (8)	0 3 (5.7) 26 (49.1) 22 (41.5) 2 (3.8)
ECG heart rate, mean beats/min (SD)		95.7 (19.8)	99.6 (20.5)	97.7 (20.4)
Systolic blood pressure, mean mmHg (SD)		136.2 (15.2)	139.7 (19.3)	137.9 (17.2)
6- minute walk distance, median metres (IQR)		372 (150-434.5)	384 (229-438)	384 (207-437)

SD=standard deviation; BMI= body mass index; COPD=chronic obstructive pulmonary disease; LVEF=left ventricular ejection fraction; IQR= interquartile range; NHYA= New York Heart Association; mEHRA=modified European Heart Rhythm Association; ECG=electrocardiogram

5.2.2 Interim review

A face-to-face review was conducted at 1-month post enrolment to assess compliance in the preceding week with the wearable device, smartphone and to check application usage. It was noted that only 36% of participants had the ability to connect to home WIFI, hence the majority were reliant on the data plans provided. Seventy-five percent of participants were able to carry the smartphone regularly, the main reason for not carrying it was forgetfulness (85%). AF recordings made using the PPG application was technically more challenging. In total 65% of participants stated they had completed their AF recordings using the PPG application provided. On checking their technique only 73% were able to do it correctly. With regards to engaging with the My AF application, 65 % were able to successfully record their quality of life and AF symptom score. The biggest reason for not engaging with applications provided was due to technical issues with the application itself and forgetfulness (see **table 24**). Of note there was consistent use of the wrist-worn wearable (more so than the smartphone), with 90.4% of participants wearing it every day in the last 7 days before interim review (**table 25**).

Table 24: Compliance with technology at interim review

Characteristic		Randomised to digoxin N= 28	Randomised to beta-blocker N= 25	Total N= 52
Carrying smartphone regularly in the last 7 days, n (%)		22	17	39 (75.0%)
Reason for not carrying smartphone, n (%)	Forgot	4	7	11 (84.6%)
	Too big to carry smartphone	2	0	2 (16.7%)
AF recordings made using PPG smartphone camera in last 7 days, n (%)		15	18	33 (63.5%)
AF recordings made using PPG smartphone camera correctly in last 7 days, n (%)		13	11	24 (72.7%)
Reason not recorded, n (%)	Forgot	2	3	5 (27.8%)
	Health limitation	1	1	2 (10.5%)
	App not working	6	3	9 (50%)
	Other	3	0	3 (16.7)
In the last week recorded Quality of life on 'my AF app', n (%)		16	17	33 (64.7)
If not reason why quality of life not recorded, n (%)	Forgot	6	6	12 (63.1%)
	Health limitation	2	1	3 (15.8%)
	App not working	0	0	0
	Other	3	1	4 (21.1%)
In the last week recorded mEHRA scores on 'my AF app', n (%)		16	17	33 (64.7%)
If not reason why mEHRA not recorded, n (%)	Forgot	9	7	16 (84.2%)
	Health limitation	2	1	3 (15.8%)
	App not working	0	0	0

AF=atrial fibrillation; PPG=photoplethysmography; mEHRA=modified European Heart Rhythm Association, *1 patient did not attend the interim review.

Table 25: Usage of wearable device at interim review

Characteristic		Randomised to Digoxin N = 28	Randomised to Beta-blocker N = 24*	Total N = 52
Number of days wearable device worn in last 7 days, n	0 days	1 (3.6%)	1 (4.2%)	2 (3.9%)
	1 day	0 (0%)	0 (0%)	0 (0%)
	2 days	0 (0%)	1 (4.2%)	1 (1.9%)
	3 days	1 (3.6%)	0 (0%)	1 (1.9%)
	4 days	0 (0 %)	0 (0%)	0 (%)
	5 days	0 (0%)	1 (4.2%)	1 (1.9%)
	6 days	0 (0%)	0 (0%)	0 (%)
	7 days	26 (92.9%)	21 (87.5%)	47 (90.4%)

*1 patient did not attend the interim review.

5.2.3 Final review

A total of 51 participants completed the final sub-study follow up, 1 participant was lost to follow up from each arm of the RATE-AF trial due to a withdrawal (beta-blocker) and unrelated trial death (digoxin). In terms of clinical status, as the RATE-AF trial reported, there was an improvement in the classification score for HF and AF symptoms. The majority of patients reported no limitation (class 1; 45.1%) or slight limitation (class 2; 51%) of physical activity, compared to class 2 or 3 (marked limitation) at baseline. In terms of AF symptoms, again the EHRA classification score was better with the majority reporting no symptoms (class 1; 49%) compared to baseline. No participants reported symptoms that were severe enough to be present at rest. In terms of cardiac function, the median NT-proBNP level was lower in the digoxin group at 947 pg/mL at final review (IQR 592-1759 pg/mL), compared to 1112 pg/mL at baseline (IQR 766-1831 pg/mL). In comparison, in the bisoprolol group an increase was seen with the NT proBNP level measuring 1288 pg/mL at

final review (IQR 894-2143 pg/mL) compared to 1057 pg/mL at baseline (IQR 829-1717pg/mL). The mean LVEF on echocardiogram improved in both groups and overall, with fewer patients having an LVEF of less than 50% (see **table 26**).

Table 26: Clinical characteristics at the end of the sub-study

Characteristic		**Randomised to Digoxin N=27**	**Randomised to Beta blocker N= 24**	**Total N= 51**
NHYA class, n (%)	I	17	6	23 (45.1%)
	II	8	18	26 (51.0%)
	III	2	0	2 (3.9%)
	IV	0	0	0
mEHRA class, n (%)	1	16	9	25 (49.0)
	2a	7	10	17 (33.3)
	2b	2	4	6 (11.8)
	3	2	1	3 (5.9)
	4	0	0	0
LVEF on echocardiogram, mean %, (SD)		60.1 (9.6)	58.9 (8.2)	59.3 (8.9)
Echocardiogram LVEF ≤ 50%, n (%)		4	2	6 (11.5%)
NTproBNP, median (pg/mL), (IQR)		947 (592-1759)	1288 (894-2143)	1158 (793-1912)

NHYA= New York Heart Association; mEHRA=modified European Heart Rhythm Association; LVEF=left ventricular ejection fraction; SD=standard deviation; IQR= interquartile range

Table 27: Usage of smartphone and applications at end of sub-study

Characteristic		Randomised to Digoxin N= 27	Randomised to Beta blocker N= 24	Total N= 51
Carrying smartphone regularly in the last 7 days, n (%)		22	17	39 (76.5%)
Smartphone was used for calls, messaging, internet and applications, n (%)		11	10	21 (41.2%)
AF recordings made using PPG smartphone camera in last 7 days, n (%)		21	19	40 (78.4%)
AF recordings made using PPG smartphone camera correctly in last 7 days, n (%)		15	13	28 (68.3%)
Reason not recorded, n (%)	Forgot	3	2	5 (45.4%)
	Due health limitation	4	1	5 (45.4%)
	App not working	0	1	1 (9.1%)
	Other	0	2	2 (18.2%)
Ease of use of My AF app, n (%)	Easy	14	13	27 (52.9%)
	Intermediate	4	3	7 (13.7%)
	Hard	9	8	17 (33.3%)
In the last week recorded Quality of life on 'my AF app', n (%)		20	13	33 (64.7%)
If not reason why quality of life not recorded, n (%)	Forgot	4	7	11 (61.1%)
	Not able to due to health limitation	0	3	3 (16.7%)
	App not working	1	0	1 (5.6%)
	Other	2	1	3 (16.7%)
In the last week recorded mEHRA scores on 'my AF app', n (%)		17	12	29 (56.9%)
If not reason why mEHRA not recorded, n (%)	Forgot	8	9	17 (77.3%)
	Not able to due to health limitation	0	3	3 (13.6%)
	App not working	0	0	0
	Other	2	0	2 (9.1%)
Use of 'my AF app' improved understanding of AF, n (%)	Yes	14	10	24 (47.1%)
	No	5	4	9 (17.7%)
	Not sure	2	3	5 (9.8%)
	Not applicable	6	7	13 (25.5%)
Use of 'my AF app' improved compliance to AF medication, n (%)	Yes	8	10	18 (35.3%)
	No	13	6	19 (37.3%)
	Not sure	1	1	2 (3.9%)

	Not applicable	5	7	12 (23.5%)

AF= atrial fibrillation; PPG= photoplethysmography; mEHRA=modified European Heart Rhythm Association

The number of participants who were able to carry the smartphone remained the same as at interim, but only 41% used the smartphone for messaging, making calls and accessing the internet. With regards to PPG recordings participant engagement was higher at 78%, and the majority (68%) were able to complete these recordings correctly. Use of the 'My AF' application for recording quality of life and mEHRA scores remained similar to interim. In terms of ease of use 52.9% found this application easy to use, 13.7% classed it as intermediate and just over one third found it difficult to use. In general, 47.1% of participants felt the application improved their understanding of AF, but this did not translate to better medication use with only 35.3% finding it improved their compliance. Again, compliance with the wearable device remained high with 92.2% wearing it every day. A total of 57.9% felt it improved their motivation to exercise and 54.9% reported higher overall physical activity whilst using the wearable device (see **tables 27 and 28**).

Table 28: Usage of wearable device measured at end of sub-study

Characteristic		Randomised to Digoxin N= 27	Randomised to Beta blocker N= 24	Total N= 51
Number of days Fitbit device worn in last 7 days n (%)	0 days	2	1	3 (5.9%)
	1 day	0	1	1 (2.0%)
	2 days	0	0	0
	3 days	0	0	0
	4 days	0	0	0
	5 days	0	0	0
	6 days	0	0	0
	7 days	25	22	47 (92.2%)
Reason for not wearing Fitbit device, n (%)	Non-compliance	0	1	1 (25.0%)
	Did not wish to continue	0	0	0
	Hospital admission	0	0	0
	Damage to Fitbit or smartphone	1	0	1 (25.0%)
	Other	1	1	2 (50.0%)
Fitbit improved motivation for physical activity, n (%)		16	13	29 (56.9%)
Fitbit increased physical activity overall, n (%)		16	12	28 (54.9%)

5.2.4 Wearable-acquired heart and physical activity

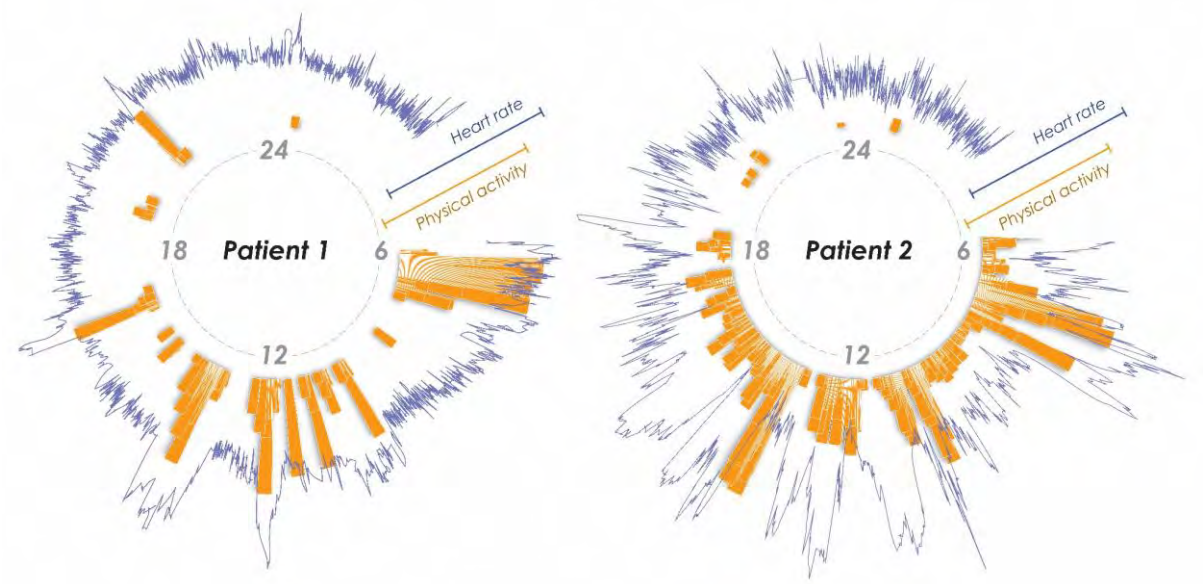
The mean duration of ambulatory sensor data collected was 20 weeks (SD 7). Across all patients there were 143,379,796 data intervals collected for heart rate and 23,704,307 for corresponding physical activity. In those randomised to digoxin, the mean data collection period was 153 days (SD 53), and the total number of combined data points for heart rate and step count was 4,746,169. In those randomised to beta-blockers the mean period was 160 days (SD 46) and 4,683,959 combined data points were obtained. Per patient, an average of 2,623,951 heart rate (SD 907,697) and 436,081 step count data points (SD 122,493) were obtained for each digoxin-treated patient; and for those allocated to beta blockers, 2,796,367 (SD 811,956) heart rate and 459,762 step count data points (SD 110,443) were collected. (**Table 29**). There was considerable variation in the correlation between heart rate and physical activity and **figure 26** demonstrates the variation that occurred in a 24-hour period for two patients. **Figure 27** shows the median value and range of correlation between heart rate and physical activity during the day (08:00 to 20:00) over a 12-week period, averaged at 10 second intervals for patients in AF (n=50) who were appropriately rate controlled with either beta-blockers or digoxin.

Table 29: Sensor data collection from wearables

Data collection	Randomised to digoxin N = 28	Randomised to beta-blocker N = 25
Total number of data points for heart rate	73,470,613	69,909,183
Total number of data points for step count	12,210,254	11,494,053
Number of combined data points for heart rate and step count	4,746,169	4,683,959
Mean number of data points for heart rate (SD)	2,623,951 (907,697)	2,796,367 (811,956)
Mean number of data points for step count (SD)	436,081 (122,493)	459,762 (110,443)
Mean number of combined data points for heart rate and step count (SD)	169,506 (71,217)	187,358 (57,784)
Mean timespan for data collection, days (SD)	153 (53)	160 (46)

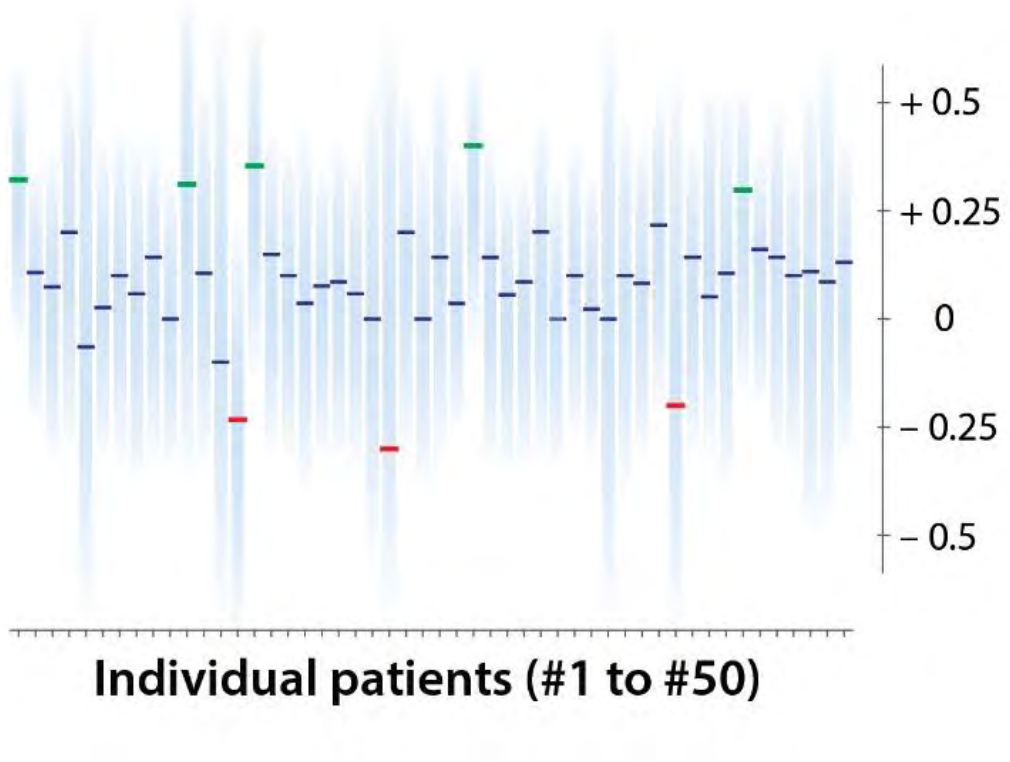
SD=standard deviation

Figure 26: Wearable device data for heart rate and physical activity over a 24-hour period



Individual examples of circadian data capture, heart rate (blue) and physical activity in terms of step count (orange). Patient 1 was allocated beta blocker therapy and patient 2 to low dose digoxin.

Figure 27: Wearable device data for heart rate and physical activity



5.2.5 Comparison of digoxin versus beta-blockers for heart rate control and physical activity

Weekly averages of heart rate were no different comparing those randomised to low-dose digoxin or beta-blockers. For total step count the weekly averages were higher in patients randomised to digoxin, especially in later periods. This was evident during weeks 11-15 where no statistically significant difference in heart rate control was seen with either treatment (p-value=0.96). However, there was a statistical difference in total step count. Those randomised to digoxin therapy had a total step count of 30,045 (SEM 2102), compared to 25,856 (SEM 2251) on beta blockers with a p-value of 0.025 (**table 30**).

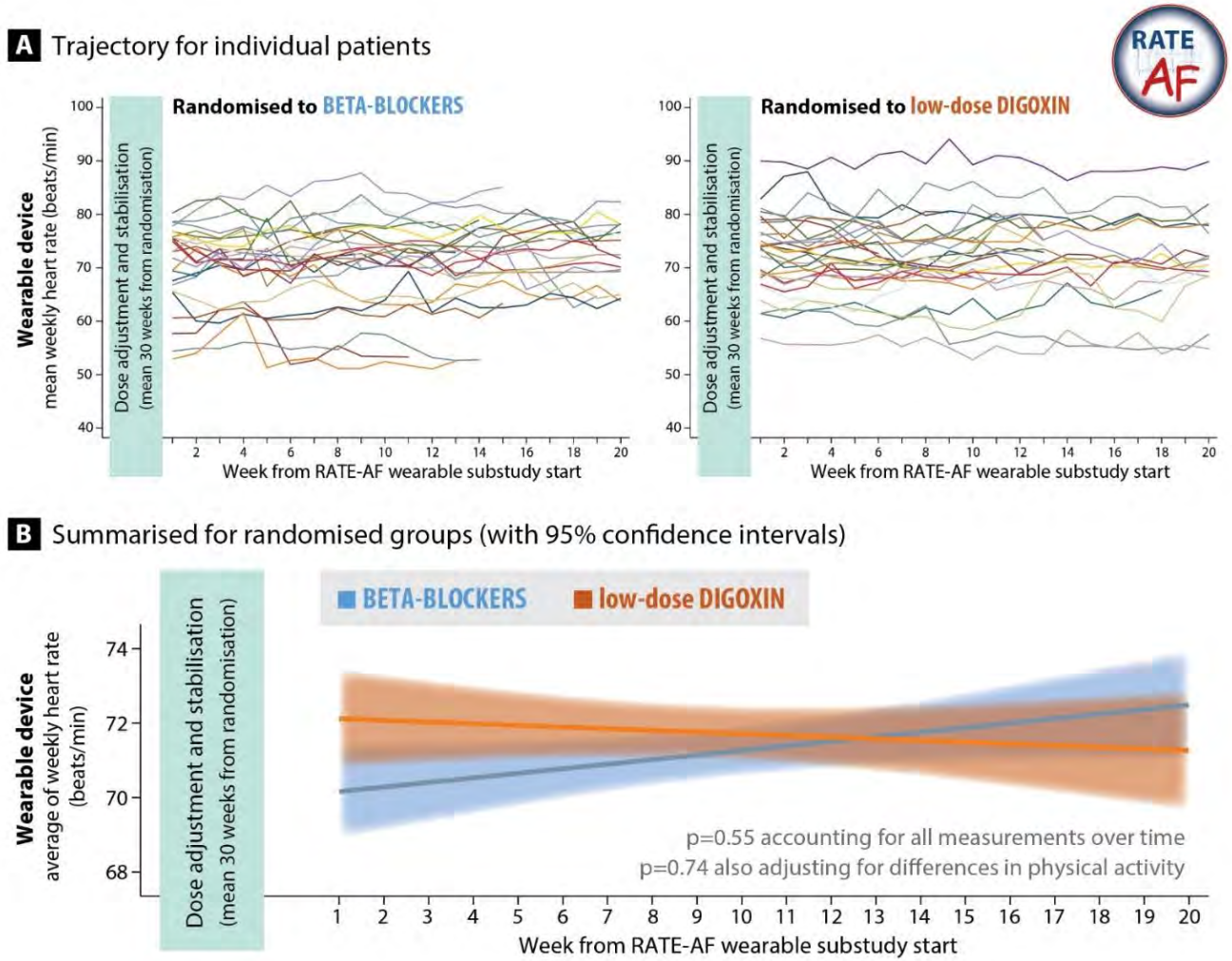
Figure 28A shows the individual trajectory of each patient randomised to digoxin (left) and bisoprolol (right). Accounting for all repeated measurements over time, there was no significant difference in heart rate comparing low-dose digoxin and beta-blocker therapy using the wearable sensors. The unadjusted regression coefficient for digoxin versus beta-blockers was 1.22 (95% CI -2.82 to 5.27; p=0.55), and 0.66 adjusted for age, gender, diagnosis of heart failure and NT-proBNP (95% CI -3.45 to 4.77; p=0.75). There remained no difference in heart rate between digoxin and beta-blockers after accounting for physical activity (p=0.74; **figure 28B**). A post-hoc adjusted subgroup analysis according to activity levels found no difference in heart rate between digoxin and beta-blockers in those with low weekly-averaged activity (<15,000 steps/week; 298 weeks from 44 patients; p=0.48), minimum recommended activity (15-30,000 steps/week; 316 weeks from 37 patients; p=0.47), or recommended activity (≥30,000 steps/week; 417 weeks from 33 patients; p=0.97).

Table 30: Wearable sensor parameters of heart rate and physical activity over sequential five-week periods

Wearable sensor output	Weeks 1-5			Weeks 6-10			Weeks 11-15			Weeks 16-20		
	Digoxin	Beta-blockers	p-value	Digoxin	Beta-blockers	p-value	Digoxin	Beta-blockers	p-value	Digoxin	Beta-blockers	p-value
Heart rate, mean beats per minute (SEM)	72.1 (0.6)	70.8 (0.7)	0.54	71.6 (0.7)	70.5 (0.8)	0.97	71.9 (0.9)	71.5 (0.7)	0.96	71.2 (1.0)	72.7 (0.6)	0.25
Total step count (SEM)	32,519 (2480)	28,287 (2052)	0.53	31,461 (2065)	30,114 (2163)	0.53	30,045 (2102)	25,856 (2251)	0.025	33,539 (2499)	29,105 (2694)	0.07

SEM= Standard error of mean

Figure 28: Wearable device heart rate according to randomised therapy

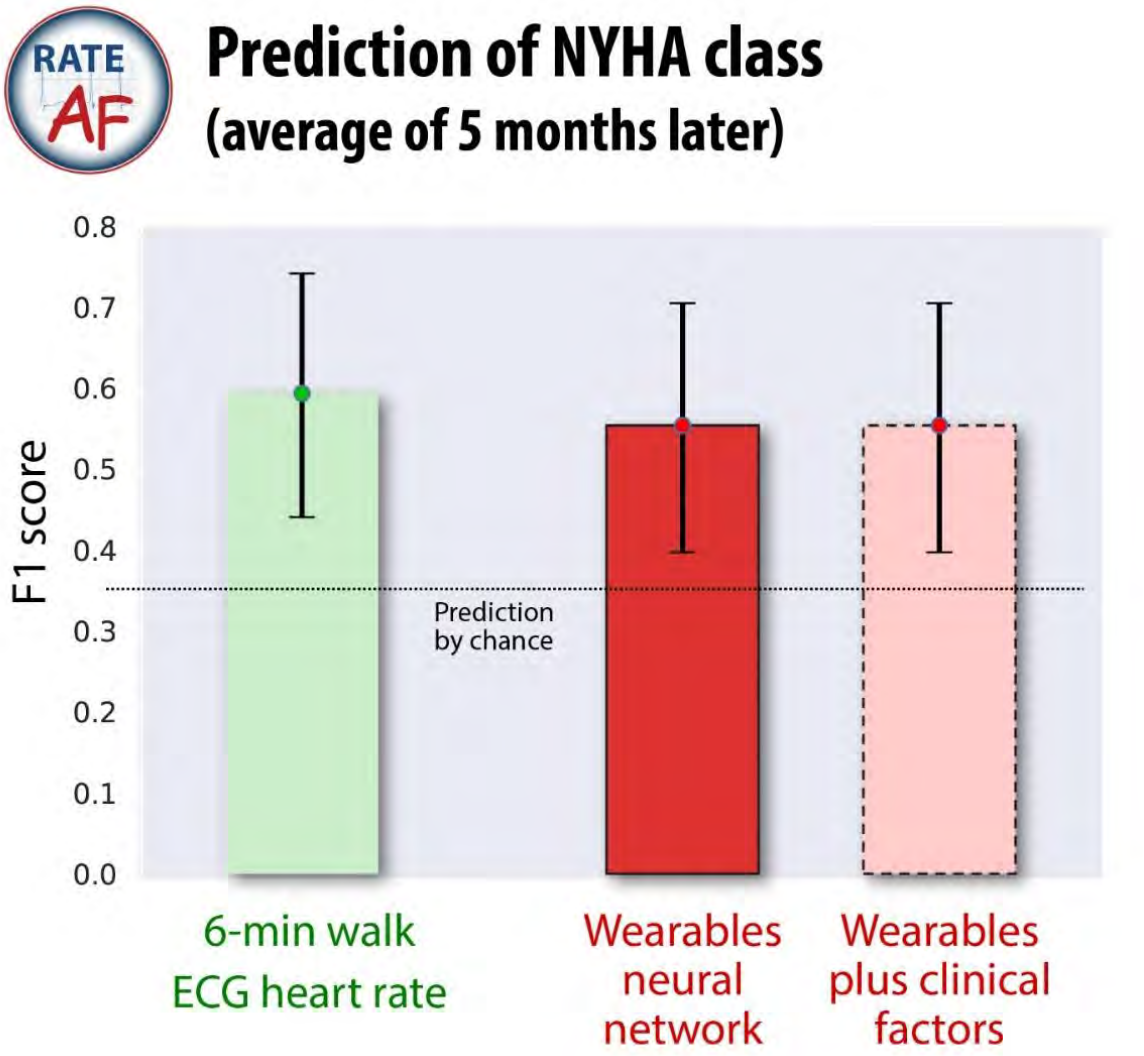


28A: Mean weekly heart rate condensed from all wearable device data for each individual patient over a 20-week period randomised to beta-blockers (left) and low-dose digoxin (right). 28B: Average weekly heart rates according to randomisation to beta-blockers or low-dose digoxin.

5.2.6 Prediction of NYHA class using wearables versus conventional trial parameters

The unsupervised CNN model was trained on heart rate and step count data using the wearable sensor output from 41 patients who were established on rate control therapy, and had sufficient time windows for analysis. For prediction of NYHA class at the end of the trial, the wearables neural network model yielded an F1 score of 0.55 (95% CI 0.40-0.70), with chance being 0.35. F1 scores combine precision and recall of each model and are presented with 95% confidence levels, see **Figure 29**. The wearable data were independent of clinical factors such as age, sex and body mass index, with no change in F1 score when combining wearable data with clinical factors (F1 score of 0.55; 95% CI 0.41-0.70). The wearables neural network was equivalent to a model generated from conventional trial parameters (ECG heart rate and 6MW test results) for NYHA class prediction, which returned an F1 score of 0.59 (95% CI 0.44-0.74); p=0.72 for comparison with wearables CNN.

Figure 29: Prediction of clinical progress with wearable data neural network



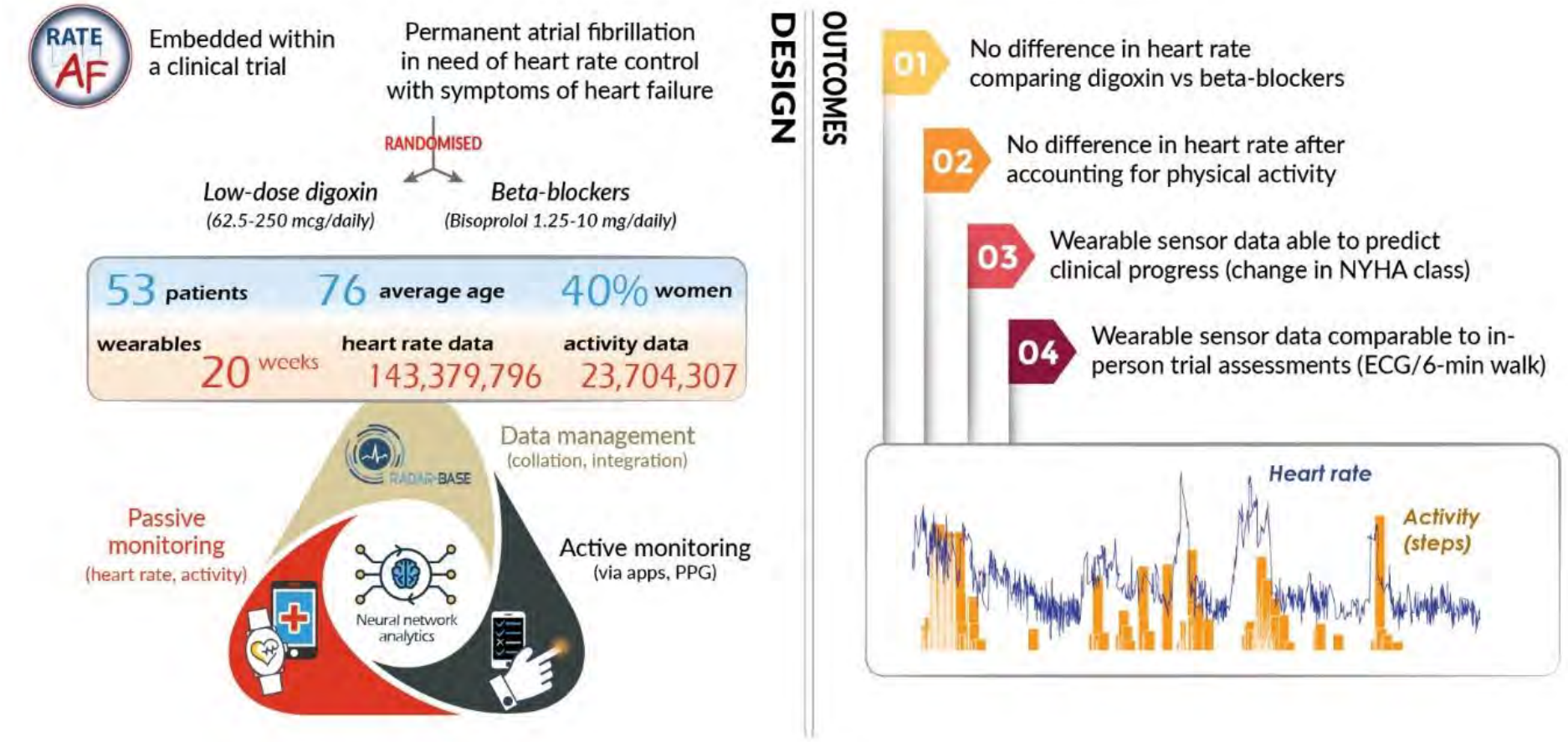
5.3 Discussion

The RATE-AF wearables study has demonstrated the potential for non-medical grade devices to contribute to research as well as clinical monitoring of progress in response to changes in therapy. Embedded with a randomised trial that was itself part of improvements to routine clinical care, the study provides robust information on the value, but also limitations of using these devices. The commonly held preconception amongst clinicians is that digoxin is inferior to beta-blockers for heart rate control in AF, which originated from early acute trials confounded by prescription bias.(64) Although these findings are based on a small sub-study

with patients on stable rate control therapy, they contradict this notion of beta-blocker superiority over digoxin, supporting the hypothesis that both treatments are comparable for rate control. No difference was seen over the study period consisting of over 140 million data points. The equivalence of low-dose digoxin to beta-blockers held even after accounting for individual differences in day-to-day physical activity, and no difference in heart rate was evident even in highly active time periods. For total step count weekly averages were higher in digoxin patients, especially in later weeks. During weeks 11-15 no statistically significant difference in heart rate control was seen with either treatment (p-value=0.96). However, there was a statistical difference in total step count in the digoxin patients compared to those on beta-blockers with a p-value of 0.025. Further work is required to evaluate the effect of beta-blockers and digoxin on physical activity.

The wearable sensor data was comparable to conventional clinical trial outcomes for prediction of clinical outcome (in this case, heart failure symptoms measured using NHYA classification), but required the development of a neural network pipeline for appropriate analytics. Again, this supported the hypothesis that trained neural networks and conventional measurements can predict future NYHA class to a similar extent. Standardisation of these approaches could potentially lead to wearable devices contributing to, or even replacing resource-intensive clinical tests and visits in the future. **Figure 30** provides a summary of the sub-study findings in the form of a graphical abstract.

Figure 30: Graphical abstract summary of the wearables sub-study



Overview of the RATE-AF randomised trial wearable study. ECG=Electrocardiogram; NYHA=New York Heart Association heart failure classification; PPG=photoplethysmography.

AF is a key public health issue, predicted to double in prevalence over the next few decades.(180) Recent evidence suggests that cognitive decline and vascular dementia should be added to the list of thromboembolic adverse events suffered by patients (and potentially avoidable).(209) New technologies have rapidly increased our ability to detect AF, but how these should be implemented in clinical practice remains unclear.(145) The approach to rate control of AF amongst clinicians is highly inconsistent, with under- and over-treatment common in routine care. Despite permanent AF being the most common 'type'(185), there is a remarkable lack of research to support decision-making and improve patient quality of life.(178, 191, 210-212) Treatment choices are often made based on evidence from heart failure trials, however the reduction in mortality from beta-blockers was not evident in double-blind trials in patients with concomitant AF.(62) Digoxin has traditionally been reserved for sicker patients or as a second-line agent. Like beta-blockers, there is no apparent mortality impact from digoxin in patients with AF and heart failure, and observational analyses are wholly inappropriate to evaluate outcomes with digoxin due to systematic prescription biases.(64) The RATE-AF trial was the first head-to-head, longer-term randomised trial of digoxin and beta-blockers, showing that low-digoxin had a similar impact on patient-report quality life as beta-blockers yet substantially improved functional class (two-class mEHRA improvement in 68% on digoxin versus 29% for beta-blockers at 12-months, and one-class NYHA improvement in 79% on digoxin versus 36% for beta-blockers). Use of digoxin led to a significant reduction in natriuretic peptides, and less than half the number of primary care and unplanned hospital visits.(196) This study now adds a further dimension, showing that low-dose digoxin can be an effective rate control agent as monotherapy. This was the case across all patient activities, including physical exertion which was poorly correlated with heart rate in these patients with AF. Although limited by sample size, weekly averages of heart rate did fall more over time in those randomised to

digoxin, accompanied by higher total step counts towards the end of the study. This may be consistent with the longer-term activity of digoxin, including pro-parasympathetic effects on cellular, electrophysiological and neurohormonal pathways.(213)

At present, the majority of clinical trials in cardiovascular research utilise in-person periodic testing that not only requires patient and staff attendance but is time-constrained in an environment unlike the patients' own surroundings. Wearables offer an exciting possibility for patient-directed data collection that better reflects real-life day-to-day variations in heart rate and physical activity.(110) In this study, a self-trained neural network was designed to provide value from the vast amounts of sensor information collected by wearables and address key issues such as inconsistent and missing data. The performance of the wearable neural network was independent of clinical factors, and equivalent for prediction of future NYHA class as conventional parameters such as a 12-lead ECG and 6MW testing. Of note, older patients were recruited (mean 76, with range of 61 to 90 years) with no or limited prior experience with health technology. Participants indicted a broad appeal to monitor and improve their own functional capacity. A study by Fabritz et al looked at the use of a wrist worn PPG sensor for atrial arrhythmia detection in 882 adults with an age range of 65 to 90 years. It found atrial arrythmias in 5% with a 95% CI of 3.6 to 6.6.(214). These findings support the use of wearable technology for atrial arrythmia detection in the older population.

Several studies have utilised wearable devices in cardiovascular research. Whilst the Apple smartwatch and Fitbit studies have received much attention using photoplethysmography to identify new AF (116, 184), there are also many different devices being used across heart failure to quantify patient physiology.(215) Early optimism has already started to fade somewhat, and one retrospective matched study in the US found that patients using wearable devices accessed significantly more healthcare but with no difference in heart rate.(216) It

remains to be seen how these consumer bought, non-medical devices will be integrated into routine care, with or without the acceptance of healthcare professionals.(206, 217) Further randomised trials are clearly needed to understand whether continuous monitoring can provide a personalised assessment of treatment response or help to identify sub-phenotypes of disease.

Similarly, the approaches to transparent and validated AI remain in their infancy.(125) AI has demonstrated an ability to go beyond our current linear understanding of disease trajectory and interactions, able to personalise diagnosis and therapeutic strategies even in multimorbid conditions where conventional statistics are unable to handle multi-dimensionality.(218, 219) Despite the potential benefits of AI, broader use in clinical practice requires an approach guided by standardised methodological principles as described in chapter 1.

5.3.1 Strengths and limitations

Although the number of patients included was limited, this study encompassed a large volume of wearable data and was able to benefit from being embedded into a randomised trial to limit extraneous bias. The wearables were implemented post-randomisation, however there was no crossover use of digoxin(196), analysis was by intention-to-treat, and balance in clinical characteristics between groups was maintained for those patients joining the wearable study. As expected, the actual usage amongst participants had considerable variation in terms of data collection time. Use in an older population (mean age 76 years, SD 8) brought about challenges, with many not previously owning a smartphone. However, good compliance and data quality was achieved by providing assistance as required to resolve technical issues. Despite this, missing data points were frequent, but their impact was minimised by using an

innovative approach to embed machine learning on the significance of missing data, rather than ignore or impute it. Diversity in ethnicity was severely limited (white British or Irish ethnicity accounted for 94% of participants) and so these results cannot be extrapolated to other ethnicity groups.

5.4 Conclusion

In summary, a wrist-worn wearable device and smartphone were successfully deployed within a randomised controlled trial to evaluate continuous, ambulatory heart rate and physical activity in patients with permanent AF and symptoms of heart failure. Low-dose digoxin and beta-blocker therapy had equivalent effects on longer-term heart rate, measured over a 20-week period with an average of two to three million data points on heart rate per patient. A neural network using the wearable sensor data had similar capability as traditional trial measures for predicting future health status.

Chapter 6: Other approaches to phenotype and determining prognosis in atrial fibrillation and heart failure

6.1 Introduction

6.1.1 Multimorbidity

Multi-morbidity is defined as the presence of two or more chronic conditions in an individual. It is associated with adverse outcomes, poorer quality of life and increased mortality, with a meta-analysis identifying those with at least two-co-morbidities having almost double the risk of death.(220) In developed countries the rise in individuals living with multiple chronic conditions is attributed to better healthcare, improved survival, and an ageing population. In CVD the prevalence of patients with five or more co-morbidities has risen four-fold to 25% from the year 2000 to 2014 (84). It is rapidly becoming a public health concern due to greater healthcare consumption and costs.(221-223)

Multi-morbidity is a complex entity where progression and severity are affected by genetic, environmental and lifestyle factors. It can arise in two ways, either through frequency or non-random association. Frequency occurs when common conditions occur randomly, this was seen in a recent study by Kuan et al which found HF co-occurred with eight other conditions and was more often seen in the older population. Non-random association is when conditions co-occur either due to shared genetic and environmental factors, or because a condition or its treatment causes another condition to develop.(221)

Despite the rising presence of multi-morbidity clinical research has remained focused on mono-disease and how this affects morbidity, and mortality with evidence-based treatment options orientated towards the management of that single disease. Research that factors in the presence of co-morbidities is limited as these individuals are usually excluded from

clinical trials. Current management of co-morbidities requires that each disease is treated separately. Often the treatment recommendations for a single disease can affect treatment efficacy and prognosis when administered in multi-morbidity due to unpredictable drug-drug and drug-disease interactions.(221, 223, 224)

In addition to the lack of multi-morbid patients recruited to clinical studies, conventional statistical methods are not able to factor for the presence of multiple co-morbidities. This is due to the limitations posed by data assumptions and their inability to process complex datasets with a large number of interacting variables. If the research hypothesis relates to a single disease, then often conventional statistical analysis will suffice. However, if the research question and analysis involve multi-morbidity and big datasets with higher dimensional interactions between variables then other approaches such as AI may be required.(83, 223)

In this chapter I will firstly show how conventional analysis can be used for phenotyping and prognostication in AF and HF, an exemplar of multimorbidity (Chapter 6.2). I will then demonstrate how AI tools can enhance accounting of multimorbidity interactions (Chapter 6.3).

6.2 The use of conventional statistical approaches

For this part of my PhD, I first looked at how conventional statistical analysis can be used to assess the impact of an additional co-morbidity on mortality (in this case renal impairment in patients with heart failure) and treatment efficacy (using beta-blockers as an example).

6.2.1 Rationale to assess heart failure and renal impairment

HF is often associated with renal impairment due to shared risk factors such as diabetes and hypertension, and through common pathophysiological pathways, including sympathetic nervous system activation and the RAAS. Renal impairment is an independent predictor for mortality and can accelerate disease progression in CVD and HF. Its presence in HF is significant due to its effect on the use of pharmacological treatments and mortality.(225-227) Just over half (51%) of the patients (n=47,716) on a Swedish HF registry (228) were found to have some form of renal dysfunction defined as an estimated glomerular filtration rate (eGFR) of <60 mL/min.(93, 229)

Renal impairment in HF patients limits the initiation, up-titration and maintenance of well-recognised guideline-based therapies.(226) There is a clear gap in our knowledge in terms of providing clinicians with evidence-backed treatment in those with HF and renal impairment. Despite this, these patients are often excluded from clinical trials, so the value of beta-blockers in renal impairment remains unknown. Previous sub-analyses from RCTs have suggested that the effectiveness of beta-blockers was preserved in patients at various levels of eGFR, but the number of participants in these studies was too small to quantify any significance.(230-233)

This meta-analysis evaluated the effect of renal dysfunction on mortality in patients with HFrEF using individual patient data (IPD) from the landmark, double-blind RCTs comparing beta-blockers with placebo.(234) A hypothesis that beta-blockers reduced mortality in patients with moderate and moderately severe renal impairment compared to placebo was tested.(93)

6.2.2 Methods for conventional statistical approach

The international Beta-blockers in Heart Failure Collaborative Group was formed to pool IPD from major HF trials to improve clinical management.(62, 63, 93, 235, 236) The background and study design has previously been described, and this review was conducted based on PRSIMA IPD guidance.(93, 237) To summarise trials eligible for inclusion were placebo-controlled, over 300 patients recruited; no other treatment investigated; follow up of more than 6 month and mortality reported as end point.(238) Eleven trials were included and are detailed below:

- Australia/New Zealand Heart Failure Study (ANZ)(239)
- Beta-Blocker Evaluation Survival Trial (BEST)(240)
- Carvedilol Post-Infarct Survival Control in LV Dysfunction Study (CAPRICORN)(241)
- Carvedilol Hibernating Reversible Ischaemia Trial: Marker of Success Study (CHRISTMAS)(242)
- Cardiac Insufficiency Bisoprolol Study (CIBIS I)(243)
- Cardiac Insufficiency Bisoprolol Study II (CIBIS-II)(244)
- Carvedilol Prospective Randomized Cumulative Survival Study (COPERNICUS)(245)
- Metoprolol in Idiopathic Dilated Cardiomyopathy Study (MDC)(246)
- Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF)(36)
- Study of the Effects of Nebivolol Intervention on Outcomes and Rehospitalisation in Seniors with Heart Failure (SENIORS)(247)

  U.S. Carvedilol Heart Failure Program (US-HF).(37)

Data collection

IPD from each trial was obtained using a standardised data request form, data was then cross-checked and compared with trial databases and published reports to ensure robustness. Trial databases were then harmonised according to the standardised form so that patient characteristics and outcomes across all trials matched. Imputation was not performed for missing data.(93)

For this analysis, the CIBIS I trial was excluded due to lack of data on renal function. Patients with a baseline creatinine and LVEF <50% from the remaining 10 trials were included and stratified by heart rhythm (sinus rhythm, AF/flutter). Those with a missing ECG or with a paced rhythm were excluded. Renal function was calculated using the Modification of Diet in Renal Disease equation to obtain eGFR, which was categorised as described in **table 31** below. Those patients with an eGFR of 30-59ml/min/1.73 m^2 were pre-specified as a group of interest, as were those with a 20% or more reduction in eGFR between baseline and follow-up.(93)

(116)

6.2.3 Statistical analysis

The main outcome for this analysis was all-cause mortality. Outcomes of the intention to treat analysis of beta-blockers versus placebo were presented using Cox proportional hazard regression, stratified by study and grouped by heart rhythm and eGFR category. Hazard ratios (HRs) and 95% confidence intervals (CIs) were provided, along with p values. Effect modification was evaluated by the use of values from interaction terms fitted in the multivariable models. Kaplan-Meier plots were selected to show pooled, unadjusted data for eGFR/treatment groups, with log-rank p values for comparison.(93)

Table 31: National Kidney foundation staging

Category	Estimated glomerular filtration rate
Stage 1(normal)	≥90ml/min/1.73m^2
Stage 2 (mildly decreased)	60-89ml/min/1.73m^2
Stage 3a (mild to moderately decreased)	45-59ml/min/1.73m^2
Stage 3b (moderate to severely decreased)	30-44ml/min/1.73m^2
Stage 4 & 5 (severely decreased or kidney failure))	<30ml/min/1.73m^2

6.2.4 Results of conventional statistical analysis

In total 16,740 patients were included with 13,861 in sinus rhythm and 2,879 in AF. The median age was 65 years, 23 % were female, median LVEF was 27% (21% to 33%) and baseline eGFR was 63 (IQR 50-77) ml/min/1.73m^2 (see **table 32** for baseline characteristics). In terms of baseline characteristics, the eGFR was higher in the sinus rhythm cohort (64ml/min/1.73m^2) compared to those in AF (60ml/min/1.73m^2). A total of 4,584 patients (27.4%) had stage 3a (mild to moderate) renal impairment, and 2,286 (13.7%) stage 3b (moderate to severely impaired). Only 448 (2.7%) had a severely impaired GFR (<3060ml/min/1.73m^2), as most were excluded from clinical trials. Increasing age, female gender, presence of diabetes and HF of ischaemic origin or prolonged duration were associated with poorer renal function.(93)

Table 32: Baseline characteristics by eGFR category for patients in sinus rhythm

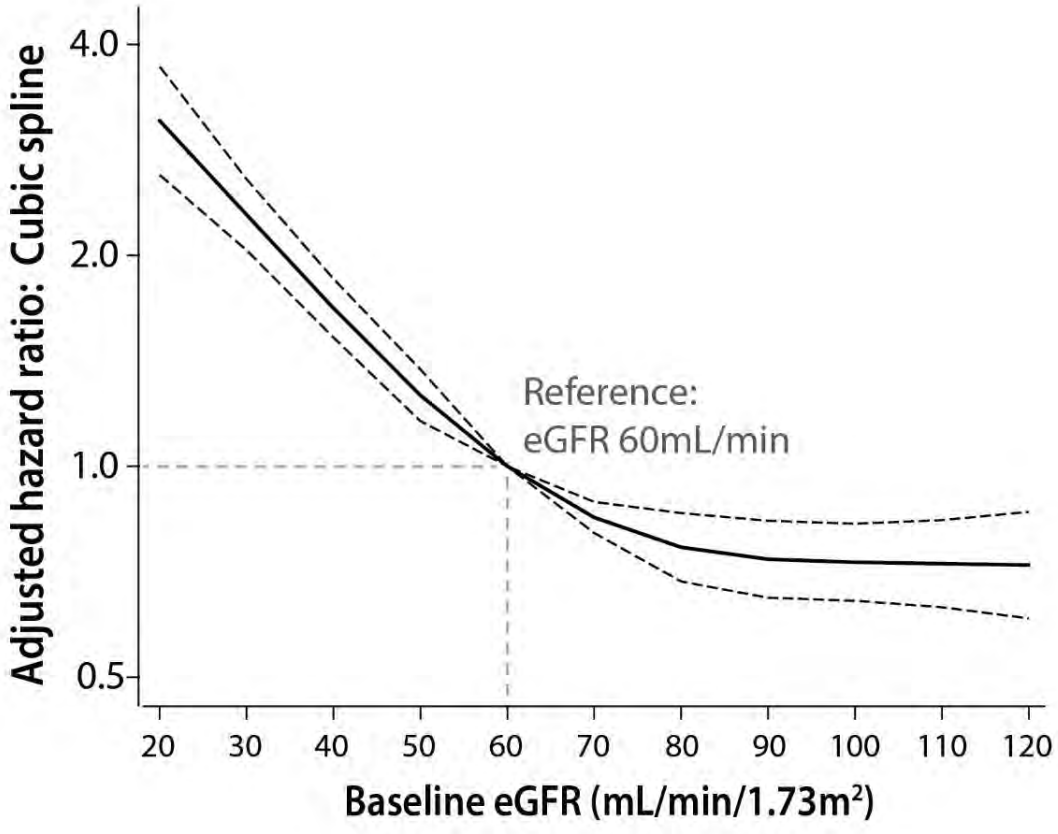
Renal stage / Characteristic	eGFR <30 n=373	eGFR 30-44 n=1,825	eGFR 45-59 n=3,712	eGFR 60-89 n=6,405	eGFR ≥90 n=1,546
Estimated GFR, median mL/min/1.73m^2 (IQR)	26 (24 - 28)	39 (36 - 42)	53 (50 - 57)	72 (66 - 79)	100 (94 - 109)
Age, median years (IQR)	71 (66 - 76)	70 (63 - 75)	67 (59 - 73)	61 (52 - 70)	54 (46 - 62)
Women, %	39.1%	36.2%	27.7%	19.8%	16.6%
Years with HF diagnosis, median (IQR)	4 (2 - 8)	4 (2 - 7)	3 (1 - 6)	2 (1 - 5)	2 (1 - 5)
Ischemic HF etiology, %	80.2%	78.2%	71.8%	66.1%	57.4%
Prior myocardial infarction, %	64.8%	67.4%	62.4%	56.0%	49.4%
Diabetes Mellitus, %	43.0%	31.9%	24.1%	22.1%	23.4%
NYHA class III/IV, %	80.5%	71.8%	67.3%	60.1%	70.3%
LVEF, median % (IQR)	24 (20 - 31)	24 (20 - 31)	25 (20 - 32)	28 (21 - 33)	28 (22 - 33)
Heart rate, median bpm (IQR)	79 (72 - 88)	80 (72 - 88)	80 (72 - 88)	80 (72 - 88)	80 (72 - 90)
Systolic BP, median mmHg (IQR)	124 (110 - 140)	124 (110 - 140)	122 (110 - 138)	122 (110 - 138)	120 (110 - 137)
Diastolic BP, median mmHg (IQR)	72 (65 - 80)	75 (68 - 80)	76 (70 - 81)	78 (70 - 83)	78 (70 - 84)
Body mass index, median kg/m2 (IQR)	25 (23 - 29)	26 (23 - 29)	26 (24 - 30)	27 (24 - 30)	27 (24 - 30)
Any diuretic therapy, %	96.2%	91.0%	87.9%	81.5%	82.2%
ACEI or ARB, %	92.2%	93.0%	95.4%	95.1%	95.6%
Aldosterone antagonists, %	13.6%	11.5%	9.7%	6.9%	5.6%
Digoxin, %	54.7%	53.1%	53.4%	53.7%	58.6%

ACEI, angiotensin converting enzyme inhibitor; ARB, angiotensin receptor blocker; BP, blood pressure; bpm, beats/minute; eGFR, estimated glomerular filtration rate; HF, heart failure; IQR, interquartile range; LVEF, left-ventricular ejection fraction; NYHA, New York Heart Association. eGFR measured in ml/min/1.73m^2

6.2.5 Effect of renal impairment on mortality in heart failure with reduced ejection fraction

In terms of renal impairment, eGFR was independently associated with mortality over a median follow up period of 1.3 years (IQR 0.8-1.9 years). For each 10ml/min/1.73m^2 decline in eGFR the risk of death increased by 12% (95% CI 10%-15%; P<0.001). In those with stage 4 and 5 renal impairment (eGFR <30ml/min/1.73m^2) mortality was higher, often due to advanced heart failure (see **figure 31**).(93)

Figure 31: Impact of renal dysfunction on mortality in heart failure with reduced ejection fraction

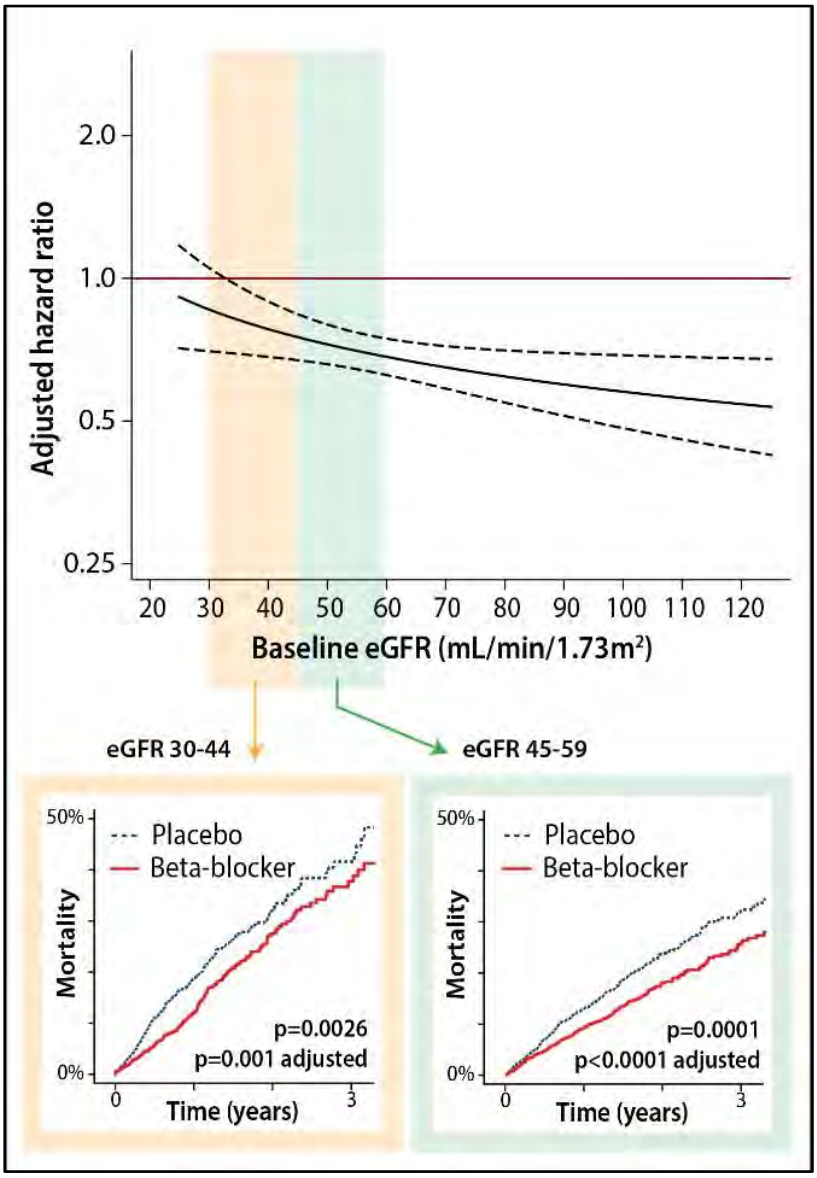


Reproduced with permission from Kotecha D, Gill S. et al, 2019 (93)
The association of renal function at baseline with all-cause mortality in patients with heart failure and reduced ejection fraction. eGFR=estimated glomerular filtration rate.

6.2.6 The efficacy of beta-blockers according to renal function

In sinus rhythm (n=13,861) beta-blockers significantly reduced mortality in all categories of baseline renal function with a HR of 0.71 (95% CI 0.66 to 0.78; p<0.0001), except for severe renal impairment (<30ml/min/1.73m^2), where the number of patients was too small. In sinus rhythm and stage 3a renal impairment (eGFR 45-59ml/min/1.73m^2), the use of beta-blockers compared to placebo yielded an adjusted HR of 0.73 (95% CI 0.62 - 0.86; p<0.0001); and in stage 3b (eGFR 30-44ml/min/1.73m^2) a HR of 0.71 (95% CI 0.58-0.87:p=0.001, see **table 33**). In sinus rhythm, a weak interaction between the efficacy of beta-blockers and renal function was noted (p=0.021: for central 99% of eGFR values; **see figure 32)**.(93)

Figure 32: Beta-blocker efficacy according to baseline renal function in sinus rhythm



Reproduced and adapted from Kotecha D, Gill S et al, 2019 (93).
All-cause mortality of beta-blockers versus placebo in patients with sinus rhythm across renal function. Upper panel: Flexible parametric survival spline plot across renal function. Lower panel: Kaplan-Meier survival plots for stage 3a (mild to moderate) renal impairment (left) and stage 3b (moderate to severe) renal impairment (right).

In AF and HFrEF patients (n=2,879), no change in mortality was seen across all categories of eGFR with the use of beta-blockers, (p value for interaction = 0.18 for the central 99% of eGFR) (see **table 34** and **figure 33**).(93)

Table 33: Hazard ratio according to renal function for beta-blocker versus placebo in sinus rhythm

Sinus rhythm	**eGFR <30**	**eGFR 30-44**	**eGFR 45-59**	**eGFR 60-89**	**eGFR ≥90**
Number of patients with complete data*	372	1,817	3,680	6,372	1,543
Number of deaths (%)	111 (29.8%)	405 (22.3%)	592 (16.1%)	834 (13.1%)	168 (10.9%)
Hazard ratio for beta-blockers versus placebo	1.28	0.71	0.73	0.66	0.64
95% confidence interval	0.87 - 1.91	0.58 - 0.87	0.62 - 0.86	0.57 - 0.76	0.47 - 0.88
p-value	0.35	0.001	<0.0001	<0.0001	0.006

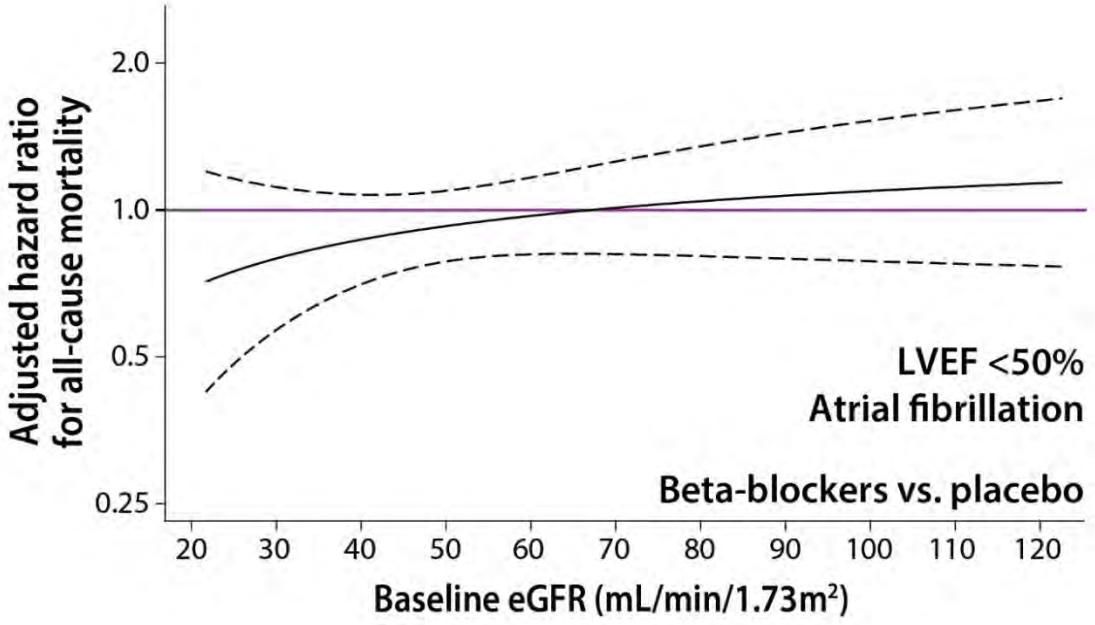
*Including adjustment at baseline for: eGFR, age, gender, left ventricular ejection fraction, history of myocardial infarction, systolic blood pressure, heart rate, use of angiotensin converting enzyme inhibitors or angiotensin receptor blockers, and diuretic therapy. eGFR, estimated glomerular filtration rate ml/min/1.73m^2

Table 34: Hazard ratio according to renal function for beta-blocker versus placebo in atrial fibrillation

Atrial fibrillation	eGFR <30	eGFR 30-44	eGFR 45-59	eGFR 60-89	eGFR ≥90
Number of patients with complete data*	74	458	869	1,230	235
Number of deaths (%)	24 (32.4%)	137 (29.9%)	172 (19.8%)	207 (16.8%)	36 (15.3%)
Hazard ratio for beta-blockers versus placebo	0.58	0.83	1.08	0.97	0.88
95% confidence interval	0.21 - 1.63	0.58 - 1.19	0.80 - 1.47	0.74 - 1.29	0.44 - 1.75
p-value	0.32	0.32	0.59	0.86	0.79

*Including adjustment at baseline for: eGFR, age, gender, left ventricular ejection fraction, history of myocardial infarction, systolic blood pressure, heart rate, use of angiotensin converting enzyme inhibitors or angiotensin receptor blockers, and diuretic therapy. eGFR, estimated glomerular filtration rate measured in ml/min/1.73m^2

Figure 33: Beta-blocker efficacy according to baseline renal function in atrial fibrillation



Reproduced and adapted from Kotecha D, Gill S et al, 2019. Flexible parametric survival plot showing all-cause mortality comparing beta-blockers versus placebo in atrial fibrillation across renal function. eGFR, estimated glomerular filtration rate; LVEF, left ventricular ejection fraction.

6.2.7 Key findings from conventional analysis

In summary beta-blockers conferred a survival benefit in patients with HFrEF and sinus rhythm with mild to moderately severe renal dysfunction. Conventional statistical methods were unable to go beyond this and demonstrate the simultaneous effect complex interactions such as AF and change in renal function over time can have on the outcome. For example, when looking at the effect of renal dysfunction on mortality in HFrEF conventional analysis was only able to detect a weak interaction between beta-blocker efficacy and eGFR with low power ($p=0.021$). Hence, this type of analysis was not capable of identifying sub-groups where the efficacy of beta-blockers may have been clinically significant.(93) With regards to AF, no mortality benefit with beta-blockers was seen in all eGFR categories and higher rates of renal dysfunction and poorer prognoses were noted when compared to placebo. Again, conventional statistical methods were unable to identify characteristics that were associated with an adverse outcome.

Following on from this work we chose to investigate if advanced AI based approaches could better assess the multi-level interactions between co-morbidities and identify clusters of patients in sinus rhythm and AF by beta-blocker efficacy.

6.3 Use of neural network-based approaches to improve phenotyping and prognoses in the presence of multi-morbidity

6.3.1 Rationale for neural network clustering of mortality in heart failure

Despite advances in treatment the mortality from HFrEF especially in the presence of co-morbidities like AF remains high. The first IPD meta-analysis from the beta-blocker collaborative group demonstrated that in HFrEF patients with sinus rhythm (n=13,956), beta-blocker therapy reduced mortality (HR 0.73, 0.67-0.800; p<0.001). In those with AF (n=3,066) this mortality benefit was not seen (HR 0.97, 0.83-1.14; p=0.73).(62, 98) A conventional sub-group analysis of the AF cohort was unable to identify any individual feature that determined efficacy.(62, 98) With the number of living with chronic conditions such as HFrEF and AF, better identification of sub-group responses in AF (and without) is key to ensuring a more personalised approach to care. This is not possible with conventional analysis, but AI-based approaches such as an unsupervised clustering technique, are able to go beyond any pre-conceived ideas and identify interactions between multiple variables. Prior to this study a PubMed search (from start to 25th May 2021) found 391 results of which, 19 studies were evaluated at full text level. Of these, 12 studies described clustering approaches in patients with heart failure. Each of these studies showed clustering methods that lacked transparency, and had a limited assessment of robustness of cluster membership. Only 1 study carried out external validation. Using IPD from the double blind RCTs the hypothesis for this study was that clustering could identify distinct clusters of HFrEF patients with similar responses to beta-blocker efficacy.(98)

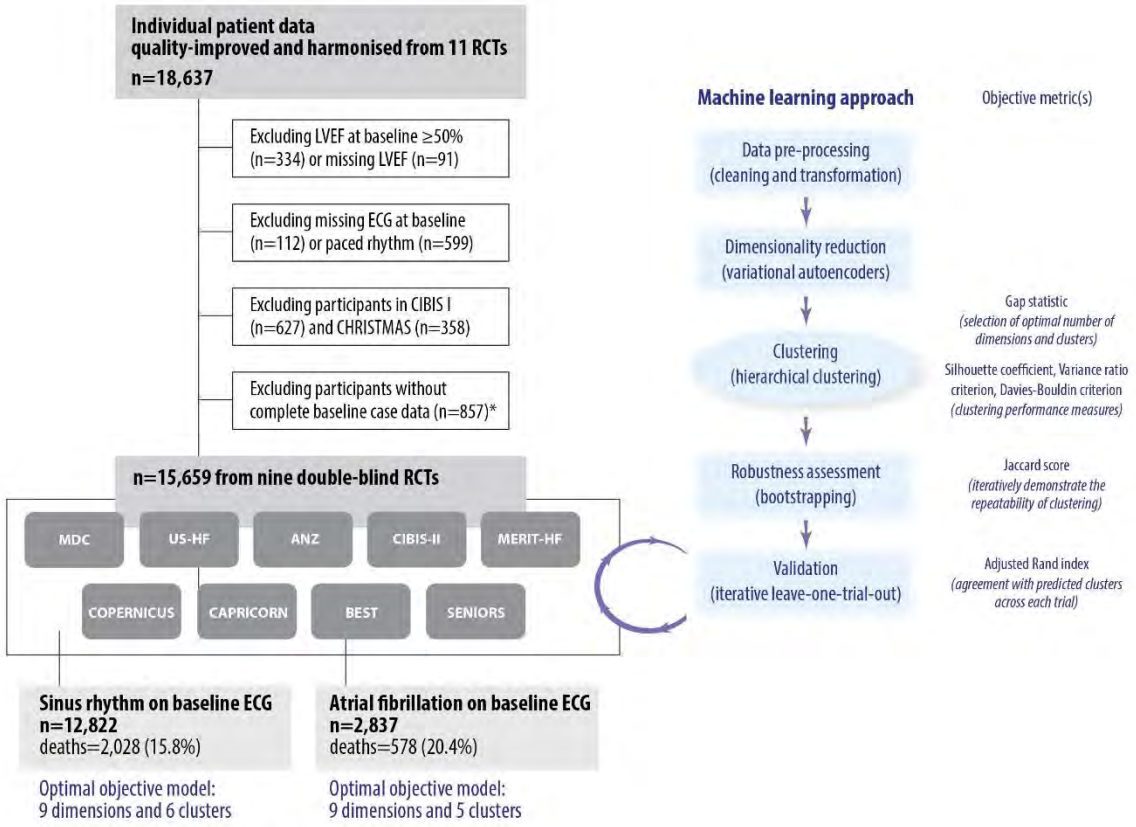
6.3.2 Methods for AI approaches

This study was performed by the card*AI*c group and Health Data Research UK Midlands team using IPD from the eleven RCTs collated by the Beta-blockers in Heart Failure

Collaborative Group.(98) The inclusion criteria and trials included have been outlined above in section 6.2.2. For this analysis nine RCTs were included and two were excluded, these were CHRISTMAS (No AF patients) and CIBIS I (renal function not recorded). IPD from each trial was obtained using a standardised data request form. All data was cross-checked and compared with published reports. To maintain IPD integrity any discrepancies identified were checked against original case report forms. Trial databases were then standardised so that patient characteristics and outcomes matched across all trials. (98)

Those participants with an LVEF ≥50% were excluded as previous analysis had shown no benefit of beta blockers in sinus rhythm or AF.(98, 236) Using their baseline ECG, participants were categorised into sinus rhythm or AF/atrial flutter cohorts, and the primary outcome was all-cause mortality. Several common baseline variables were identified prior to analysis that could help with the identification of cluster phenotypes. These variables were chosen due to their as potential effect on outcome in HFrEF patients and included: age, sex, body mass index, heart rate, systolic blood pressure, LVEF, previous myocardial infarction, NHYA classification, creatinine and baseline medication.(98) See **figure 34** for study flowchart.

Figure 34: Study flowchart



Reproduced with permission from Karwarth A, Bunting K, Gill S. et al, 2021(98)
ECG = electrocardiogram; LVEF = left ventricular ejection fraction; RCT = randomised controlled trials

Data pre-processing and dimensionality reduction

The first step in this analysis involved data pre-processing where any inconsistencies or noise were removed, and missing values identified and handled appropriately. The next stage was to perform dimensionality reduction using autoencoders, a neural network-based approach. An autoencoder consists of several layers, the first (encoder) layer is where the ML model learns to reduce input dimensions and compresses the data. The second is the bottleneck (hidden) layer that contains the compressed (lowest dimension) version of the input data. The third (decoder) layer where the model learns how to reconstruct the original input data from the encoded version, so it resembles the original data.(98)

Clustering approach

Following dimensionality reduction, an unsupervised ML approach called clustering was conducted (by colleague Dr Karwath). Clustering involves identifying a group of data points (clusters) with similar characteristics, which are dissimilar to data points in other groups. There are different clustering techniques available, in this study hierarchical clustering and k-means++ were used. Hierarchical clustering involves uses a tree like structure where new clusters are formed from previous formed ones. The k-means approach partitions datapoints into k-clusters, with each datapoint assigned to its nearest mean (centre of cluster). The number of clusters and dimensions were defined by iteratively repeating the process until iterations reached a maximum or clusters converged.(98, 248). Performance measures were used to evaluate the clustering technique and the model with the highest gap statistic selected.(98)

Evaluation and validation protocol

The final step in the ML analysis was to assess the robustness and validation of the clustering approach. In terms of measuring robustness, a method called bootstrapping where repeated random clustering of data subsets was performed and compared against random cluster assignments.(98, 249) As each of the 9 trials used in this analysis were unique, each one could be used for external validation where clusters were repeatedly generated using the leave one study out approach. The agreement between clustering approaches and a random model was measured using the Adjusted Rand Index (ARI), where 0 is completely random and 1 is perfect matching). The Jaccard index is used to define the similarity of two clusters. A score of 0 indicates no overlap, and 1 complete overlap between sets.(98)

6.3.3 Results of the neural network clustering analysis

A total of 15,659 patients with HFrEF where included (12,822 were in sinus rhythm and 2837 were in AF), the median age was 64 (IQR 55-72), 24% were female and median LVEF was 27% (IQR 21-33), (see **table 35**).

Table 35: Baseline characteristics

Characteristic	All patients n=15,659	Sinus rhythm n=12,822	Atrial fibrillation n=2,837
Age, median years (IQR)	64 (55-72)	64 (54-71)	69 (60-74)
Women, n (%)	3708 (23.7%)	3185 (24.8%)	523 (18.4%)
Body mass index, median kg/m^2 (IQR)	26.6 (24.0-29.8)	26.6 (24.0-29.7)	26.9 (24.3-30.1)
Heart rate, median beats/minute (IQR)	80 (72-88)	80 (72-88)	81 (72-92)
Systolic BP, median mmHg (IQR)	124 (110-140)	123 (110-139)	126 (113-140)
LVEF, median % (IQR)	27 (21-33)	27 (21-33)	27 (21-33)
Prior myocardial infarction, n (%)	8538 (54.5%)	7411 (57.8%)	1127 (39.7%)
NYHA class III/IV, n (%)	8802 (63.7%)	7048 (61.9%)	1754 (72.6%)
Creatinine, µmol/L median (IQR)	105 (88-124)	104 (88-124)	108 (90-131)
ACEI or ARB, n (%)	14877 (95.0%)	12188 (95.1%)	2689 (94.8%)
Any diuretic therapy, n (%)	13563 (86.6%)	10914 (85.1%)	2649 (93.4%)
Anticoagulation therapy, n (%)	5033 (32.1%)	3379 (26.4%)	1654 (58.3%)
Digoxin, n (%)	9299 (59.4%)	6919 (54.0%)	2380 (83.9%)

ACEI = angiotensin converting enzyme inhibitor; ARB = angiotensin receptor blocker; BP = blood pressure; IQR, interquartile range; LVEF, left-ventricular ejection fraction; NYHA, New York Heart Association.

As expected overall the sinus rhythm cohort demonstrated that beta blockers reduced mortality compared to placebo (OR 0.74: 95% CI 0.67-0.81; p<0.001). The clustering approach produced six clusters in sinus rhythm (SR1-SR6) of which all, but one cluster showed that beta-blockers reduced mortality (see **table 36**). In particular, the benefit of beta blockers was seen in clusters with younger non-ischaemic patients (SR5) and those with significantly impaired LVEF and higher mortality rates (SR6). Cluster SR4 showed no significant mortality benefit of beta-blockers (OR 0.86: 95% CI 0.67-1.1; p=0.22). This cluster consisted of 2,537 patients who tended to be older with lower heart rate and less

severe HF symptoms. **Figure 35** below shows the clusters identified in sinus rhythm against all-cause mortality. The mortality risk is represented by green circles, with the size indicating the number of patients in a cluster. The odds ratio below the dotted line represents a mortality benefit from beta-blockers, and the Radar plots summarise the variables used in each cluster, values closer to the outer ring implying a higher-than-average presence.(98)

Overall, in the AF cohort, compared to placebo beta-blockers did not significantly reduce mortality (OR 0.92: CI 077-1.10; p=0.37). The clustering approach produced five clusters (AF1-5) with all but one cluster demonstrating a lack of efficacy with beta blocker therapy. Cluster AF2 (n=659) demonstrated than compared to placebo beta-blockers significantly reduced mortality (OR 0.57: 95% CI 0.35-0.93; p=0.023). This cluster was made up of mostly younger patients with lowers rates of previous myocardial infarction.(98) **Figure 36** below shows the clusters identified in AF against all-cause mortality. The mortality risk is represented by blue circles, with the size indicating the number of patients in a cluster. The odds ratio below the dotted line represents a mortality benefit from beta-blockers, and the Radar plots summarise the variables used in each cluster, values closer to the outer ring implying a higher-than-average presence.

The robustness of each cluster membership was confirmed using bootstrap methods with a mean Jaccard score (defines similarity of clusters) and was compared with a random assignment. Validity of the clustering technique was demonstrated using the leave one study out approach and ARI, which was compared to a random model. (See **table 37 and 38**).(98) In **table 37** shows the robustness of clustering and validity protocol in the 9 dimensions and 6 clusters model for all-cause mortality in sinus rhythm. Both the Jaccard score, and ARI were significantly different to random cluster assignment using repeated random forest models ($p<0.0001$ and $p=0.0198$, respectively). In **table 38** shows the robustness of clustering and

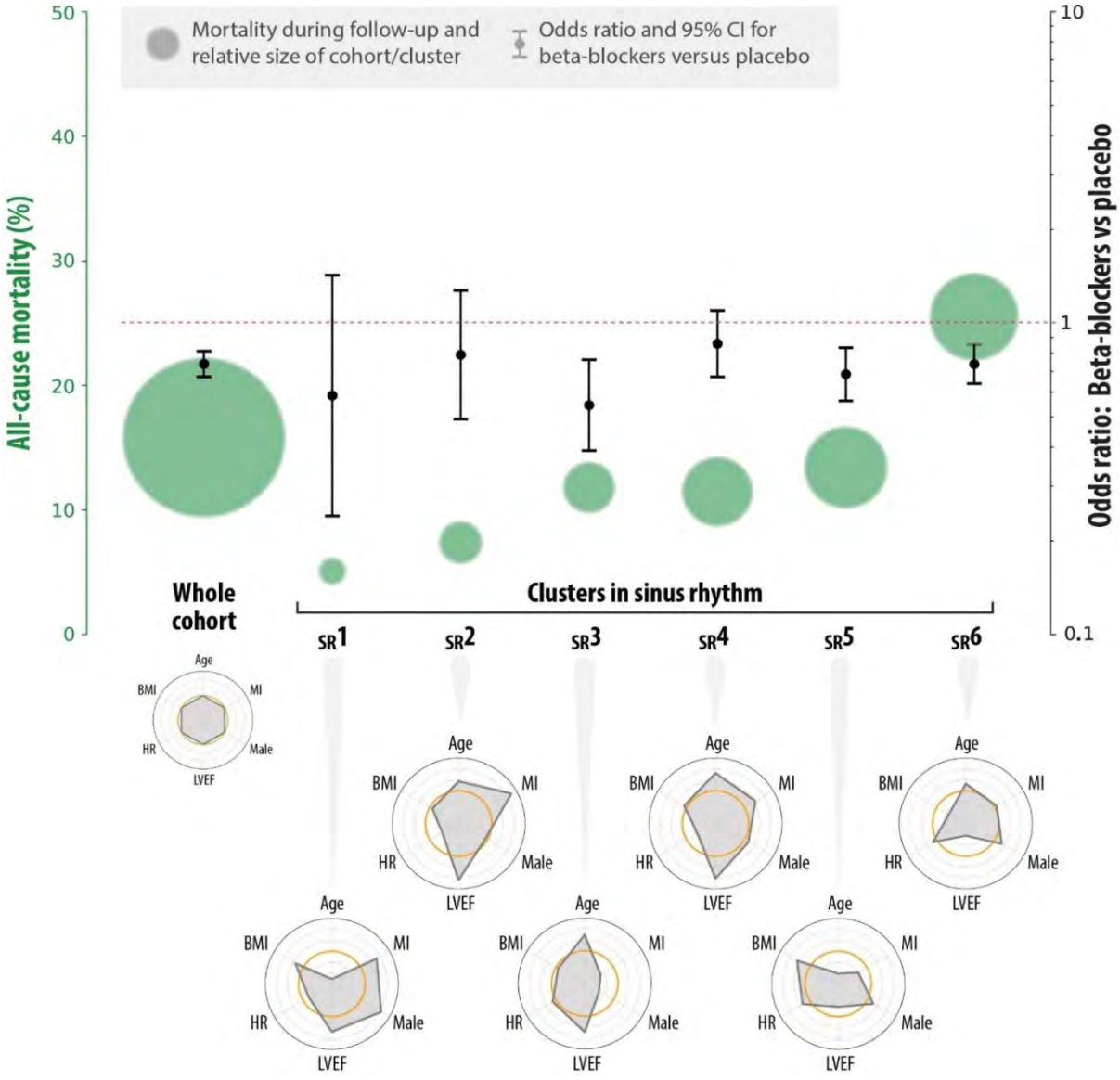
validity protocol in the 9 dimensions and 5 clusters model for all-cause mortality in AF. Both the Jaccard score, and ARI were significantly different to random cluster assignment using repeated random forest models ($p<0.0001$ and $p=0.0264$, respectively).

Table 36: Cluster-specific all-cause mortality results for sinus rhythm and atrial fibrillation

Cluster	**Annualised mortality (%)**	**Placebo Death / Number of patients (% died)**	**Beta-blockers Death / Number of patients (% died)**	**Odds ratio (95% CI); p-value**	**Number needed to treat (95% CI)**
Sinus rhythm					
$_{SR}$ 1	3.9%	14 / 222 (6.3%)	8 / 211 (3.8%)	0.59 (0.24 to 1.43); p=0.24	39.8 (-62.8 to 15.1)
$_{SR}$ 2	5.7%	40 / 487 (8.2%)	34 / 514 (6.6%)	0.79 (0.49 to 1.27); p=0.34	62.5 (-60.6 to 20.6)
$_{SR}$ 3	9.1%	108 / 731 (14.8%)	59 / 683 (8.6%)	0.54 (0.39 to 0.76); p<0.0001	16.3 (10.6 to 35.6)
$_{SR}$ 4	8.8%	151 / 1231 (12.3%)	140 / 1306 (10.7%)	0.86 (0.67 to 1.10); p=0.22	64.7 (-106.6 to 24.8)
$_{SR}$ 5	10.3%	267 / 1706 (15.7%)	202 / 1791 (11.3%)	0.69 (0.56 to 0.83); p<0.0001	22.9 (15.1 to 47.4)
$_{SR}$ 6	19.6%	541 / 1899 (28.5%)	464 / 2041 (22.7%)	0.74 (0.64 to 0.85); p<0.0001	17.4 (11.8 to 33.0)
Atrial fibrillation					
$_{AF}$ 1	13.8%	50 / 307 (16.3%)	59 / 301 (19.6%)	1.25 (0.83 to 1.90); p=0.29	-30.2 (-10.6 to 35.9)
$_{AF}$ 2	9.2%	50 / 338 (14.8%)	29 / 321 (9.0%)	0.57 (0.35 to 0.93); p=0.024	17.4 (9.4 to 118.6)
$_{AF}$ 3	15.1%	68 / 348 (19.5%)	69 / 348 (19.8%)	1.02 (0.70 to 1.48); p=0.92	n/a
$_{AF}$ 4	28.4%	81 / 201 (40.3%)	68 / 202 (33.7%)	0.75 (0.50 to 1.13); p=0.17	15.1 (-36.1 to 6.2)
$_{AF}$ 5	17.0%	51 / 231 (22.1%)	53 / 240 (22.1%)	1.00 (0.65 to 1.55); p=1.0	n/a

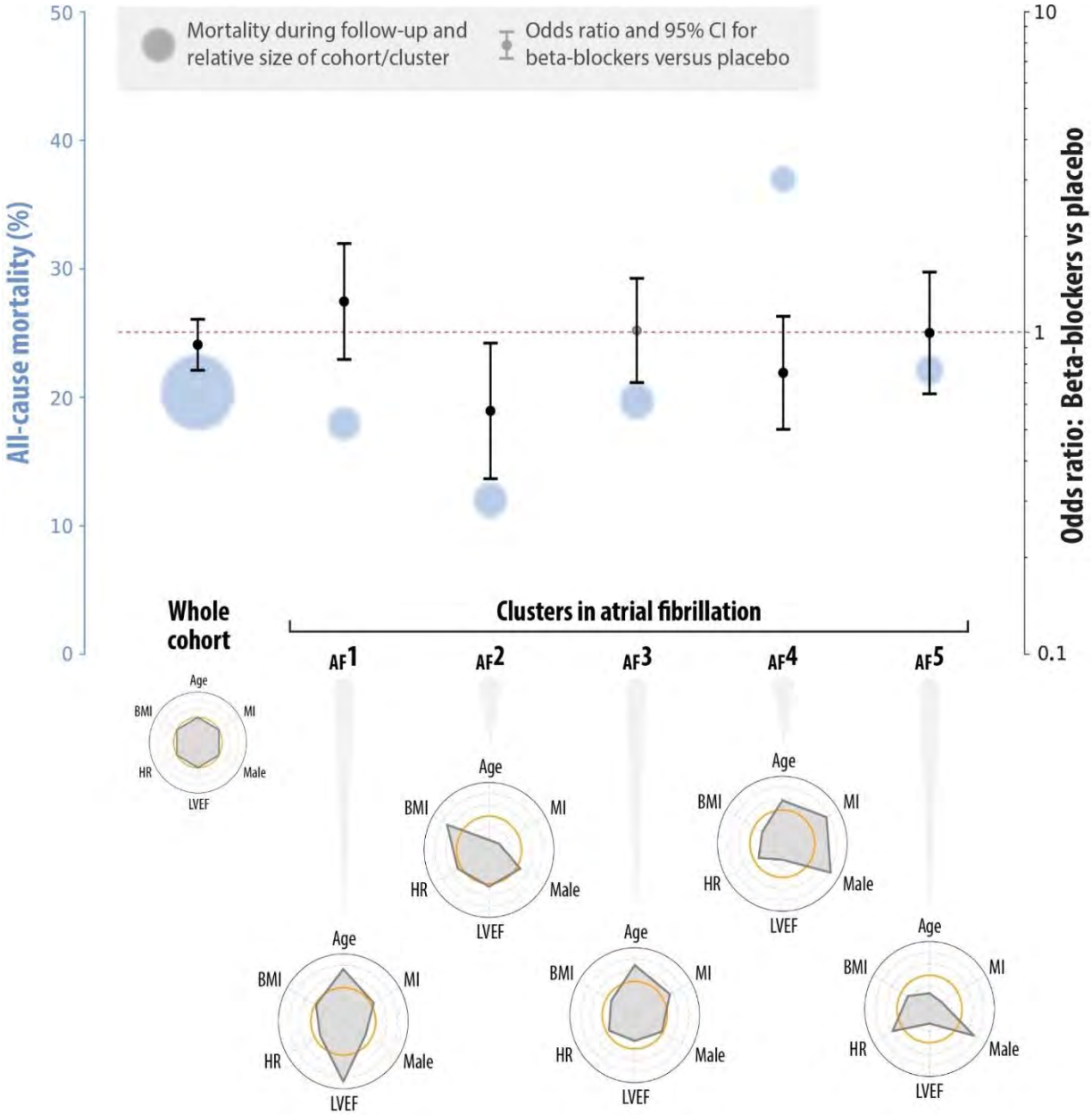
Results based on objective assessment for the number of dimensions and clusters for sinus and atrial fibrillation, as defined by the gap statistic.
n/a = not applicable as the absolute risk reduction with beta-blockers is close or equal to zero.

Figure 35: Clustering for all-cause mortality and beta-blocker efficacy in sinus rhythm



Reproduced with permission from Karwarth A, Bunting K, Gill S. et al, 2021(98)
Mortality risk is represented by the green circles, size indicates number of patients in a cluster. Radar plots summarise the variables used in each cluster. OR=odds ratio; BMI=body mass index; HR=heart rate; LVEF=left ventricular ejection fraction; MI=myocardial infarction; SR =sinus rhythm

Figure 36: Clustering for all-cause mortality and beta-blocker efficacy in atrial fibrillation



Reproduced with permission from Karwarth A, Bunting K, Gill S et al, 2021(98)
Mortality risk is represented by the blue circles, size indicates number of patients in a cluster. Radar plots summarise the variables used in each cluster. OR=odds ratio; BMI=body mass index; HR=heart rate; LVEF=left ventricular ejection fraction; MI=myocardial infarction; SR =sinus rhythm

Table 37: Validation protocol in the sinus rhythm cohort

Study	Number of participants		Adjusted Jaccard score		Adjusted Rand Index	
	Iteration set (all minus trial x)	Prediction set (study x)	Iteration set (all minus trial x)	Prediction set (study x)	Iteration set (all minus trial x)	Prediction set (study x)
CAPRICORN	11249	1573	0.589	0.471	0.533	0.459
US Carvedilol	12072	750	0.532	0.500	0.519	0.416
SENIORS	11693	1129	0.574	0.546	0.577	0.419
MERIT-HF	9625	3197	0.386	0.429	0.484	0.453
CIBIS II	10861	1961	0.708	0.692	0.608	0.591
ANZ	12468	354	0.598	0.372	0.610	0.393
MDC	12560	262	0.644	0.683	0.625	0.467
COPERNICUS	11393	1429	0.558	0.784	0.539	0.640
BEST	10655	2167	0.618	0.688	0.623	0.601
Average			**0.579**	**0.574**	**0.569**	**0.493**

Robustness of clustering and validity protocol in the 9 dimensions and 6 clusters model for all-cause mortality in sinus rhythm. Both the Jaccard score, and Adjusted Rand Index were significantly different to random cluster assignment using repeated random forest models ($p<0.0001$ and $p=0.0198$, respectively).

Table 38: Validation protocol in the atrial fibrillation cohort

Trial (*x*)	Number of participants	Adjusted Jaccard score	Adjusted Rand Index

	Iteration set (all minus trial *x*)	**Prediction set (study *x*)**	**Iteration set (all minus trial *x*)**	**Prediction set (study *x*)**	**Iteration set (all minus trial *x*)**	**Prediction set (study *x*)**
CAPRICORN	2701	136	0.658	0.887	0.585	0.928
US Carvedilol	2703	134	0.528	0.490	0.501	0.360
SENIORS	2337	500	0.592	0.828	0.493	0.873
MERIT-HF	2176	661	0.359	0.539	0.302	0.586
CIBIS II	2329	508	0.471	0.610	0.406	0.535
ANZ	2788	49	0.778	0.712	0.831	0.710
MDC	2787	50	0.646	0.337	0.637	0.073
COPERNICUS	2417	420	0.542	0.381	0.522	0.318
BEST	2458	379	0.718	0.521	0.683	0.400
Average			**0.588**	**0.590**	**0.551**	**0.532**

6.3.4 Key findings from the neural network clustering analysis

Using data from RCTs this AI-based clustering technique identified clusters of patients with HFrEF and sinus rhythm or AF with different treatment responses to beta blockers therapy. In sinus rhythm a cluster of patients was identified where beta-blockers did not confer a mortality benefit, and in AF a cluster where beta-blockers significantly reduced mortality was found. The study demonstrated how AI-based approaches go beyond conventional statistical analysis and assessed how interactions between multiple variables can affect treatment efficacy of beta-blockers in patients with HFrEF. This personalisation identified a cluster of patients in sinus rhythm with HFrEF who did not benefit prognostically from beta-blockers. This finding could address why some patients with HFrEF and sinus rhythm have poorer outcomes despite optimal medical therapy. Avoiding the use of beta-blockers in this group could reduce the risk of side effects and potentially improve their prognoses through optimization of other established therapies. In AF, a cluster of patients was identified who had a significant reduction in mortality with beta-blocker therapy. Radar plots showed that these patients were often younger with less co-morbidities. Identifying these patients early and commencing beta-blocker therapy has the potential to improve their outcome. Further studies in the form of robust RCTs are required to personalise disease management in other chronic cardiac and non-cardiac conditions.

With the use of unsupervised ML algorithms an element of randomness would have occurred during the dimensionality reduction and embedding steps which is inevitable. However, external validation using the leave one study out approach provided reliably and reproducibility of our findings and AI approach. This cohort consisted on HFrEF patients with established AF or sinus rhythm on their baseline ECG. Therefore, these findings may

not be applicable to those with other HF phenotypes, those with lower AF burdens such as paroxysmal AF, and those with AF who do not have HF.

The utilization of AI based approaches to analyse large complex datasets in healthcare remains a relatively novel but rapidly growing field. During this clustering study it became quite clear that there was a lack of clarity amongst healthcare professionals regarding how, when and why AI-based approaches can be used in research. In addition, there is limited information available to educate healthcare professionals and improve their understanding, which may hinder their application and impact patient care in the future. This information can be found in chapter 1.(83)

6.5 Conclusion

To summarise conventional statistical analysis showed the overall efficacy of beta-blockers in patients with HFrEF and renal dysfunction, however these methods were unable to process the interaction between multiple variables simultaneously and could not identify sub-groups where the efficacy of beta-blockers may have been clinically significant.

The clustering technique was able to show how a robust and validated AI-based approach could identify sub-groups with different responses to beta-blocker therapy. Moving forward, further validation of these methods across other conditions and therapies is required, followed by prospective studies to confirm these techniques can improve patient outcomes.(83)

The application of artificial intelligence in healthcare has the potential to significantly improve patient care. An understanding of the overall AI framework from data modalities, selecting appropriate machine learning algorithms, and ensuring model evaluation and validation, can yield findings that can personalise and enhance the delivery of healthcare. In

terms of the future AI is moving so rapidly that it seems nearly impossible to predict where it will be in a few years. With the capability of AI models improving exponentially there is no telling what kind of insights one might obtain. However, there is increasing public concerns about privacy data and safety, especially given recent AI advances and these will need to address centrally with updated General Data Protection Regulation (GDPR).

Chapter 7: Discussion

7.1 Summary of findings

The work described in this thesis has shown potential applications of big data and ML in cardiovascular research, these included: use of smartphone PPG technology for AF detection; AI-based analysis of wearable sensor data as a comparator to conventional trial outcomes, and the identification of new/previously unknown sub-phenotypes based on their prognostic response to established therapies. The main findings which addressed the aims of this PhD are summarised below:

1. A systematic review and meta-analysis of smartphone PPG technology for AF detection, compared to a 1,3 or 12-lead ECG showed unrealistically high sensitivity and specificity on meta-analysis (94% and 97 % respectively), with significant heterogeneity seen across studies ($p<0.0001$). Current evidence was limited to small, mostly single-centre, poor studies. A high risk of bias was noted in these studies, mostly due to patient selection and study design, flow and timings. Additionally, there was evidence of publication bias, with negative findings less likely to be published.
2. Non-medical grade wearable devices were successfully embedded into the RATE-AF trial and collected over 143 million heart rate recordings and over 23 million corresponding physical activity intervals. Low-dose digoxin was found to be non-inferior to beta-blockers for long-term heart rate control, and this was not affected by individual physical activity. The performance of a CNN trained on sensor data was comparable to conventional trial assessments for predicting future NYHA classification.

3. Conventional statistical analyses to assess the impact on beta-blockers in patients with HFrEF and renal impairment showed in sinus rhythm betablockers significantly reduced mortality but was unable to process interactions between multiple variables. Therefore, conventional statistics was unable to identify sub-groups where the efficacy of established therapies such as beta-blockers may or may not be clinically significant. We demonstrated that a robust and well-validated AI-based clustering technique was able to define clusters of beta-blocker efficacy.(98) Overall, patients with HFrEF and sinus rhythm had a mortality benefit with beta-blockers, but our clustering technique identified a sub-group (with associated characteristics), who did not benefit from beta-blockers. Similarly, in AF and HFrEF overall, beta-blockers had no effect on mortality, but one cluster of younger patients were found where they did cause a significant reduction in mortality.

7.2 Smartphone PPG detection of AF

With the rising prevalence of AF, there is a need to increase community screening, especially given the silent sequelae of this condition.(180) Smartphone or smartwatch PPG technology has the potential to offer a large-scale, non-invasive AF screening tool where patients can monitor their own heart rhythm repeatedly. The Apple Heart study,(116) Huawei heart study (250) and Fitbit study (184) have all shown the potential that PPG technology can offer in terms of remote AF detection. However, there are several limitations which hinder its widespread use. A common trend seen in the aforementioned studies was that participants tended to be younger. In the Apple Heart study, the mean age was 41, Huawei Heart study was 35 years and in the Fitbit study it was 47. (116, 184, 250) Hence, remote digital studies incorporating PPG technology (especially smartwatches) leans itself more towards the young, who are also more likely to own this type of technology. This brings with it concerns

regarding low AF detection rates, anxiety and potential morbidity from misdiagnosis and unnecessary treatment.(145) The older population, who are more likely to benefit from AF screening, may decline to participate due to difficulties engaging with technology.(251) In addition, the use of PPG technology requires the consumer to make a costly upfront investment in hardware, excluding those of lower socioeconomic class.

Many of the PPG software applications used for AF detection are from commercial providers and studies to date have provided limited information regarding the ML algorithms used to determine if AF is present. Information regarding how the data is pre-processed (a key step in this type of noisy data), alongside algorithms used to ensure data quality is maintained, and AF is delineated from other common arrhythmias are key if these methods are to be reproduced and validated.

Current ESC guidance describes smartphone and smartwatch PPG technology as opportunistic AF screening tools. If AF is detected using PPG technology a confirmatory ECG is required, due to the lack of validation studies. The accuracy of PPG based wearables for AF detection has been questioned as some studies have found a less favourable performance of these devices when compared with studies reported by industry partners.(16, 252-254) If smartphone PPG is to play a larger role in AF screening, large-scale randomised clinical trials independent from commercial providers are required.

7.3 Use of wearable technology

As described in chapter 5, a sub-study evaluating the use of wearable technology for PPG heart rate monitoring and physical activity was successfully integrated into an on-going RCT. This confirmed the main trial findings that PPG heart rate from wearable devices did not

demonstrate a significant difference in long term heart rate control using low-dose digoxin or beta-blockers.

Traditional clinical trials utilise sporadic single time-point measures to assess treatment efficacy. This is both resource intensive and time-consuming as it requires in-patient visits, and only provides a snapshot of a patient's clinical status. Wearables devices can provide continuous data on heart rate and physical activity that better reflects a patient's day-to-day circadian variation.(110) The vast amount of data generated from these devices was then analysed using a self-training AI-based neural network, whose performance was comparable to a 12-lead ECG and 6MW test for predicting future NYHA classification.

The use of wearable technology in cardiovascular research is growing. From a patient's perspective several advantages exist including; reduced hospital visits, earlier detection of diseases through screening, and continuous monitoring of chronic diseases. From a healthcare perspective there are several cautions that should be noted. As mentioned earlier embracing wearable technology in clinical studies may produce a selection bias by under-representing certain groups, these include: the older age population, ethnic minorities, and those of a lower socioeconomic status. These groups may have less access to digital technology, resulting in a healthcare inequality. (254) Another issue relates to concerns around data privacy and security. The research team will need to uphold data privacy, provide adequate consent options and limit those with data access.(83)

7.4 Conventional and neural network-based analysis

HF is associated with the presence of multiple co-morbidities which complicates the management of this condition and creates a challenge in terms of knowing which patients will benefit from established HF therapies, and which will not.(98) This has been reflected in the

work outlined in this thesis where conventional analysis demonstrated beta-blockers reduce mortality in HFrEF patients in sinus rhythm with renal dysfunction, but this was not seen in those with AF. Chapter 6 discussed the limitations of conventional analysis and how this could be overcome using an AI-based approach capable of processing multiple and higher-dimension interactions that are present in multi-morbidity. An AI-based clustering approach was developed and was able to define clusters based on beta-blocker efficacy in patients with HFrEF and sinus rhythm or AF. The extensive evaluation and validation processes performed, provided reassurance to these findings. HFrEF patients with co-morbidities are often excluded from clinical trials, hence the efficacy of established therapies in this cohort remains unknown. Further development of these processes could facilitate a more personalised approach to treatment in those with multi-morbidity.(98)

Whilst recognising the potential AI-based approaches can offer, careful consideration should be given to its limitations. Adequate pre-processing of data is essential to ensure outputs predicted are relevant clinically. Another issue with AI-based approaches is their lack of transparency, which is known as the black-box effect. Here information regarding the internal workings involved in the development of an AI model are lacking. This is commonly seen in unsupervised learning where the point of choosing this approach is to go beyond data pre-conceptions and identify hidden patterns or associations within data. Hence, it is not always possible to transfer these non-linear workings into a linear form. Despite this, in the clustering approach we were able to follow an AI-based workflow, which was developed in collaboration with Health Data Research-Midlands to standardise methodology.(83)

The speed at which AI technology is evolving is leaving a significant gap in knowledge amongst clinicians, healthcare professionals and the general public. An AI based workflow,

alongside case-based explanations was developed to tackle this and improve understanding, therefore encouraging uptake of AI in clinical studies where appropriate, with adequate consideration of internal and external validation.(83)

7.5: Future work

These initial studies have highlighted the potential wearable devices and AI-based analysis can offer in terms of improving patient care. Going forward, prospective randomised controlled trials that incorporate wearable devices are necessary to assess if continuous data collection from these devices can take the place of conventional trial assessments. In designing any future studies, attempts should be made to ensure underrepresented populations are included. Early involvement of PPI groups is important to ensure studies incorporate the participant viewpoint in all aspects of study design and provide a societal mandate for the use of AI in healthcare.

The BigData@heart initiative, utilised some of the findings of my thesis and moved forward to develop a global framework for how healthcare data can be used. International stakeholders including industry partners, journal editors, regulators, scientists, clinicians and patient representatives collaborated to develop a CODE-EHR framework. This was created to improve study design and provide clarification so that the use of healthcare data in research is utilised more effectively. The CODE-EHR framework is a minimum standards checklist that encourages researchers to check certain criteria (ideally a minimum of five) are met when using EHRs in clinical research.(129)

The follow on the BigData@heart initiative called HYPERMARKER has received funding and will incorporate the findings of my thesis. This EU/UKRI funded programme is

designed to develop clinical decision support for metabolomic-guided, AI-supported precision treatment in patients with hypertension, with patient-devices at home.

References

1. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. Nat Rev Cardiol. 2016;13(6):350-9.

2. Hemingway H, Asselbergs FW, Danesh J, Dobson R, Maniadakis N, Maggioni A, et al. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. Eur Heart J. 2018;39(16):1481-95.

3. Hayashi H, Abe Y, Morita Y, Yamaji Y, Nakane E, Haruna Y, et al. Impact of stroke volume on prognostic outcome in patients with atrial fibrillation and concomitant heart failure with preserved ejection fraction. J Cardiol. 2019;73(4):307-12.

4. Steinberg JS, O'Connell H, Li S, Ziegler PD. Thirty-Second Gold Standard Definition of Atrial Fibrillation and Its Relationship With Subsequent Arrhythmia Patterns: Analysis of a Large Prospective Device Database. Circ Arrhythm Electrophysiol. 2018;11(7):e006274.

5. Andrade J, Khairy P, Dobrev D, Nattel S. The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. Circ Res. 2014;114(9):1453-68.

6. Brundel B, Ai X, Hills MT, Kuipers MF, Lip GYH, de Groot NMS. Atrial fibrillation. Nat Rev Dis Primers. 2022;8(1):21.

7. Zoni-Berisso M, Lercari F, Carazza T, Domenicucci S. Epidemiology of atrial fibrillation: European perspective. Clin Epidemiol. 2014;6:213-20.

8. Krijthe BP, Kunst A, Benjamin EJ, Lip GY, Franco OH, Hofman A, et al. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. Eur Heart J. 2013;34(35):2746-51.

9. Di Carlo A, Bellino L, Consoli D, Mori F, Zaninelli A, Baldereschi M, et al. Prevalence of atrial fibrillation in the Italian elderly population and projections from 2020 to 2060 for Italy and the European Union: the FAI Project. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2019;21(10):1468-75.

10. Staerk L, Sherer JA, Ko D, Benjamin EJ, Helm RH. Atrial Fibrillation: Epidemiology, Pathophysiology, and Clinical Outcomes. Circ Res. 2017;120(9):1501-17.

11. Aune D, Feng T, Schlesinger S, Janszky I, Norat T, Riboli E. Diabetes mellitus, blood glucose and the risk of atrial fibrillation: A systematic review and meta-analysis of cohort studies. J Diabetes Complications. 2018;32(5):501-11.

12. Cadby G, McArdle N, Briffa T, Hillman DR, Simpson L, Knuiman M, et al. Severity of OSA is an independent predictor of incident atrial fibrillation hospitalization in a large sleep-clinic cohort. Chest. 2015;148(4):945-52.

13. Hobbelt AH, Siland JE, Geelhoed B, Van Der Harst P, Hillege HL, Van Gelder IC, et al. Clinical, biomarker, and genetic predictors of specific types of atrial fibrillation in a community-based cohort: data of the PREVEND study. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2017;19(2):226-32.

14. Staerk L, Preis SR, Lin H, Lubitz SA, Ellinor PT, Levy D, et al. Protein Biomarkers and Risk of Atrial Fibrillation: The FHS. Circ Arrhythm Electrophysiol. 2020;13(2):e007607.

15. Markides V, Schilling RJ. Atrial fibrillation: classification, pathophysiology, mechanisms and drug treatment. Heart. 2003;89(8):939-43.

16. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. Eur Heart J. 2021;42(5):373-498.

17. Boriani G, Pettorelli D. Atrial fibrillation burden and atrial fibrillation type: Clinical significance and impact on the risk of stroke and decision making for long-term anticoagulation. Vascul Pharmacol. 2016;83:26-35.

18. Rho RW, Page RL. Asymptomatic atrial fibrillation. Prog Cardiovasc Dis. 2005;48(2):79-87.

19. Furberg CD, Psaty BM, Manolio TA, Gardin JM, Smith VE, Rautaharju PM. Prevalence of atrial fibrillation in elderly subjects (the Cardiovascular Health Study). Am J Cardiol. 1994;74(3):236-41.

20. Wynn GJ, Todd DM, Webber M, Bonnett L, McShane J, Kirchhof P, et al. The European Heart Rhythm Association symptom classification for atrial fibrillation: validation and improvement through a simple modification. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2014;16(7):965-72.

21. Pisters R, Lane DA, Marin F, Camm AJ, Lip GY. Stroke and thromboembolism in atrial fibrillation. Circ J. 2012;76(10):2289-304.

22. Killu AM, Granger CB, Gersh BJ. Risk stratification for stroke in atrial fibrillation: a critique. Eur Heart J. 2019;40(16):1294-302.

23. Rivera-Caravaca JM, Roldán V, Esteve-Pastor MA, Valdés M, Vicente V, Lip GYH, et al. Long-Term Stroke Risk Prediction in Patients With Atrial Fibrillation: Comparison of the ABC-Stroke and CHA(2)DS(2)-VASc Scores. J Am Heart Assoc. 2017;6(7).

24. Fox KAA, Lucas JE, Pieper KS, Bassand JP, Camm AJ, Fitzmaurice DA, et al. Improved risk stratification of patients with atrial fibrillation: an integrated GARFIELD-AF tool for the prediction of mortality, stroke and bleed in patients with and without anticoagulation. BMJ open. 2017;7(12):e017157.

25. Zhu W, Fu L, Ding Y, Huang L, Xu Z, Hu J, et al. Meta-analysis of ATRIA versus CHA(2)DS(2)-VASc for predicting stroke and thromboembolism in patients with atrial fibrillation. International journal of cardiology. 2017;227:436-42.

26. 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):263-72.

27. Kirchhof P, Camm AJ, Goette A, Brandes A, Eckardt L, Elvan A, et al. Early Rhythm-Control Therapy in Patients with Atrial Fibrillation. The New England journal of medicine. 2020;383(14):1305-16.

28. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res. 2023;118(17):3272-87.

29. Castiglione V, Aimo A, Vergaro G, Saccaro L, Passino C, Emdin M. Biomarkers for the diagnosis and management of heart failure. Heart Fail Rev. 2022;27(2):625-43.

30. Kemp CD, Conte JV. The pathophysiology of heart failure. Cardiovasc Pathol. 2012;21(5):365-71.

31. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J. 2021;42(36):3599-726.

32. van Riet EE, Hoes AW, Wagenaar KP, Limburg A, Landman MA, Rutten FH. Epidemiology of heart failure: the prevalence of heart failure and ventricular dysfunction in older adults over time. A systematic review. Eur J Heart Fail. 2016;18(3):242-52.

33. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, et al. Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. Circulation. 2018;137(12):e67-e492.

34. Miranda-Silva D, Lima T, Rodrigues P, Leite-Moreira A, Falcão-Pires I. Mechanisms underlying the pathophysiology of heart failure with preserved ejection fraction: the tip of the iceberg. Heart Fail Rev. 2021;26(3):453-78.

35. Zhang R, Ma S, Shanahan L, Munroe J, Horn S, Speedie S. Discovering and identifying New York heart association classification from electronic health records. BMC Med Inform Decis Mak. 2018;18(Suppl 2):48.

36. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). Lancet. 1999;353(9169):2001-7.

37. Packer M, Bristow MR, Cohn JN, Colucci WS, Fowler MB, Gilbert EM, et al. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. U.S. Carvedilol Heart Failure Study Group. The New England journal of medicine. 1996;334(21):1349-55.

38. Enalapril for congestive heart failure. The New England journal of medicine. 1987;317(21):1349-51.

39. Garg R, Yusuf S. Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials. Jama. 1995;273(18):1450-6.

40. McMurray JJ, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, et al. Angiotensin-neprilysin inhibition versus enalapril in heart failure. The New England journal of medicine. 2014;371(11):993-1004.

41. Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A, et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. The New England journal of medicine. 1999;341(10):709-17.

42. Zannad F, McMurray JJ, Krum H, van Veldhuisen DJ, Swedberg K, Shi H, et al. Eplerenone in patients with systolic heart failure and mild symptoms. The New England journal of medicine. 2011;364(1):11-21.

43. McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, et al. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. The New England journal of medicine. 2019;381(21):1995-2008.

44. Glikson M, Nielsen JC, Kronborg MB, Michowitz Y, Auricchio A, Barbash IM, et al. 2021 ESC Guidelines on cardiac pacing and cardiac resynchronization therapy. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2022;24(1):71-164.

45. Tang AS, Wells GA, Talajic M, Arnold MO, Sheldon R, Connolly S, et al. Cardiac-resynchronization therapy for mild-to-moderate heart failure. The New England journal of medicine. 2010;363(25):2385-95.

46. Cleland JG, Daubert JC, Erdmann E, Freemantle N, Gras D, Kappenberger L, et al. The effect of cardiac resynchronization on morbidity and mortality in heart failure. The New England journal of medicine. 2005;352(15):1539-49.

47. Køber L, Thune JJ, Nielsen JC, Haarbo J, Videbæk L, Korup E, et al. Defibrillator Implantation in Patients with Nonischemic Systolic Heart Failure. The New England journal of medicine. 2016;375(13):1221-30.

48. Acosta J, Fernández-Armenta J, Borràs R, Anguera I, Bisbal F, Martí-Almor J, et al. Scar Characterization to Predict Life-Threatening Arrhythmic Events and Sudden Cardiac Death in Patients With Cardiac Resynchronization Therapy: The GAUDI-CRT Study. JACC Cardiovasc Imaging. 2018;11(4):561-72.

49. Leyva F, Zegard A, Acquaye E, Gubran C, Taylor R, Foley PWX, et al. Outcomes of Cardiac Resynchronization Therapy With or Without Defibrillation in Patients With Nonischemic Cardiomyopathy. J Am Coll Cardiol. 2017;70(10):1216-27.

50. Anker SD, Butler J, Filippatos G, Ferreira JP, Bocchi E, Böhm M, et al. Empagliflozin in Heart Failure with a Preserved Ejection Fraction. The New England journal of medicine. 2021;385(16):1451-61.

51. Vaduganathan M, Docherty KF, Claggett BL, Jhund PS, de Boer RA, Hernandez AF, et al. SGLT-2 inhibitors in patients with heart failure: a comprehensive meta-analysis of five randomised controlled trials. Lancet. 2022;400(10354):757-67.

52. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, et al. 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. J Am Coll Cardiol. 2022;79(17):1757-80.

53. Reddy YNV, Borlaug BA, Gersh BJ. Management of Atrial Fibrillation Across the Spectrum of Heart Failure With Preserved and Reduced Ejection Fraction. Circulation. 2022;146(4):339-57.

54. Santhanakrishnan R, Wang N, Larson MG, Magnani JW, McManus DD, Lubitz SA, et al. Atrial Fibrillation Begets Heart Failure and Vice Versa: Temporal Associations and Differences in Preserved Versus Reduced Ejection Fraction. Circulation. 2016;133(5):484-92.

55. Kotecha D, Piccini JP. Atrial fibrillation in heart failure: what should we do? European heart journal. 2015;36(46):3250-7.

56. Kotecha D, Lam CS, Van Veldhuisen DJ, Van Gelder IC, Voors AA, Rienstra M. Heart Failure With Preserved Ejection Fraction and Atrial Fibrillation: Vicious Twins. J Am Coll Cardiol. 2016;68(20):2217-28.

57. Verhaert DVM, Brunner-La Rocca HP, van Veldhuisen DJ, Vernooy K. The bidirectional interaction between atrial fibrillation and heart failure: consequences for the management of both diseases. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2021;23(23 Suppl 2):ii40-ii5.

58. Trulock KM, Narayan SM, Piccini JP. Rhythm control in heart failure patients with atrial fibrillation: contemporary challenges including the role of ablation. J Am Coll Cardiol. 2014;64(7):710-21.

59. Deedwania PC, Lardizabal JA. Atrial fibrillation in heart failure: a comprehensive review. Am J Med. 2010;123(3):198-204.

60. Hartupee J, Mann DL. Neurohormonal activation in heart failure with reduced ejection fraction. Nat Rev Cardiol. 2017;14(1):30-8.

61. Anter E, Jessup M, Callans DJ. Atrial fibrillation and heart failure: treatment considerations for a dual epidemic. Circulation. 2009;119(18):2516-25.

62. Kotecha D, Holmes J, Krum H, Altman DG, Manzano L, Cleland JG, et al. Efficacy of beta blockers in patients with heart failure plus atrial fibrillation: an individual-patient data meta-analysis. Lancet. 2014;384(9961):2235-43.

63. Kotecha D, Flather MD, Altman DG, Holmes J, Rosano G, Wikstrand J, et al. Heart Rate and Rhythm and the Benefit of Beta-Blockers in Patients With Heart Failure. Journal of the American College of Cardiology. 2017;69(24):2885-96.

64. Ziff OJ, Lane DA, Samra M, Griffith M, Kirchhof P, Lip GY, et al. Safety and efficacy of digoxin: systematic review and meta-analysis of observational and controlled trial data. Bmj. 2015;351:h4451.

65. Cleland JGF, Bunting KV, Flather MD, Altman DG, Holmes J, Coats AJS, et al. Beta-blockers for heart failure with reduced, mid-range, and preserved ejection fraction: an individual patient-level analysis of double-blind randomized trials. Eur Heart J. 2018;39(1):26-35.

66. Goldstein RE, Boccuzzi SJ, Cruess D, Nattel S. Diltiazem increases late-onset congestive heart failure in postinfarction patients with early reduction in ejection fraction. The Adverse Experience Committee; and the Multicenter Diltiazem Postinfarction Research Group. Circulation. 1991;83(1):52-60.

67. Elkayam U. Calcium channel blockers in heart failure. Cardiology. 1998;89 Suppl 1:38-46.

68. Hagens VE, Crijns HJ, Van Veldhuisen DJ, Van Den Berg MP, Rienstra M, Ranchor AV, et al. Rate control versus rhythm control for patients with persistent atrial fibrillation with mild to moderate heart failure: results from the RAte Control versus Electrical cardioversion (RACE) study. Am Heart J. 2005;149(6):1106-11.

69. Chen S, Pürerfellner H, Meyer C, Acou WJ, Schratter A, Ling Z, et al. Rhythm control for patients with atrial fibrillation complicated with heart failure in the contemporary era of catheter ablation: a stratified pooled analysis of randomized data. Eur Heart J. 2020;41(30):2863-73.

70. Rillig A, Magnussen C, Ozga AK, Suling A, Brandes A, Breithardt G, et al. Early Rhythm Control Therapy in Patients With Atrial Fibrillation and Heart Failure. Circulation. 2021;144(11):845-58.

71. Olsson LG, Swedberg K, Ducharme A, Granger CB, Michelson EL, McMurray JJ, et al. Atrial fibrillation and risk of clinical events in chronic heart failure with and without left ventricular systolic dysfunction: results from the Candesartan in Heart failure-Assessment of Reduction in Mortality and morbidity (CHARM) program. J Am Coll Cardiol. 2006;47(10):1997-2004.

72. Yusuf S, Healey JS, Pogue J, Chrolavicius S, Flather M, Hart RG, et al. Irbesartan in patients with atrial fibrillation. The New England journal of medicine. 2011;364(10):928-38.

73. Ţica O, Khamboo W, Kotecha D. Breaking the Cycle of Heart Failure With Preserved Ejection Fraction and Atrial Fibrillation. Card Fail Rev. 2022;8:e32.

74. Packer M, Butler J, Zannad F, Filippatos G, Ferreira JP, Pocock SJ, et al. Effect of Empagliflozin on Worsening Heart Failure Events in Patients With Heart Failure and Preserved Ejection Fraction: EMPEROR-Preserved Trial. Circulation. 2021;144(16):1284-94.

75. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. J Am Coll Cardiol. 2017;69(21):2657-64.

76. Alonso SG, de la Torre Diez I, Rodrigues J, Hamrioui S, Lopez-Coronado M. A Systematic Review of Techniques and Sources of Big Data in the Healthcare Sector. J Med Syst. 2017;41(11):183.

77. Austin C, Kusumoto F. The application of Big Data in medicine: current implications and future directions. J Interv Card Electrophysiol. 2016;47(1):51-9.

78. Kim J. Big Data, Health Informatics, and the Future of Cardiovascular Medicine. J Am Coll Cardiol. 2017;69(7):899-902.

79. Hermon R, Williams PA. Big data in healthcare: What is it used for? 2014.

80. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? Heart. 2018;104(14):1156-64.

81. Silverio A, Cavallo P, De Rosa R, Galasso G. Big Health Data and Cardiovascular Diseases: A Challenge for Research, an Opportunity for Clinical Care. Front Med (Lausanne). 2019;6:36.

82. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2095-128.

83. Gill SK, Karwath A, Uh HW, Cardoso VR, Gu Z, Barsky A, et al. Artificial intelligence to enhance clinical value across the spectrum of cardiovascular healthcare. Eur Heart J. 2023.

84. Tran J, Norton R, Conrad N, Rahimian F, Canoy D, Nazarzadeh M, et al. Patterns and temporal trends of comorbidity among adult patients with incident cardiovascular disease in the UK between 2000 and 2014: A population-based cohort study. PLoS Med. 2018;15(3):e1002513.

85. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. Nat Biotechnol. 2014;32(1):40-51.

86. Denaxas SC, Morley KI. Big biomedical data and cardiovascular disease research: opportunities and challenges. Eur Heart J Qual Care Clin Outcomes. 2015;1(1):9-16.

87. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019;20(5):e262-e73.

88. Lumbers RT, Shah S, Lin H, Czuba T, Henry A, Swerdlow DI, et al. The genomics of heart failure: design and rationale of the HERMES consortium. ESC Heart Fail. 2021;8(6):5531-41.

89. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun. 2020;11(1):163.

90. Benjamins JW, Hendriks T, Knuuti J, Juarez-Orozco LE, van der Harst P. A primer in artificial intelligence in cardiovascular medicine. Neth Heart J. 2019;27(9):392-402.

91. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Medical Research Methodology. 2019;19(1):64.

92. Sánchez de la Nava AM, Atienza F, Bermejo J, Fernández-Avilés F. Artificial intelligence for a personalized diagnosis and treatment of atrial fibrillation. Am J Physiol Heart Circ Physiol. 2021;320(4):H1337-h47.

93. Kotecha D, Gill SK, Flather MD, Holmes J, Packer M, Rosano G, et al. Impact of Renal Impairment on Beta-Blocker Efficacy in Patients With Heart Failure. J Am Coll Cardiol. 2019;74(23):2893-904.

94. Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography. J Am Coll Cardiol. 2016;68(21):2287-95.

95. Seetharam K, Min JK. Artificial Intelligence and Machine Learning in Cardiovascular Imaging. Methodist Debakey Cardiovasc J. 2020;16(4):263-71.

96. Mathur P, Srivastava S, Xu X, Mehta JL. Artificial Intelligence, Machine Learning, and Cardiovascular Disease. Clin Med Insights Cardiol. 2020;14:1179546820927404.

97. Krittanawong C, Johnson KW, Rosenson RS, Wang Z, Aydar M, Baber U, et al. Deep learning for cardiovascular medicine: a practical primer. Eur Heart J. 2019;40(25):2058-73.

98. Karwath A, Bunting KV, Gill SK, Tica O, Pendleton S, Aziz F, et al. Redefining β-blocker response in heart failure patients with sinus rhythm and atrial fibrillation: a machine learning cluster analysis. Lancet. 2021;398(10309):1427-35.

99. Agrawal R, Prabakaran S. Big data in digital healthcare: lessons learnt and recommendations for general practice. Heredity (Edinb). 2020;124(4):525-34.

100. Camacho DM, Collins KM, Powers RK, Costello JC, Collins JJ. Next-Generation Machine Learning for Biological Networks. Cell. 2018;173(7):1581-92.

101. Cuocolo R, Perillo T, De Rosa E, Ugga L, Petretta M. Current applications of big data and machine learning in cardiology. J Geriatr Cardiol. 2019;16(8):601-7.

102. Khurshid S, Weng LC, Al-Alusi MA, Halford JL, Haimovich JS, Benjamin EJ, et al. Accelerometer-derived physical activity and risk of atrial fibrillation. Eur Heart J. 2021;42(25):2472-83.

103. Vardas PE, Asselbergs FW, van Smeden M, Friedman P. The year in cardiovascular medicine 2021: digital health and innovation. Eur Heart J. 2022.

104. Alaa AM, Bolton T, Di Angelantonio E, Rudd JHF, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. PLoS One. 2019;14(5):e0213653.

105. Krittanawong C, Johnson KW, Tang WW. How artificial intelligence could redefine clinical trials in cardiovascular medicine: lessons learned from oncology. Per Med. 2019;16(2):83-8.

106. Lee CS, Lee AY. How Artificial Intelligence Can Transform Randomized Controlled Trials. Transl Vis Sci Technol. 2020;9(2):9.

107. Godfrey A, Hetherington V, Shum H, Bonato P, Lovell NH, Stuart S. From A to Z: Wearable technology explained. Maturitas. 2018;113:40-7.

108. Jin D, Adams H, Cocco AM, Martin WG, Palmer S. Smartphones and wearable technology: benefits and concerns in cardiology. Medical Journal of Australia. 2020;212(2):54-6.e1.

109. Perry B, Herrington W, Goldsack JC, Grandinetti CA, Vasisht KP, Landray MJ, et al. Use of Mobile Devices to Measure Outcomes in Clinical Research, 2010-2016: A Systematic Literature Review. Digital biomarkers. 2018;2(1):11-30.

110. Bayoumy K, Gaber M, Elshafeey A, Mhaimeed O, Dineen EH, Marvel FA, et al. Smart wearable devices in cardiovascular care: where we are and how to move forward. Nat Rev Cardiol. 2021;18(8):581-99.

111. Yang CC, Hsu YL. A review of accelerometry-based wearable motion detectors for physical activity monitoring. Sensors (Basel). 2010;10(8):7772-88.

112. Cobos Gil M. Standard and Precordial Leads Obtained With an Apple Watch. Ann Intern Med. 2020;172(6):436-7.

113. Bae S, Borac S, Emre Y, Wang J, Wu J, Kashyap M, et al. Prospective validation of smartphone-based heart rate and respiratory rate measurement algorithms. Commun Med (Lond). 2022;2:40.

114. Pelliccia A, Sharma S, Gati S, Bäck M, Börjesson M, Caselli S, et al. 2020 ESC Guidelines on sports cardiology and exercise in patients with cardiovascular disease. Eur Heart J. 2021;42(1):17-96.

115. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. Jama. 2018;320(2):146-55.

116. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, et al. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. New England journal of medicine. 2019;381(20):1909-17.

117. Borgundvaag E, Janssen I. Objectively Measured Physical Activity and Mortality Risk Among American Adults. Am J Prev Med. 2017;52(1):e25-e31.

118. Dohrn IM, Sjöström M, Kwak L, Oja P, Hagströmer M. Accelerometer-measured sedentary time and physical activity-A 15 year follow-up of mortality in a Swedish population-based cohort. J Sci Med Sport. 2018;21(7):702-7.

119. Lee IM, Shiroma EJ, Evenson KR, Kamada M, LaCroix AZ, Buring JE. Accelerometer-Measured Physical Activity and Sedentary Behavior in Relation to All-Cause Mortality: The Women's Health Study. Circulation. 2018;137(2):203-5.

120. Ballinger B, Hsieh J, Singh A, Sohoni N, Wang J, Tison G, et al., editors. DeepHeart: semi-supervised sequence learning for cardiovascular risk prediction. Proceedings of the AAAI conference on artificial intelligence; 2018.

121. Gibson CM, Steinhubl S, Lakkireddy D, Turakhia MP, Passman R, Jones WS, et al. Does Early Detection of Atrial Fibrillation Reduce the Risk of Thromboembolic Events? Rationale and Design of the Heartline Study. Am Heart J. 2023.

122. Anderson L, Sharp GA, Norton RJ, Dalal H, Dean SG, Jolly K, et al. Home-based versus centre-based cardiac rehabilitation. Cochrane Database Syst Rev. 2017;6(6):Cd007130.

123. Pereira T, Tran N, Gadhoumi K, Pelter MM, Do DH, Lee RJ, et al. Photoplethysmography based atrial fibrillation detection: a review. NPJ Digit Med. 2020;3:3.

124. Elul Y, Rosenberg AA, Schuster A, Bronstein AM, Yaniv Y. Meeting the unmet needs of clinicians from AI systems showcased for cardiology with deep-learning-based ECG analysis. Proc Natl Acad Sci U S A. 2021;118(24).

125. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-e97.

126. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. BMJ. 2020;370:m3164.

127. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2(10):e549-e60.

128. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018.

129. Kotecha D, Asselbergs FW, European Society of Cardiology, BigData@Heart consortium. Best practice framework for the use of coded electronic healthcare records in clinical research (CoDe-EHR checklist): International stakeholder consensus from the European Society of Cardiology and the BigData@Heart consortium. 2021: https://www.escardio.org/Research/Big-Data-Heart [Accessed 07 Apr 2021].

130. Karwath A, Bunting KV, Gill SK, Tica O, Pendleton S, Aziz F, et al. Redefining beta-blocker response in heart failure patients with sinus rhythm and atrial fibrillation: a machine learning cluster analysis. Lancet. 2021.

131. Kornej J, Börschel CS, Benjamin EJ, Schnabel RB. Epidemiology of Atrial Fibrillation in the 21st Century: Novel Methods and New Insights. Circ Res. 2020;127(1):4-20.

132. Schnabel RB, Haeusler KG, Healey JS, Freedman B, Boriani G, Brachmann J, et al. Searching for Atrial Fibrillation Poststroke: A White Paper of the AF-SCREEN International Collaboration. Circulation. 2019;140(22):1834-50.

133. Andersson T, Magnuson A, Bryngelsson IL, Frøbert O, Henriksson KM, Edvardsson N, et al. All-cause mortality in 272,186 patients hospitalized with incident atrial fibrillation 1995-2008: a Swedish nationwide long-term case-control study. Eur Heart J. 2013;34(14):1061-7.

134. Kim MH, Johnston SS, Chu BC, Dalal MR, Schulman KL. Estimation of total incremental health care costs in patients with atrial fibrillation in the United States. Circ Cardiovasc Qual Outcomes. 2011;4(3):313-20.

135. Wattigney WA, Mensah GA, Croft JB. Increased atrial fibrillation mortality: United States, 1980-1998. Am J Epidemiol. 2002;155(9):819-26.

136. Svennberg E, Engdahl J, Al-Khalili F, Friberg L, Frykman V, Rosenqvist M. Mass Screening for Untreated Atrial Fibrillation: The STROKESTOP Study. Circulation. 2015;131(25):2176-84.

137. Lubitz SA, Atlas SJ, Ashburner JM, Lipsanopoulos ATT, Borowsky LH, Guan W, et al. Screening for Atrial Fibrillation in Older Adults at Primary Care Visits: VITAL-AF Randomized Controlled Trial. Circulation. 2022;145(13):946-54.

138. Dilaveris PE, Kennedy HL. Silent atrial fibrillation: epidemiology, diagnosis, and clinical impact. Clin Cardiol. 2017;40(6):413-8.

139. Lip GY, Tse HF, Lane DA. Atrial fibrillation. Lancet. 2012;379(9816):648-61.

140. Freedman B. Screening for Atrial Fibrillation Using a Smartphone: Is There an App for That? J Am Heart Assoc. 2016;5(7).

141. Lopez Perales CR, Van Spall HGC, Maeda S, Jimenez A, Laţcu DG, Milman A, et al. Mobile health applications for the detection of atrial fibrillation: a systematic review. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2020.

142. Lowres N, Neubeck L, Redfern J, Freedman SB. Screening to identify unknown atrial fibrillation. A systematic review. Thromb Haemost. 2013;110(2):213-22.

143. Grieten L, Haelst RV, Mortelmans C, Vaes B. The diagnostic accuracy of smartphone applications to detect atrial fibrillation: A head-to-head comparison between fibricheck and alivecor. Heart Rhythm. 2018;15(5):S71.

144. Engdahl J, Andersson L, Mirskaya M, Rosenqvist M. Stepwise screening of atrial fibrillation in a 75-year-old population: implications for stroke prevention. Circulation. 2013;127(8):930-7.

145. Gill S, Bunting KV, Sartini C, Cardoso VR, Ghoreishi N, Uh HW, et al. Smartphone detection of atrial fibrillation using photoplethysmography: a systematic review and meta-analysis. Heart. 2022;108(20):1600-7.

146. Rossello X, Stanbury M, Beeri R, Kirchhof P, Casadei B, Kotecha D. Digital learning and the future cardiologist. European heart journal. 2019;40(6):499-501.

147. Castaneda D, Esparza A, Ghamari M, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. Int J Biosens Bioelectron. 2018;4(4):195-202.

148. Brasier N, Raichle CJ, Dorr M, Becke A, Nohturfft V, Weber S, et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2019;21(1):41-7.

149. Chan PH, Wong CK, Poh YC, Pun L, Leung WW, Wong YF, et al. Diagnostic Performance of a Smartphone-Based Photoplethysmographic Application for Atrial Fibrillation Screening in a Primary Care Setting. J Am Heart Assoc. 2016;5(7).

150. Fan YY, Li YG, Li J, Cheng WK, Shan ZL, Wang YT, et al. Diagnostic Performance of a Smart Device With Photoplethysmography Technology for Atrial Fibrillation Detection: Pilot Study (Pre-mAFA II Registry). JMIR Mhealth Uhealth. 2019;7(3):e11437.

151. McManus DD, Lee J, Maitas O, Esa N, Pidikiti R, Carlucci A, et al. A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. Heart Rhythm. 2013;10(3):315-9.

152. Mc MD, Chong JW, Soni A, Saczynski JS, Esa N, Napolitano C, et al. PULSE-SMART: Pulse-Based Arrhythmia Discrimination Using a Novel Smartphone Application. J Cardiovasc Electrophysiol. 2016;27(1):51-7.

153. Mutke MR, Brasier N, Raichle C, Ravanelli F, Doerr M, Eckstein J. Comparison and Combination of Single-Lead ECG and Photoplethysmography Algorithms for Wearable-Based Atrial Fibrillation Screening. Telemed J E Health. 2021;27(3):296-302.

154. Poh MZ, Poh YC, Chan PH, Wong CK, Pun L, Leung WW, et al. Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms. Heart. 2018;104(23):1921-8.

155. Proesmans T, Mortelmans C, Van Haelst R, Verbrugge F, Vandervoort P, Vaes B. Mobile Phone-Based Use of the Photoplethysmography Technique to Detect Atrial Fibrillation in Primary Care: Diagnostic Accuracy Study of the FibriCheck App. JMIR Mhealth Uhealth. 2019;7(3):e12284.

156. Rozen G, Vaid J, Hosseini SM, Kaadan MI, Rafael A, Roka A, et al. Diagnostic Accuracy of a Novel Mobile Phone Application for the Detection and Monitoring of Atrial Fibrillation. Am J Cardiol. 2018;121(10):1187-91.

157. Yan BP, Lai WHS, Chan CKY, Chan SC, Chan LH, Lam KM, et al. Contact-Free Screening of Atrial Fibrillation by a Smartphone Using Facial Pulsatile Photoplethysmographic Signals. J Am Heart Assoc. 2018;7(8).

158. Grieten L, Van Der Auwera J, Vandervoort P, Dujardin K. Using smartphone enabled technologies for detection atrial fibrillation: Is there a difference in signal quality between ECG and PPG? Heart Rhythm. 2018;15(5):S70-S1.

159. Karim N, Aral M, Eduawor S, Finlay M. AF detection using smartphone apps. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2017;19:iii234.

160. Kuan DJ, Chung HJ, Lai WH, Chan CKY, To OTL, Poh YC, et al. A novel strategy combining hand-held ECG and contact-free facial pulsatile photoplethysmographic detection for atrial fibrillation screening. Heart Rhythm. 2018;15(5):S88-S9.

161. Maitas O, Lee J, Robotis D, Bourell P, Chon KH, McManus DD. Detection of atrial fibrillation using an iphone 4S. Circulation2012.

162. Maitas O, Lee J, Robotis D, Chon KH, McManus DD. Detection of atrial fibrillation using a smartphone camera. Circ Res2012.

163. Mortelmans C, Van Haelst R, Van Der Auwera J, Grieten L, Vandervoort P, Vaes B. Validation of a new smartphone application for the diagnosis of atrial fibrillation in primary care. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2017;19:iii16.

164. Mutke MR, Hoelz B, Brasier N, Ravanelli F, Doerr M, Eckstein J. Non-invasive rhythm diagnostics: bringing old and new together. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology2019. p. 779-87.

165. Napolitano CA, Chong JW, Saczynski JS, Esa N, Darling CE, Boyer E, et al. Pulse-based arrhythmia discrimination using a novel smartphone application. Circulation. 2015;132.

166. Proesmans T, Smeets C, Dreesen P, Vanhaeren J, Vandervoort P. Preliminary results of the FLASH-AF: Validation of the device independent nature of a pulse deriving smartphone application for the detection of atrial fibrillation. Eur Heart J2019. p. 765.

167. Proesmans T, Vandenberk T, Vandervoort P, Anne W, Dujardin K. Evaluation of screening technologies and assessments in a voluntary screening programme in the general belgian population. Eur Heart J2018. p. 1290-1.

168. Rozen G, Vaid J, Hosseini SM, Rafael A, Roka A, Kaadan I, et al. Diagnostic accuracy of a novel mobile phone application in monitoring atrial fibrillation. Eur Heart J. 2017;38(suppl_1):451-2.

169. Siu D, Wong CK, Chan PH, Poh YK, Poh MZ. Head-to-head comparison of a camera-based smartphone application cardiio rhythmtm with alivecor® heart monitor for atrial fibrillation screening in primary healthcare setting. J Am Coll Cardiol2016. p. 696.

170. Smeets C, Houben E, Vranken J, Proesmans T, Vandervoort P. The diagnostic accuracy of a pulse-deriving smartphone application is device independent. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology2019. p. 1024-9.

171. Vaid J, Poh MZ, Saleh A, Kalantarian S, Poh YKC, Rafael A, et al. Diagnostic accuracy of a novel mobile application (cardiio rhythm) for detecting atrial fibrillation. J Am Coll Cardiol2015. p. A361.

172. Vandenberk T, Mortelmans C, Van Haelst R, Van Schelvergem G, Pelckmans C, Storms V, et al. Diagnosis of atrial fibrillation based on a PPG signal compared to one lead ecg. Heart Rhythm. 2018;15(5):S505.

173. Vandenberk T, Pelckmans C, Mortelmans C, Van Haelst R, Storms V, Thijs IM, et al. Atrial fibrillation diagnosis based on a smartphone derived PPG waveform: A diagnostic accuracy study versus singlelead. Circulation. 2018;138.

174. Yan BP, Chan S, Lai WH, Lau BH, Lam EK, Yip AK, et al. Validation of a novel contact-free atrial fibrillation screening method using an iPhone camera to detect facial pulsatile photoplethysmographic signals. Circulation. 2016;134.

175. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.

176. Dwamena B. MIDAS: A program for Meta-analytical Integration of Diagnostic Accuracy Studies in Stata Division of Nuclear Medicine, Department of Radiology, University of Michigan Medical School, Ann Arbor, Michigan2007.

177. Kotecha D, Breithardt G, Camm AJ, Lip GYH, Schotten U, Ahlsson A, et al. Integrating new approaches to atrial fibrillation management: the 6th AFNET/EHRA Consensus Conference. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2018;20(3):395-407.

178. Jones NR, Taylor CJ, Hobbs FDR, Bowman L, Casadei B. Screening for atrial fibrillation: a call for evidence. Eur Heart J. 2020;41(10):1075-85.

179. Friberg L, Rosenqvist M, Lindgren A, Terent A, Norrving B, Asplund K. High prevalence of atrial fibrillation among patients with ischemic stroke. Stroke. 2014;45(9):2599-605.

180. Lane DA, Skjoth F, Lip GYH, Larsen TB, Kotecha D. Temporal Trends in Incidence, Prevalence, and Mortality of Atrial Fibrillation in Primary Care. J Am Heart Assoc. 2017;6(5).

181. Kelli HM, Witbrodt B, Shah A. The Future of Mobile Health Applications and Devices in Cardiovascular Health. Euro Med J Innov. 2017;2017:92-7.

182. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, et al. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. The New England journal of medicine. 2019;381(20):1909-17.

183. Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y, et al. Mobile Photoplethysmographic Technology to Detect Atrial Fibrillation. J Am Coll Cardiol. 2019;74(19):2365-75.

184. Lubitz SA, Faranesh AZ, Selvaggi C, Atlas SJ, McManus DD, Singer DE, et al. Detection of Atrial Fibrillation in a Large Population Using Wearable Devices: The Fitbit Heart Study. Circulation. 2022;146(19):1415-24.

185. Chiang CE, Naditch-Brule L, Murin J, Goethals M, Inoue H, O'Neill J, et al. Distribution and risk profile of paroxysmal, persistent, and permanent atrial fibrillation in routine clinical practice: insight from the real-life global survey evaluating patients with atrial fibrillation international registry. Circ Arrhythm Electrophysiol. 2012;5(4):632-9.

186. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J. 2016;37(38):2893-962.

187. Sethi NJ, Feinberg J, Nielsen EE, Safi S, Gluud C, Jakobsen JC. The effects of rhythm control strategies versus rate control strategies for atrial fibrillation and atrial flutter: A systematic review with meta-analysis and Trial Sequential Analysis. PLoS One. 2017;12(10):e0186856.

188. Kirchhof P, Ammentorp B, Darius H, De Caterina R, Le Heuzey JY, Schilling RJ, et al. Management of atrial fibrillation in seven European countries after the publication of the 2010 ESC Guidelines on atrial fibrillation: primary results of the PREvention oF thromboemolic events--European Registry in Atrial Fibrillation (PREFER in AF). Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2014;16(1):6-14.

189. Ulimoen SR, Enger S, Carlson J, Platonov PG, Pripp AH, Abdelnoor M, et al. Comparison of four single-drug regimens on ventricular rate and arrhythmia-related symptoms in patients with permanent atrial fibrillation. Am J Cardiol. 2013;111(2):225-30.

190. Clemo HF, Wood MA, Gilligan DM, Ellenbogen KA. Intravenous amiodarone for acute heart rate control in the critically ill patient with atrial tachyarrhythmias. Am J Cardiol. 1998;81(5):594-8.

191. Kotecha D, Calvert M, Deeks JJ, Griffith M, Kirchhof P, Lip GY, et al. A review of rate control in atrial fibrillation, and the rationale and protocol for the RATE-AF trial. BMJ open. 2017;7(7):e015099.

192. Hallberg P, Lindbäck J, Lindahl B, Stenestrand U, Melhus H. Digoxin and mortality in atrial fibrillation: a prospective cohort study. Eur J Clin Pharmacol. 2007;63(10):959-71.

193. Turakhia MP, Santangeli P, Winkelmayer WC, Xu X, Ullal AJ, Than CT, et al. Increased mortality associated with digoxin in contemporary patients with atrial fibrillation: findings from the TREAT-AF study. J Am Coll Cardiol. 2014;64(7):660-8.

194. Whitbeck MG, Charnigo RJ, Khairy P, Ziada K, Bailey AL, Zegarra MM, et al. Increased mortality among patients taking digoxin--analysis from the AFFIRM study. Eur Heart J. 2013;34(20):1481-8.

195. Aguirre Dávila L, Weber K, Bavendiek U, Bauersachs J, Wittes J, Yusuf S, et al. Digoxin-mortality: randomized vs. observational comparison in the DIG trial. Eur Heart J. 2019;40(40):3336-41.

196. Kotecha D, Bunting KV, Gill SK, Mehta S, Stanbury M, Jones JC, et al. Effect of Digoxin vs Bisoprolol for Heart Rate Control in Atrial Fibrillation on Patient-Reported Quality of Life: The RATE-AF Randomized Clinical Trial. Jama. 2020;324(24):2497-508.

197. Bunting KV, Stanbury M, Tica O, Kotecha D. Transforming clinical research by involving and empowering patients- the RATE-AF randomized trial. European heart journal. 2021;42(25):2411-4.

198. Pevnick JM, Birkeland K, Zimmer R, Elad Y, Kedan I. Wearable technology for cardiology: An update and framework for the future. Trends in cardiovascular medicine. 2018;28(2):144-50.

199. Redfield MM, Anstrom KJ, Levine JA, Koepp GA, Borlaug BA, Chen HH, et al. Isosorbide Mononitrate in Heart Failure with Preserved Ejection Fraction. The New England journal of medicine. 2015;373(24):2314-24.

200. Hemingway H, Asselbergs FW, Danesh J, Dobson R, Maniadakis N, Maggioni A, et al. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. Eur Heart J. 2017.

201. Moore K, O'Shea E, Kenny L, Barton J, Tedesco S, Sica M, et al. Older Adults' Experiences With Using Wearable Devices: Qualitative Systematic Review and Meta-synthesis. JMIR Mhealth Uhealth. 2021;9(6):e23832.

202. Fitbit Research Library. Fitabase. 2022. [2022-09-23]. https://www.fitabase.com/research-library/ [

203. Balbim GM, Marques IG, Marquez DX, Patel D, Sharp LK, Kitsiou S, et al. Using Fitbit as an mHealth Intervention Tool to Promote Physical Activity: Potential Challenges and Solutions. JMIR Mhealth Uhealth. 2021;9(3):e25289.

204. Benedetto S, Caldato C, Bazzan E, Greenwood DC, Pensabene V, Actis P. Assessment of the Fitbit Charge 2 for monitoring heart rate. PLoS One. 2018;13(2):e0192691.

205. Patridge EF, Bardyn TP. Research Electronic Data Capture (REDCap). J Med Libr Assoc. 2018;106(1):142-4.

206. Izmailova ES, Wagner JA, Perakslis ED. Wearable Devices in Clinical Trials: Hype and Hypothesis. Clinical pharmacology and therapeutics. 2018;104(1):42-52.

207. Bakobaki JM, Rauchenberger M, Joffe N, McCormack S, Stenning S, Meredith S. The potential for central monitoring techniques to replace on-site monitoring: findings from an international multi-centre clinical trial. Clin Trials. 2012;9(2):257-64.

208. Huhn S, Axt M, Gunga HC, Maggioni MA, Munga S, Obor D, et al. The Impact of Wearable Technologies in Health Research: Scoping Review. JMIR Mhealth Uhealth. 2022;10(1):e34384.

209. Schnabel RB, Marinelli EA, Arbelo E, Boriani G, Boveda S, Buckley CM, et al. Early diagnosis and better rhythm management to improve outcomes in patients with atrial fibrillation: the 8th AFNET/EHRA consensus conference. Europace. 2022.

210. Steg PG, Alam S, Chiang CE, Gamra H, Goethals M, Inoue H, et al. Symptoms, functional status and quality of life in patients with controlled and uncontrolled atrial fibrillation: data from the RealiseAF cross-sectional international registry. Heart. 2012;98(3):195-201.

211. Nabauer M, Gerth A, Limbourg T, Schneider S, Oeff M, Kirchhof P, et al. The Registry of the German Competence NETwork on Atrial Fibrillation: patient characteristics and initial management. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2009;11(4):423-34.

212. Lip GY, Laroche C, Dan GA, Santini M, Kalarus Z, Rasmussen LH, et al. A prospective survey in European Society of Cardiology member countries of atrial fibrillation management: baseline results of EURObservational Research Programme Atrial Fibrillation (EORP-AF) Pilot General Registry. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2014;16(3):308-19.

213. Ziff OJ, Kotecha D. Digoxin: The good and the bad. Trends Cardiovasc Med. 2016;26(7):585-95.

214. Fabritz L, Connolly DL, Czarnecki E, Dudek D, Guasch E, Haase D, et al. Smartphone and wearable detected atrial arrhythmias in Older Adults: Results of a fully digital European Case finding study. Eur Heart J Digit Health. 2022;3(4):610-25.

215. DeVore AD, Wosik J, Hernandez AF. The Future of Wearables in Heart Failure Patients. JACC Heart Fail. 2019;7(11):922-32.

216. Wang L, Nielsen K, Goldberg J, Brown JR, Rumsfeld JS, Steinberg BA, et al. Association of Wearable Device Use With Pulse Rate and Health Care Use in Adults With Atrial Fibrillation. JAMA Netw Open. 2021;4(5):e215821.

217. Sana F, Isselbacher EM, Singh JP, Heist EK, Pathik B, Armoundas AA. Wearable Devices for Ambulatory Cardiac Monitoring: JACC State-of-the-Art Review. Journal of the American College of Cardiology. 2020;75(13):1582-92.

218. Yao X, Rushlow DR, Inselman JW, McCoy RG, Thacher TD, Behnken EM, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. Nat Med. 2021;27(5):815-9.

219. Karwath A, Bunting KV, Gill SK, Tica O, Pendleton S, Aziz F, et al. Redefining β-blocker response in heart failure patients with sinus rhythm and atrial fibrillation: a machine learning cluster analysis. Lancet. 2021.

220. Nunes BP, Flores TR, Mielke GI, Thumé E, Facchini LA. Multimorbidity and mortality in older adults: A systematic review and meta-analysis. Arch Gerontol Geriatr. 2016;67:130-8.

221. Kuan V, Denaxas S, Patalay P, Nitsch D, Mathur R, Gonzalez-Izquierdo A, et al. Identifying and visualising multimorbidity and comorbidity patterns in patients in the English National Health Service: a population-based study. Lancet Digit Health. 2023;5(1):e16-e27.

222. Siah KW, Wong CH, Gupta J, Lo AW. Multimorbidity and mortality: A data science perspective. J Multimorb Comorb. 2022;12:26335565221105431.

223. Majnarić LT, Babič F, O'Sullivan S, Holzinger A. AI and Big Data in Healthcare: Towards a More Comprehensive Research Framework for Multimorbidity. J Clin Med. 2021;10(4).

224. Cesario A, D'Oria M, Calvani R, Picca A, Pietragalla A, Lorusso D, et al. The Role of Artificial Intelligence in Managing Multimorbidity and Cancer. J Pers Med. 2021;11(4).

225. Schefold JC, Filippatos G, Hasenfuss G, Anker SD, von Haehling S. Heart failure and kidney dysfunction: epidemiology, mechanisms and management. Nature Reviews Nephrology. 2016;12:610.

226. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J. 2016;37(27):2129-200.

227. Metra M, Cotter G, Gheorghiade M, Dei Cas L, Voors AA. The role of the kidney in heart failure. European heart journal. 2012;33(17):2135-42.

228. Lofman I, Szummer K, Hagerman I, Dahlstrom U, Lund LH, Jernberg T. Prevalence and prognostic impact of kidney disease on heart failure patients. Open heart. 2016;3(1):e000324.

229. Hillege HL, Nitsch D, Pfeffer MA, Swedberg K, McMurray JJ, Yusuf S, et al. Renal function as a predictor of outcome in a broad spectrum of patients with heart failure. Circulation. 2006;113(5):671-8.

230. Wali RK, Iyengar M, Beck GJ, Chartyan DM, Chonchol M, Lukas MA, et al. Efficacy and safety of carvedilol in treatment of heart failure with chronic kidney disease: a meta-analysis of randomized trials. Circ Heart Fail. 2011;4(1):18-26.

231. Ghali JK, Wikstrand J, Van Veldhuisen DJ, Fagerberg B, Goldstein S, Hjalmarson A, et al. The influence of renal function on clinical outcome and response to beta-blockade in

systolic heart failure: insights from Metoprolol CR/XL Randomized Intervention Trial in Chronic HF (MERIT-HF). J Card Fail. 2009;15(4):310-8.

232. Cohen-Solal A, Kotecha D, van Veldhuisen DJ, Babalis D, Bohm M, Coats AJ, et al. Efficacy and safety of nebivolol in elderly heart failure patients with impaired renal function: insights from the SENIORS trial. Eur J Heart Fail. 2009;11(9):872-80.

233. Castagno D, Jhund PS, McMurray JJ, Lewsey JD, Erdmann E, Zannad F, et al. Improved survival with bisoprolol in patients with heart failure and renal impairment: an analysis of the cardiac insufficiency bisoprolol study II (CIBIS-II) trial. Eur J Heart Fail. 2010;12(6):607-16.

234. Kotecha D, Manzano L, Altman DG, Krum H, Erdem G, Williams N, et al. Individual patient data meta-analysis of beta-blockers in heart failure: rationale and design. Systematic reviews. 2013;2(1):7.

235. Kotecha D, Manzano L, Krum H, Rosano G, Holmes J, Altman DG, et al. Effect of age and sex on efficacy and tolerability of beta blockers in patients with heart failure with reduced ejection fraction: individual patient data meta-analysis. BMJ. 2016;353:i1855.

236. Cleland JGF, Bunting KV, Flather MD, Altman DG, Holmes J, Coats AJS, et al. Beta-blockers for heart failure with reduced, mid-range, and preserved ejection fraction: an individual patient-level analysis of double-blind randomized trials. European heart journal. 2018;39(1):26-35.

237. Kotecha D, Manzano L, Altman DG, Krum H, Erdem G, Williams N, et al. Individual patient data meta-analysis of beta-blockers in heart failure: rationale and design. Systematic reviews. 2013;2:7.

238. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. JAMA. 2015;313(16):1657-65.

239. Australia/New Zealand Heart Failure Research Collaborative Group. Randomised, placebo-controlled trial of carvedilol in patients with congestive heart failure due to ischaemic heart disease. Lancet. 1997;349(9049):375-80.

240. Beta-Blocker Evaluation of Survival Trial Investigators. A trial of the beta-blocker bucindolol in patients with advanced chronic heart failure. The New England journal of medicine. 2001;344(22):1659-67.

241. Dargie HJ. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. Lancet. 2001;357(9266):1385-90.

242. Cleland JG, Pennell DJ, Ray SG, Coats AJ, Macfarlane PW, Murray GD, et al. Myocardial viability as a determinant of the ejection fraction response to carvedilol in patients with heart failure (CHRISTMAS trial): randomised controlled trial. Lancet. 2003;362(9377):14-21.

243. CIBIS Investigators and Committees. A randomized trial of beta-blockade in heart failure. The Cardiac Insufficiency Bisoprolol Study (CIBIS). Circulation. 1994;90(4):1765-73.

244. The Cardiac Insufficiency Bisoprolol Study II (CIBIS-II): a randomised trial. Lancet. 1999;353(9146):9-13.

245. Packer M, Coats AJ, Fowler MB, Katus HA, Krum H, Mohacsi P, et al. Effect of carvedilol on survival in severe chronic heart failure. The New England journal of medicine. 2001;344(22):1651-8.

246. Waagstein F, Bristow MR, Swedberg K, Camerini F, Fowler MB, Silver MA, et al. Beneficial effects of metoprolol in idiopathic dilated cardiomyopathy. Metoprolol in Dilated Cardiomyopathy (MDC) Trial Study Group. Lancet. 1993;342(8885):1441-6.

247. Flather MD, Shibata MC, Coats AJ, Van Veldhuisen DJ, Parkhomenko A, Borbola J, et al. Randomized trial to determine the effect of nebivolol on mortality and cardiovascular hospital admission in elderly patients with heart failure (SENIORS). Eur Heart J. 2005;26(3):215-25.

248. Han J, Kamber M, Pei J. 10 - Cluster Analysis: Basic Concepts and Methods. In: Han J, Kamber M, Pei J, editors. Data Mining (Third Edition). Boston: Morgan Kaufmann; 2012. p. 443-95.

249. Yu H, Chapman B, Di Florio A, Eischen E, Gotz D, Jacob M, et al. Bootstrapping estimates of stability for clusters, observations and model selection. Computational Statistics. 2019;34(1):349-72.

250. Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y, et al. Mobile Health Technology for Atrial Fibrillation Screening Using Photoplethysmography-Based Smart Devices: The HUAWEI Heart study. Journal of the American College of Cardiology. 2019.

251. Lee L, Maher ML. Factors Affecting the Initial Engagement of Older Adults in the Use of Interactive Technology. Int J Environ Res Public Health. 2021;18(6).

252. Seshadri DR, Bittel B, Browsky D, Houghtaling P, Drummond CK, Desai MY, et al. Accuracy of Apple Watch for Detection of Atrial Fibrillation. Circulation. 2020;141(8):702-3.

253. Ford C, Xie CX, Low A, Rajakariar K, Koshy AN, Sajeev JK, et al. Comparison of 2 Smart Watch Algorithms for Detection of Atrial Fibrillation and the Benefit of Clinician Interpretation: SMART WARS Study. JACC Clin Electrophysiol. 2022;8(6):782-91.

254. Hughes A, Shandhi MMH, Master H, Dunn J, Brittain E. Wearable Devices in Cardiovascular Medicine. Circ Res. 2023;132(5):652-70.

Appendices

Appendix 1: Table A1: Search strategy for EMBASE and MEDLINE databases

Search terms included (total hits 781)
Mobile applications/mobile phone/ cellular phone/ smartphone
('mobile application'/exp OR 'mobile app':ti,ab,de,tn,dn,kw OR 'mobile application':ti,ab,de,tn,dn,kw OR 'mobile applications':ti,ab,de,tn,dn,kw OR 'mobile apps':ti,ab,de,tn,dn,kw OR 'portable software app':ti,ab,de,tn,dn,kw OR 'portable software application':ti,ab,de,tn,dn,kw OR 'portable software applications':ti,ab,de,tn,dn,kw OR 'portable software apps':ti,ab,de,tn,dn,kw OR 'cell phone use'/exp OR 'cell phone usage':ti,ab,de,tn,dn,kw OR 'cell phone use':ti,ab,de,tn,dn,kw OR 'cell phone utilization':ti,ab,de,tn,dn,kw OR 'cellphone usage':ti,ab,de,tn,dn,kw OR 'cellphone use':ti,ab,de,tn,dn,kw OR 'cellphone utilization':ti,ab,de,tn,dn,kw OR 'mobile phone usage':ti,ab,de,tn,dn,kw OR 'mobile phone use':ti,ab,de,tn,dn,kw OR 'mobile phone utilization':ti,ab,de,tn,dn,kw OR 'mobile phone'/exp OR 'cell phone':ti,ab,de,tn,dn,kw OR 'cell phones':ti,ab,de,tn,dn,kw OR 'cellphone':ti,ab,de,tn,dn,kw OR 'cellphones':ti,ab,de,tn,dn,kw OR 'cellular phone':ti,ab,de,tn,dn,kw OR 'cellular telephone':ti,ab,de,tn,dn,kw OR 'mobile phone':ti,ab,de,tn,dn,kw OR 'mobile telephone':ti,ab,de,tn,dn,kw OR 'smartphone'/exp OR 'smart phone*':ti,ab,de,tn,dn,kw OR smartphone*:ti,ab,de,tn,dn,kw OR 'mobile app*':ti,ab,de,tn,dn,kw OR mobileapp*:ti,ab,de,tn,dn,kw OR 'portable software app*':ti,ab,de,tn,dn,kw OR 'portablesoftware app*':ti,ab,de,tn,dn,kw OR ((mobile NEAR/2 app*):ti,ab,de,tn,dn,kw) OR (('cell phone*' NEAR/2 app*):ti,ab,de,tn,dn,kw) OR (('cell phone*' NEAR/2 software*):ti,ab,de,tn,dn,kw) OR (('cellphone*' NEAR/2 app*):ti,ab,de,tn,dn,kw) OR (('cellphone*' NEAR/2 software*):ti,ab,de,tn,dn,kw) OR (('cellular phone*' NEAR/2 app*):ti,ab,de,tn,dn,kw) OR (('cellular phone*' NEAR/2 software*):ti,ab,de,tn,dn,kw) OR (('cellularphone*' NEAR/2 app*):ti,ab,de,tn,dn,kw) OR (('cellularphone*' NEAR/2 software*):ti,ab,de,tn,dn,kw) OR (('cellular telephone*' NEAR/2 app*):ti,ab,de,tn,dn,kw) OR (('cellular telephone*' NEAR/2 software*):ti,ab,de,tn,dn,kw) OR (('cellulartelephone*' NEAR/2 app*):ti,ab,de,tn,dn,kw) OR (('cellulartelephone*' NEAR/2 software*):ti,ab,de,tn,dn,kw) OR ((mobile NEAR/2 (app* OR software*)):ti,ab,de,tn,dn,kw) OR 'cell phone*':ti,ab,tn,dn,de,kw OR 'cellphone*':ti,ab,tn,dn,de,kw OR 'cellular phone*':ti,ab,tn,dn,de,kw OR 'cellular telephone*':ti,ab,tn,dn,de,kw OR 'cellularphone*':ti,ab,tn,dn,de,kw OR 'cellulartelephone*':ti,ab,tn,dn,de,kw OR 'mobile phone*':ti,ab,tn,dn,de,kw OR 'mobile telephone*':ti,ab,tn,dn,de,kw OR 'mobilephone*':ti,ab,tn,dn,de,kw OR 'mobiletelephone*':ti,ab,tn,dn,de,kw OR 'smart phone*':ti,ab,tn,dn,de,kw OR 'smartphone*':ti,ab,tn,dn,de,kw
Tachycardia
AND ('tachycardia'/exp OR 'catecholaminergic polymorphic ventricular tachycardia'/exp OR 'experimental tachycardia'/exp OR 'isoprenaline-induced tachycardia'/exp OR 'heart ventricle tachycardia'/exp OR 'hyperkinetic heart syndrome'/exp OR 'monomorphic ventricular tachycardia'/exp OR 'pacemaker mediated tachycardia'/exp OR 'paroxysmal tachycardia'/exp OR 'paroxysmal supraventricular tachycardia'/exp OR 'wolff parkinson white syndrome'/exp OR 'polymorphic ventricular tachycardia'/exp OR 'postural orthostatic tachycardia syndrome'/exp OR 'reentry tachycardia'/exp OR 'atrioventricular nodal reentry tachycardia'/exp OR 'sinoatrial nodal reentry tachycardia'/exp OR 'sinus tachycardia'/exp OR 'supraventricular tachycardia'/exp OR 'ectopic atrial tachycardia'/exp OR 'junctional ectopic tachycardia'/exp OR 'tachycardia induced cardiomyopathy'/exp OR 'acute atrial fibrillation':ti,ab,tn,dn,kw,de
Atrial fibrillation/atrial flutter

OR 'acute heart atrium fibrillation':ti,ab,tn,dn,kw,de OR 'atrial fibrillation':ti,ab,tn,dn,kw,de OR 'atrium fibrillation':ti,ab,tn,dn,kw,de OR 'auricular fibrilation':ti,ab,tn,dn,kw,de OR 'auricular fibrillation':ti,ab,tn,dn,kw,de OR 'cardiac atrial fibrillation':ti,ab,tn,dn,kw,de OR 'cardiac atrium fibrillation':ti,ab,tn,dn,kw,de OR 'chronic atrial fibrillation':ti,ab,tn,dn,kw,de OR 'chronic atrium fibrillation':ti,ab,tn,dn,kw,de OR 'experimental atrial fibrillation':ti,ab,tn,dn,kw,de OR 'experimentally induced atrial fibrillation':ti,ab,tn,dn,kw,de OR 'fibrillation, heart atrium':ti,ab,tn,dn,kw,de OR 'heart atrial fibrillation':ti,ab,tn,dn,kw,de OR 'heart atrium fibrillation':ti,ab,tn,dn,kw,de OR 'heart fibrillation atrium':ti,ab,tn,dn,kw,de OR 'new-onset atrial fibrillation':ti,ab,tn,dn,kw,de OR 'nonvalvular atrial fibrillation':ti,ab,tn,dn,kw,de OR 'non-valvular atrial fibrillation':ti,ab,tn,dn,kw,de OR 'paroxysmal atrial fibrillation':ti,ab,tn,dn,kw,de OR 'paroxysmal heart atrium fibrillation':ti,ab,tn,dn,kw,de OR 'permanent atrial fibrillation':ti,ab,tn,dn,kw,de OR 'permanent atrium fibrillation':ti,ab,tn,dn,kw,de OR 'persistent atrial fibrillation':ti,ab,tn,dn,kw,de OR 'persistent atrium fibrillation':ti,ab,tn,dn,kw,de OR 'persistent heart atrium fibrillation':ti,ab,tn,dn,kw,de OR 'recent-onset atrial fibrillation':ti,ab,tn,dn,kw,de OR 'atrial fibrillation'/exp OR 'chronic atrial fibrillation'/exp OR 'experimental atrial fibrillation'/exp OR 'new-onset atrial fibrillation'/exp OR 'paroxysmal atrial fibrillation'/exp OR 'permanent atrial fibrillation'/exp OR 'persistent atrial fibrillation'/exp OR (((atrial* OR atrium* OR auricular*) NEXT/2 fibrillat*):ti,ab,tn,dn,kw,de) OR 'heart atrium flutter'/exp OR 'atrial flutter':ti,ab,dn,tn,kw,de OR 'atrium flutter':ti,ab,dn,tn,kw,de OR 'atrium flutter, heart':ti,ab,dn,tn,kw,de OR 'auricular flutter':ti,ab,dn,tn,kw,de OR 'cardiac atrial flutter':ti,ab,dn,tn,kw,de OR 'cardiac atrium flutter':ti,ab,dn,tn,kw,de OR 'flutter, heart atrium':ti,ab,dn,tn,kw,de OR 'heart atrial flutter':ti,ab,dn,tn,kw,de OR 'heart atrium flutter':ti,ab,dn,tn,kw,de OR 'supraventricular flutter':ti,ab,dn,tn,kw,de OR (((atrial* OR atrium* OR auricular*) NEXT/2 flutter*):ti,ab,tn,dn,kw,de)
Ventricular fibrillation and flutter/ heart beat/extra beats/ectopics
OR 'heart ventricle flutter'/exp OR 'flutter, heart ventricle':ti,ab,dn,tn,kw,de OR 'heart ventricle flutter':ti,ab,dn,tn,kw,de OR 'ventricular flutter':ti,ab,dn,tn,kw,de OR ((ventri* NEXT/2 flutter*):ti,ab,tn,dn,kw,de) OR 'heart ventricle fibrillation'/exp OR 'cardiac ventricle fibrillation':ti,ab,tn,dn,kw,de OR 'cardiac ventricular fibrillation':ti,ab,tn,dn,kw,de OR 'fibrillation, heart ventricle':ti,ab,tn,dn,kw,de OR 'heart ventricle fibrillation':ti,ab,tn,dn,kw,de OR 'heart ventricular fibrillation':ti,ab,tn,dn,kw,de OR 'ventricle fibrillation':ti,ab,tn,dn,kw,de OR 'ventricle fibrillation, heart':ti,ab,tn,dn,kw,de OR 'ventricular fibrillation':ti,ab,tn,dn,kw,de OR 'experimental ventricular fibrillation'/exp OR 'experimental ventricular fibrillation':ti,ab,tn,dn,kw,de OR 'experimentally induced ventricular vibrillation':ti,ab,tn,dn,kw,de OR 'electrically induced ventricular fibrillation'/exp OR 'electrically induced ventricular fibrillation':ti,ab,tn,dn,kw,de OR ((ventri* NEXT/2 fibrill*):ti,ab,tn,dn,kw,de) OR 'heart beat'/exp OR 'beat, heart':ti,ab,tn,dn,kw,de OR 'cardiac beat':ti,ab,tn,dn,kw,de OR 'heart beat':ti,ab,tn,dn,kw,de OR 'diastole'/exp OR 'cardiac diastole':ti,ab,tn,dn,kw,de OR 'diastole':ti,ab,tn,dn,kw,de OR 'diastolic force':ti,ab,tn,dn,kw,de OR 'heart diastole':ti,ab,tn,dn,kw,de OR 'extrasystole'/exp OR 'beat, ectopic':ti,ab,tn,dn,kw,de OR 'cardiac complexes, premature':ti,ab,tn,dn,kw,de OR 'ectopic beat':ti,ab,tn,dn,kw,de OR 'extrasystole':ti,ab,tn,dn,kw,de OR 'extrasystolic beat':ti,ab,tn,dn,kw,de OR 'heart extrasystole':ti,ab,tn,dn,kw,de OR 'heart premature beat':ti,ab,tn,dn,kw,de OR 'nodal extrasystole':ti,ab,tn,dn,kw,de OR 'premature beat':ti,ab,tn,dn,kw,de OR 'systole'/exp OR 'heart systole':ti,ab,tn,dn,kw,de OR 'systole':ti,ab,tn,dn,kw,de OR 'systolic index':ti,ab,tn,dn,kw,de OR 'systolic phase':ti,ab,tn,dn,kw,de OR ((heart NEXT/3 beat):ti,ab,tn,dn,kw,de) OR ((irregular* NEAR/3 (puls* OR heart* OR rhythm*)):ti,ab,tn,dn,kw,de) OR (((atypical* OR unexpect* OR unsuspect* OR unforeseeabl* OR abnormal* OR uneven* OR fitful* OR erratic*OR sporadi*) NEAR/3 (puls* OR heart* OR rhythm*)):ti,ab,de,tn,dn,kw))

Appendix 2: Table A2: Search strategy for Cochrane database

ID	Search	Hits
#1	MeSH descriptor: [Atrial Fibrillation] explode all trees	4665
#2	(atrial* near/2 fibrillation*):ti,ab,kw	12671
#3	#1 OR #2	12671
#4	MeSH descriptor: [Atrial Flutter] explode all trees	375
#5	(atrial* near/2 flutter*):ti,ab,kw	978
#6	#4 OR #5	978
#7	MeSH descriptor: [Ventricular Fibrillation] explode all trees	544
#8	(ventri* near/2 fibrillat*):ti,ab,kw	1713
#9	#7 OR #8	1713
#10	MeSH descriptor: [Tachycardia] explode all trees	1799
#11	tachycardi*:ti,ab,kw	8496
#12	#10 OR #11	8525
#13	MeSH descriptor: [Heart Rate] explode all trees	19310
#14	(heart rate):ti,ab,kw	75355
#15	#13 OR #14	75360
#16	(Irregular* or unpredictable* or atypical*) near/3 (puls* or heart* or rhythm*)	244
#17	(atypical* or unexpect* or unsuspect* or unforeseeabl* or abnormal* or uneven* or	946

	fitful* or erratic* or sporadi*) near/3 (puls* or heart* or rhythm*)	
#18	#16 OR #17	1157
#19	((smart* near/2 phone*) or (mobil* near/2 phone*) or mobilphone* or cellphone* or (cell near/2 phone*) or mobilephone or mobile app*):ti,ab,kw	8221
#20	MeSH descriptor: [Mobile Applications] explode all trees	660
#21	MeSH descriptor: [Smartphone] explode all trees	407
#22	MeSH descriptor: [Cell Phone] explode all trees	1373
#23	#19 OR #20 OR #21 OR #22	8851
#24	MeSH descriptor: [Ventricular Flutter] explode all trees	2
#25	(ventri* near/2 flutter*):ti,ab,kw	28
#26	#24 OR #25	28
#27	#3 OR #6 OR #9 OR #12 OR #15 OR #18 OR #26	90651
#28	#23 AND #27 in Trials	370

Appendix 3: REDCap wearables sub-study baseline review survey

Rate AF wearables sub-study baseline survey

Please complete the survey below.

Thank you!

1)	Random participant ID	_____
2)	RATE AF ID number	_____
3)	Date of visit	_____
4)	Has patient given consent to be a part of sub-study	◯ Yes ◯ No
5)	Date of birth	_____
6)	Drug randomised to	◯ Beta blocker ◯ Digoxin
7)	Echo baseline	◯ HF ◯ No HF
8)	Have you been diagnosed with heart failure?	◯ Yes ◯ No
9)	E-mail address given	_____
10)	Provided mobile number	_____
11)	Personal landline phone number	_____
12)	Do you own a smart phone?	◯ Yes ◯ No
13)	If so do you use the following features of a smart phone (internet, apps)	◯ Yes ◯ No ◯ N/a
14)	Personal mobile phone number	_____
15)	ECG pdf	
16)	AF on ECG?	◯ Yes ◯ No
17)	Weight (kg)	_____

18)	Height (cm)	_____
19)	Hand dominance	◯ Right ◯ Left
20)	Wrist fitbit worn	◯ Right ◯ Left
21)	EHRA score baseline	_____
22)	NHYA baseline	_____
23)	NT-ProBNP	_____ (ng/L)
24)	Do you have access to WIFI at home?	◯ Yes ◯ No
25)	SF36Q1-General health over last 4 weeks?	◯ Excellent ◯ Very Good ◯ Good ◯ Fair ◯ Poor
26)	SF36Q2 - Compared to 1 year ago, rate your general health now?	◯ Much better now than one year ago ◯ Somewhat better now than one year ago ◯ About the same ◯ Somewhat worse now than one year ago ◯ Much worse than one year ago
27)	SF36Q3a- Does health limit vigorous activities (running, lifting heavy objects, strenuous sports)?	◯ Yes, limited a lot ◯ Yes, limited a little ◯ No, not limited at all
28)	SF36Q3b- Does health limit moderate activities (moving a table, pushing a vacuum cleaner, bowling, golf)?	◯ Yes, limited a lot ◯ Yes, limited a little ◯ No, not limited at all
29)	SF36Q3c- Does health limit lifting/carrying groceries ?	◯ Yes, limited a lot ◯ Yes, limited a little ◯ No, not limited at all
30)	SF36Q3d- Does health limit climbing several fights of stairs	◯ Yes, limited a lot ◯ Yes, limited a little ◯ No, not limited at all
31)	SF36Q3e- Does health limit climbing one flight of stairs	◯ Yes, limited a lot ◯ Yes, limited a little ◯ No, not limited at all
32)	SF36Q3f- Does health limit bending, kneeling or stooping ?	◯ Yes, limited a lot ◯ Yes, limited a little ◯ No, not limited at all

33) SF36Q3g- Does health limit walking more than 1 mile?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

34) SF36Q3h- Does health limit walking several hundred yards ?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

35) SF36Q3i- Does health limit walking one hundred yards?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

36) SF36Q3j- Does health limit bathing or dressing yourself?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

37) SF36Q4a- As a result of your physical health, how much have you cut down on the amount of time spent on work or other activities ?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

38) SF36Q4b- Accomplished less than you would like?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

39) SF36Q4c- Were limited in the kind of work or other activities?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

40) SF36Q4d- Had difficulty performing work or other activities (extra effort)?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

41) SF36Q5a -How much have you cut down on the amount of time spent on work or other activities as a result of emotional problems (depressed or anxious)?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

42) SF36Q5b - Accomplished less than you would like due to your emotional health?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

43) SF36Q5c- Did work or other activities less carefully than usual as a result of your emotional health?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

44) SF36Q6- During last four weeks what extent has physical health or emotional problems interfered with normal social activities?
- ◯ Not at all
- ◯ Slightly
- ◯ Moderately
- ◯ Quite a bit
- ◯ Extremely

45) SF36Q7- How much bodily pain have you had in the last four weeks?
- ◯ None
- ◯ Very mild
- ◯ Mild
- ◯ Moderate
- ◯ Severe
- ◯ Very severe

46) SF36Q8- How much did pain interfere with normal work (work outside home and housework)?
- ◯ Not at all
- ◯ A little bit
- ◯ Moderately
- ◯ Quite a bit
- ◯ Extremely

47) SF36Q9a- How much of the time did you feel full of life?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

48) SF36Q9b- Have you been very nervous ?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

49) SF36Q9c- Have you felt so down in the dumps, nothing could cheer you up?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

50) SF36Q9d- Have you felt calm and peaceful?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

51) SF36Q9e- Did you have a lot of energy?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

52) SF36Q9f- Have you felt downhearted and low?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

53) SF36Q9g- Did you feel worn out?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

54) SF369h- Have you been happy?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

55) SF36Q9i- Did you feel tired?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

56) SF36Q10 - In the past 4 weeks how much of the time has your physical or emotional health interfered with your social activities?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

57) SF36Q11a- Get ill more easily?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

58) SF36Q11b- Healthy as anybody I know ?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

59) SF36Q11c- Expect health to get worse?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

60) SF36Q11d- My health is excellent?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

61) EQ5DLQ1- Mobility
- ◯ I have no problems in walking about
- ◯ I have slight problems in walking about
- ◯ I have moderate problems in walking about
- ◯ I have severe problems in walking about
- ◯ I am unable to walk about

62) EQ5DLQ2- Self care
- ◯ I have no problems washing or dressing myself
- ◯ I have slight problems washing or dressing myself
- ◯ I have moderate problems washing or dressing myself
- ◯ I have severe problems washing or dressing myself
- ◯ I am unable to wash or dress myself

63) EQ5DLQ3 - Usual activities
- ◯ I have no problems doing my usual activities
- ◯ I have slight problems doing my usual activities
- ◯ I have moderate problems doing my usual activities
- ◯ I have severe problems doing my usual activities
- ◯ I am unable to do my usual activities

64) EQ5DLQ4- Pain/discomfort
- ◯ I have no pain or discomfort
- ◯ I have slight pain or discomfort
- ◯ I have moderate pain or discomfort
- ◯ I have severe pain or discomfort
- ◯ I have extreme pain or discomfort

65) EQ5DLQ5- Anxiety/depression
- ◯ I am not anxious or depressed
- ◯ I am slightly anxious or depressed
- ◯ I am moderately anxious or depressed
- ◯ I am severely anxious or depressed
- ◯ I am extremely anxious or depressed

66) EQ5DLQ6- Visual analogue health score today

0 50 100

(Place a mark on the scale above)

67) AFEQTQ1- Palpitations
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

68) AFEQTQ2- Irregular heartbeat
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

69) AFEQTQ3- Pause in heart activity
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

70) AFEQTQ4- Lightheadedness or dizziness
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

71) AFEQTQ5- Ability to have recreational pastimes, sports and hobbies
- ◯ Not at all limited
- ◯ Hardly limited
- ◯ A little limited
- ◯ Moderately limited
- ◯ Quite a bit limited
- ◯ Very limited
- ◯ Extremely limited

72) AFEQTQ6 - Ability to have a relationship and do things with friends and family
- ◯ Not at all limited
- ◯ Hardly limited
- ◯ A little limited
- ◯ Moderately limited
- ◯ Quite a bit limited
- ◯ Very limited
- ◯ Extremely limited

73) AFEQTQ7- Difficulty doing activity due to tiredness, fatigue or low on energy
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

74) AFEQTQ8- Doing physical activity due to SOB
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

75) AFEQTQ9- During exercise
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

76) AFEQTQ10- Walking briskly
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

77) AFEQTQ11- Walking briskly uphill or carrying groceries or other items, up a flight of stairs without stopping
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

78) AFEQTQ12- During vigorous activity such as lifting or moving heavy furniture, running, or participating in strenuous sports
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

79) AFEQTQ13- Anxious that AF can start anytime
- ◯ Not at all bothered
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

80) AFEQTQ14- Anxious AF may worsen other medical conditions in the long run
- ◯ Not at all bothered
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

81) AFEQTQ15- Worried about the treatment side effects from medications
- ◯ Not at all bothered
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

82) AFEQTQ16- Worried about complications from procedures (ablation, surgery, pacemaker therapy)
- ◯ Not at all bothered
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

83) AFEQTQ17- Worried about side effects from blood thinners (bleeding, bruises)
- ◯ Not at all bothered
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

84) AFEQTQ18-Worried or feeling anxious that treatment interferes with daily activities
- ◯ Not at all bothered
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

85) AFEQTQ19- How well does your current treatment control your atrial fibrillation?

- ◯ Extremely satisfied
- ◯ Very satisfied
- ◯ Somewhat satisfied
- ◯ Mixed with satisfied or dissatisfied
- ◯ Somewhat dissatisfied
- ◯ Very dissatisfied
- ◯ Extremely dissatisfied

86) AFEQTQ20- To what extent has the treatment relieved your symptoms of atrial fibrillation?

- ◯ Extremely satisfied
- ◯ Very satisfied
- ◯ Somewhat satisfied
- ◯ Mixed with satisfied or dissatisfied
- ◯ Somewhat dissatisfied
- ◯ Very dissatisfied
- ◯ Extremely dissatisfied

Appendix 4: REDCap wearables sub-study interim review survey

Rate AF wearables sub-study interim survey

Please complete the survey below.

Thank you!

1) Random participant ID _____

2) RATE AF ID number _____

3) Date of follow up _____

4) Do you have WIFI?
- ○ Yes
- ○ No

5) Have you connected to WIFI?
- ☐ Yes
- ☐ No
- ☐ N/a

6) Are you carrying the phone with you regularly in the last 7 days?
- ○ Yes
- ○ No

7) If yes where do you carry the phone? _____

8) If not why?
- ○ Forgot
- ○ Not aware of need
- ○ Too big to carry
- ○ Other

9) In the past 7 days are you recording your AF PPG measurements twice weekly?
- ○ Yes
- ○ No

10) If yes, are these done correctly ?
- ○ Yes
- ○ No
- ○ N/a

11) If not reason why not measured?
- ○ Forgot
- ○ Not able to do it correctly
- ○ App not working
- ○ Other

12) In the past 7 days are you recording your my AF quality of life score twice weekly?
- ○ Yes
- ○ No

13) If not why?
- ○ Forgot
- ○ Not able to do it
- ○ App not working
- ○ Other

14) In the past 7 days are you recording your my AF symptom score twice weekly?
- ○ Yes
- ○ No

15) If not why?
- ◯ Forgot
- ◯ Not able to do it
- ◯ App not working
- ◯ Other

16) Q1- In the past 7 days how many days did you wear your fitbit device for?
- ◯ 0 days
- ◯ 1 day
- ◯ 2 days
- ◯ 3 days
- ◯ 4 days
- ◯ 5 days
- ◯ 6 days
- ◯ Everyday

17) Q2 - On a daily basis, how many hours did you wear your fitbit device for (to the nearest hour)?

_____

18) Reason for not wearing fitbit device regularly
- ◯ Non-compliance
- ◯ Did not wish to continue in study
- ◯ Hospitalization
- ◯ Damage to fitbit or phone
- ◯ Other

19) If other please specify

_____

Appendix 5: REDCap wearables sub-study final review survey

Rate AF wearables sub-study final survey

Please complete the survey below.

Thank you!

1) Random participant ID _____

2) RATE AF ID number _____

3) Date of visit _____

4) ECG pdf

5) ECG today - AF
- ○ Yes
- ○ No

6) EHRA score _____

7) NYHA score _____

8) NT-ProBNP _____ (pg/mL)

9) Echo at final review
- ○ HF
- ○ No HF

10) Have you been diagnosed with heart failure
- ○ Yes
- ○ No

11) Are you carrying the phone with you regularly in the last 7 days?
- ○ Yes
- ○ No

12) If yes where do you carry the phone? _____

13) If not why?
- ○ Forgot
- ○ Not aware of need
- ○ Too big to carry
- ○ Other

14) Did you use for the smartphone for calls/texts/internet?
- ○ Yes
- ○ No

15) if no why?
- ○ Forgot
- ○ Didn't need to
- ○ Too difficult to use

16) Did you find the phone easy to use in this study ?
- ○ Yes
- ○ No

17) Are you recording AF PPG measurements twice weekly?
- ◯ Yes
- ◯ No

18) If yes, are these done correctly ?
- ◯ Yes
- ◯ No

19) Reason why not measured?
- ◯ Forgot
- ◯ Not able to do it correctly
- ◯ App not working
- ◯ Other

20) How did you find using the apps?
- ◯ Easy
- ◯ Intermediate
- ◯ Hard

21) Are you recording your my AF quality of life score twice weekly?
- ◯ Yes
- ◯ No

22) If not why?
- ◯ Forgot
- ◯ Not able to do it
- ◯ App not working
- ◯ Other

23) Are you recording your my AF symptom score twice weekly?
- ◯ Yes
- ◯ No

24) If not why?
- ◯ Forgot
- ◯ Not able to do it
- ◯ App not working
- ◯ Other

25) Did you find the use of the my AF app improved your understanding of AF?
- ◯ Yes
- ◯ No
- ◯ Don't know
- ◯ N/a did not use app

26) Did you find using the my AF app improved your compliance to AF medication?
- ◯ Yes
- ◯ No
- ◯ Don't know
- ◯ N/a did not use

27) Q1-In the past 7 days how many days did you wear your fitbit device for?
- ◯ 0 days
- ◯ 1 day
- ◯ 2 days
- ◯ 3 days
- ◯ 4 days
- ◯ 5 days
- ◯ 6 days
- ◯ Everyday

28) Q2 - On a daily basis, how many hours did you wear your fitbit device for (to the nearest hour)?

_____

29) If not worn, reason for not wearing fitbit device
- ◯ Non-compliance
- ◯ Did not wish to continue in study
- ◯ Hospitalization
- ◯ Damage to fitbit or phone
- ◯ Other

30) Did wearing the fitbit device improve your motivation in terms of physical activity?
- ◯ Yes
- ◯ No

31) Did wearing a fitbit device increase your physical activity?
- ◯ Yes
- ◯ No

32) SF36Q1-General health over last 4 weeks?
- ◯ Excellent
- ◯ Very Good
- ◯ Good
- ◯ Fair
- ◯ Poor

33) SF36Q2 - Compared to 1 year ago, rate your general health now?
- ◯ Much better now than one year ago
- ◯ Somewhat better now than one year ago
- ◯ About the same
- ◯ Somewhat worse now than one year ago
- ◯ Much worse than one year ago

34) SF36Q3a- Does health limit vigorous activities (running, lifting heavy objects, strenuous sports)?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

35) SF36Q3b- Does health limit moderate activities (moving a table, pushing a vacuum cleaner, bowling, golf)?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

36) SF36Q3c- Does health limit lifting/carrying groceries ?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

37) SF36Q3d- Does health limit climbing several fights of stairs
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

38) SF36Q3e- Does health limit climbing one flight of stairs
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

39) SF36Q3f- Does health limit bending, kneeling or stooping ?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

40) SF36Q3g- Does health limit walking more than 1 mile?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

41) SF36Q3h- Does health limit walking several hundred yards ?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

42) SF36Q3i- Does health limit walking one hundred yards?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

43) SF36Q3j- Does health limit bathing or dressing yourself?
- ◯ Yes, limited a lot
- ◯ Yes, limited a little
- ◯ No, not limited at all

44) SF36Q4a- As a result of your physical health, how much have you cut down on the amount of time spent on work or other activities ?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

45) SF36Q4b- Accomplished less than you would like?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

46) SF36Q4c- Were limited in the kind of work or other activities?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

47) SF36Q4d- Had difficulty performing work or other activities (extra effort)?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

48) SF36Q5a -How much have you cut down on the amount of time spent on work or other activities as a result of emotional problems (depressed or anxious)?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

49) SF36Q5b - Accomplished less than you would like due to your emotional health?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

50) SF36Q5c- Did work or other activities less carefully than usual as a result of your emotional health?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

51) SF36Q6- During last four weeks what extent has physical health or emotional problems interfered with normal social activities?
- ◯ Not at all
- ◯ Slightly
- ◯ Moderately
- ◯ Quite a bit
- ◯ Extremely

52) SF36Q7- How much bodily pain have you had in the last four weeks?
- ◯ None
- ◯ Very mild
- ◯ Mild
- ◯ Moderate
- ◯ Severe
- ◯ Very severe

53) SF36Q8- How much did pain interfere with normal work (work outside home and housework)?
- ◯ Not at all
- ◯ A little bit
- ◯ Moderately
- ◯ Quite a bit
- ◯ Extremely

54) SF36Q9a- How much of the time did you feel full of life?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

55) SF36Q9b- Have you been very nervous ?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

56) SF36Q9c- Have you felt so down in the dumps, nothing could cheer you up?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

57) SF36Q9d- Have you felt calm and peaceful?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

58) SF36Q9e- Did you have a lot of energy?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

59) SF36Q9f- Have you felt downhearted and low?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

60) SF36Q9g- Did you feel worn out?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

61) SF369h- Have you been happy?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

62) SF36Q9i- Did you feel tired?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

63) SF36Q10 - In the past 4 weeks how much of the time has your physical or emotional health interfered with your social activities?
- ◯ All of the time
- ◯ Most of the time
- ◯ Some of the time
- ◯ A little of the time
- ◯ None of the time

64) SF36Q11a- Get ill more easily?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

65) SF36Q11b- Healthy as anybody I know ?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

66) SF36Q11c- Expect health to get worse?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

67) SF36Q11d- My health is excellent?
- ◯ Definitely true
- ◯ Mostly true
- ◯ Don't know
- ◯ Mostly false
- ◯ Definitely false

68) EQ5DLQ1- Mobility
- ◯ I have no problems in walking about
- ◯ I have slight problems in walking about
- ◯ I have moderate problems in walking about
- ◯ I have severe problems in walking about
- ◯ I am unable to walk about

69) EQ5DLQ2- Self care
- ◯ I have no problems washing or dressing myself
- ◯ I have slight problems washing or dressing myself
- ◯ I have moderate problems washing or dressing myself
- ◯ I have severe problems washing or dressing myself
- ◯ I am unable to wash or dress myself

70) EQ5DLQ3 - Usual activities
- ◯ I have no problems doing my usual activities
- ◯ I have slight problems doing my usual activities
- ◯ I have moderate problems doing my usual activities
- ◯ I have severe problems doing my usual activities
- ◯ I am unable to do my usual activities

71) EQ5DLQ4- Pain/discomfort
- ◯ I have no pain or discomfort
- ◯ I have slight pain or discomfort
- ◯ I have moderate pain or discomfort
- ◯ I have severe pain or discomfort
- ◯ I have extreme pain or discomfort

72) EQ5DLQ5- Anxiety/depression
- ◯ I am not anxious or depressed
- ◯ I am slightly anxious or depressed
- ◯ I am moderately anxious or depressed
- ◯ I am severely anxious or depressed
- ◯ I am extremely anxious or depressed

73) EQ5DLQ6- Visual analogue health score today

0 50 100

(Place a mark on the scale above)

74) AFEQTQ1- Palpitations

- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

75) AFEQTQ2- Irregular heartbeat

- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

76) AFEQTQ3- Pause in heart activity

- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

77) AFEQTQ4- Lightheadedness or dizziness

- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

78) AFEQTQ5- Ability to have recreational pastimes, sports and hobbies

- ◯ Not at all limited
- ◯ Hardly limited
- ◯ A little limited
- ◯ Moderately limited
- ◯ Quite a bit limited
- ◯ Very limited
- ◯ Extremely limited

79) AFEQTQ6 - Ability to have a relationship and do things with friends and family

- ◯ Not at all limited
- ◯ Hardly limited
- ◯ A little limited
- ◯ Moderately limited
- ◯ Quite a bit limited
- ◯ Very limited
- ◯ Extremely limited

80) AFEQTQ7- Difficulty doing activity due to tiredness, fatigue or low on energy

- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

81) AFEQTQ8- Doing physical activity due to SOB
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

82) AFEQTQ9- During exercise
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

83) AFEQTQ10- Walking briskly
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

84) AFEQTQ11- Walking briskly uphill or carrying groceries or other items, up a flight of stairs without stopping
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

85) AFEQTQ12- During vigorous activity such as lifting or moving heavy furniture, running, or participating in strenuous sports
- ◯ No difficulty at all
- ◯ Hardly any difficulty
- ◯ A little difficulty
- ◯ Moderate difficulty
- ◯ Quite a bit of difficulty
- ◯ A lot of difficulty
- ◯ Extreme difficulty

86) AFEQTQ13- Anxious that AF can start anytime
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

87) AFEQTQ14- Anxious AF may worsen other medical conditions in the long run
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

88) AFEQTQ15- Worried about the treatment side effects from medications
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

89) AFEQTQ16- Worried about complications from procedures (ablation, surgery, pacemaker therapy)
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

90) AFEQTQ17- Worried about side effects from blood thinners (bleeding, bruises)
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

91) AFEQTQ18-Worried or feeling anxious that treatment interferes with daily activities
- ◯ Not at all bothered, do no have this symptom
- ◯ Hardly bothered
- ◯ A little bothered
- ◯ Moderately bothered
- ◯ Quite a bit bothered
- ◯ Very bothered
- ◯ Extremely bothered

92) AFEQTQ19- How well does your current treatment control your atrial fibrillation?
- ◯ Extremely satisfied
- ◯ Very satisfied
- ◯ Somewhat satisfied
- ◯ Mixed with satisfied or dissatisfied
- ◯ Somewhat dissatisfied
- ◯ Very dissatisfied
- ◯ Extremely dissatisfied

93) AFEQTQ20- To what extent has the treatment relieved your symptoms of atrial fibrillation?
- ◯ Extremely satisfied
- ◯ Very satisfied
- ◯ Somewhat satisfied
- ◯ Mixed with satisfied or dissatisfied
- ◯ Somewhat dissatisfied
- ◯ Very dissatisfied
- ◯ Extremely dissatisfied

Appendix 6: Patient information leaflet for RATE AF wearables sub-study

Optional sub-study of physical activity and heart rate monitoring in the RATE-AF trial

Atrial fibrillation affects both heart rate and the physical activity that patients can manage day-to-day. Very little is known about how heart rate and physical activity interact, or how the treatments we use can benefit patients.

As part of the RATE-AF study, patients are asked to wear a wrist-worn device which collects information about your heart rate, movements and sleep pattern. The device connects to a mobile phone which we will give you for the duration of this study. You can use the mobile phone as a regular telephone.

This study is an extra part of the RATE-AF trial – it is entirely optional. Please take time to read the following information carefully, and feel free to ask us any questions you may have. If you decide not to participate, it will not affect your involvement in the RATE-AF trial or your NHS care, and you are free to withdraw at any time without giving a reason.

What is the purpose of the study?

The RATE-AF trial is testing how medications can be used to improve quality of life for patients with atrial fibrillation. We can only measure heart rate when patients come to visit us, and physical activity is often difficult to describe – for example, how much time you spend walking or doing any exercise. Using a digital device worn on the wrist gives us the opportunity to measure these precisely, and could lead to improvements in how we manage patients in the NHS in the future.

What will happen to me if I take part?

You will be given a wrist device and mobile phone with full instructions. You will be asked to wear it on your wrist for 24 hours a day for up to 1 year, taking it off only for charging (every few days) and when taking a shower, bath or swimming. It measures your heart rate and activity level and connects to the phone automatically using wireless technology. The mobile phone will send the information to the collaborators, along with information about your surroundings, phone use, activity and movements. The phone also has an application that enables you to record your symptoms and quality of life. You can use the phone to make calls to all UK landlines and mobiles, or access the internet, all free of charge.

Who will have access to my data and how will it be used?

Information from the wrist device and phone will be encrypted (scrambled). This study is part of a collaboration with University of Birmingham, University College London, and a European Union funded project called BigData@Heart, which aims to improve the care of patients with heart disease. Data will be securely hosted on computers managed by this collaboration. The information will be coded and cannot be linked to your identifiable personal health details stored at University Hospitals Birmingham NHS Trust. For example, your location will be masked in a way which does not reveal your actual location to prevent identification of your home address or your precise location. In the same way, while we collect information about the number and length of calls or texts, we will not be able to see the content of the messages, or the name or number of the person you have contacted. Any information you provide us with via these devices will not be checked or acted upon by members of the research team.

If we lose contact with the device or mobile phone, or if you ask for assistance, a team from this collaboration are able to access your name and contact details to provide support.

The research data collected for this study may be used for future studies not related to the RATE-AF trial. The research data will be held confidentially and will not be linked to your identifiable personal health details.

More information?

Please ask the research team – we are happy to answer any questions. The RATE-AF trial information leaflet gives further information on your rights, oversight by the Ethics Committee and who to contact in case of any problems.

Thank you for taking the time to read this information leaflet and for considering taking part in this research project.

Appendix 7: Optional consent form for RATE-AF wearables sub-study

University Hospitals Birmingham NHS
NHS Foundation Trust

CONFIDENTIAL WHEN COMPLETE

RATE-AF Trial
Participant Optional Consent Form

Participant Trial Number: ☐☐☐☐ **Please initial each box to confirm consent ↓**

1. I agree that a sample of my blood donated by me may be stored for future genetic (DNA) tests to help understand effects of the study treatment and genetic factors that might contribute to atrial fibrillation. This sample will be for medical research **only** and my results will be kept confidential. Any study using this material would require Research Ethics Committee approval. I understand that these samples may be analysed in research laboratories outside of this hospital, in the UK and oversees, within Europe. ☐

2. I agree that a sample of my blood donated by me may be stored for future biochemical tests to help understand effects of the study treatment and biochemical factors that might contribute to atrial fibrillation. This sample will be for medical research **only** and my results will be kept confidential. Any study using this material would require Research Ethics Committee approval. I understand that these samples may be analysed in research laboratories outside of this hospital, in the UK and oversees, within Europe. ☐

3. I understand that the information held and maintained by NHS Digital and other central UK bodies may be used to help contact me or provide information about my health status in the future. I understand that my name, postcode and date of birth will be shared with these central bodies. ☐

4. I am happy to be contacted by the Research Team to participate in a focus group to discuss my experience of participating in the trial. ☐

5. I am happy to be contacted in the future about relevant studies. ☐

6. I agree to participate in the **sub-study of physical activity and heart rate monitoring** in the RATE-AF trial after reading the additional participant information leaflet (v1.0 23-Jan-2018). I understand my participation is voluntary and I can withdraw at any time. I also understand that the research data collected as part of this physical activity sub-study will be kept confidential and may be used in future research unrelated to the RATE-AF trial. ☐

7. I agree to participate in **the sub-study of nerve activity and heart rate in the RATE-AF trial** after reading the additional participant information leaflet (v1.0 23-Jan-2018). I understand my participation is voluntary and I can withdraw at any time. I also understand that the research data collected as part of this nerve activity sub-study will be kept confidential and may be used in future research unrelated to the RATE-AF trial. ☐

Points 1-7 above are OPTIONAL for participation in RATE-AF; please initial these boxes only if you agree to them.

Name of participant	Date	Signature
		Witness Signature *(if applicable)*
Name of person receiving consent	Date	Signature

Original to be filed in the Investigator's Site File; 1 copy for patient; 1 copy to be kept with patient's hospital record; 1 copy to be sent to BCTU