Study Title: Predicting AF after Cardiac Surgery - the PARADISE Score. A Clinical Prediction Rule for Post-operative Atrial Fibrillation in Patients Undergoing Cardiac Surgery

Internal Reference Number / Short title: PARADISE

Ethics Ref: 21/EE/0166

CAG Ref: 21/CAG/0097

IRAS Project ID: 296508

Date and Version No: 03/07/2023 Version 5

Joint Research Office Ref: PID15669

ClinicalTrials.gov: NCT05255224

HTA Project: NIHR131227

Chief Investigator:	Professor Peter Watkinson, University of Oxford
Lead Investigator:	Professor Benjamin O'Brien, Perioperative Medicine Barts Heart Centre (Joint Lead Applicant)
Investigators:	Dr Rui Providencia, Clinical Scientist, University College London
	Professor Gregory Lip, Price-Evans Chair of Cardiovascular Medicine, Liverpool Heart and Chest Hospital NHS Foundation Trust
	Professor Jochen Muehlschlegel, Distinguished Chair of Anesthesiology, Brigham and Women's Hospital, Associate Professor, Harvard Medical School
	Professor Gary Collins, Professor of Medical Statistics University of Oxford
	Professor David Clifton, Professor of Clinical Machine Learning, University of Oxford
	Dr Oliver C Redfern, Clinical Researcher, University of Oxford
	Dr Jonathan Bedford, NIHR Doctoral Fellow, University of Oxford
Sponsor:	University of Oxford

Clinical Trials & Research Governance,

Boundary Brook House,

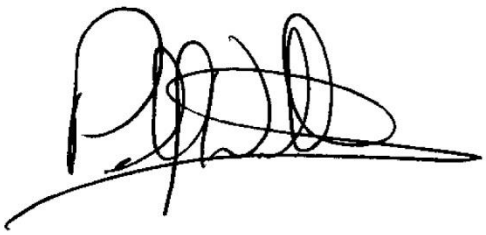
Oxford,

OX3 7GB,

United Kingdom

Funder: NIHR National Institute of Health Research, Health Technology Assessment Programme HTA Project: NIHR131227

Chief Investigator Signature:



Statistician Signature:

Conflicts of interest:

BOB has received funding from the British Heart Foundation for the Tight K trial Grant Number: CS/18/3/34063 clinicaltrials.gov: NCT04053816

GYHL: Consultant for Bayer/Janssen, BMS/Pfizer, Medtronic, Boehringer Ingelheim, Novartis, Verseon and Daiichi-Sankyo. Speaker for Bayer, BMS/Pfizer, Medtronic, Boehringer Ingelheim, and Daiichi-Sankyo. No fees are directly received personally.

PW has received funding from the NIHR for the CAFE study and the RRAM study. He was previously Chief Medical officer for Sensyne Health and has shares in the company. Sensyne Health have provided research support to his university department.

GSC was a member of the NIHR HTA commissioning board from 2016 to 2020.

Confidentiality Statement

This document contains confidential information that must not be disclosed to anyone other than the Sponsor, the Investigator Team, HRA, host organisation, and members of the Research Ethics Committee, unless authorised to do so.

TABLE OF CONTENTS

1 KEY CONTACTS

Chief Investigator	Professor Peter Watkinson Kadoorie Centre, University of Oxford, Level 3, John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU 01865 857613 peter.watkinson@ndcn.ox.ac.uk
Sponsor	University of Oxford Clinical Trials & Research Governance, Boundary Brook House, Oxford, OX3 7GB, United Kingdom Telephone: +44 1865 616480 Email: ctrg@admin.ox.ac.uk
Funder(s)	NIHR National Institute of Health Research, Health Technology Assessment Programme
Clinical Trials Unit	Oxford Clinical Trials Research Unit Botnar Research Centre University of Oxford
Lead Statistician	Professor Gary Collins Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal sciences 01865 223463 gary.collins@csm.ox.ac.uk
Committees	**Project Oversight Group**

	The project oversight group will comprise the CI (Peter Watkinson) and co-lead investigator (Ben O'Brien), two PPI members, the Founder & CEO of the Atrial Fibrillation Association, a Consultant Cardiologist, and Professor of Cardiovascular Medicine, from the East and North Hertfordshire NHS Trust and University of Hertfordshire and a Professor of Stroke and Older People's Care at the University of Central Lancashire. The final member is still to be recruited.

2 LAY SUMMARY

Atrial Fibrillation (AF) is a common abnormal heart rhythm. AF causes the heart to beat irregularly and sometimes very rapidly. About 30-50% of patients develop AF after heart surgery. These patients stay longer on the Intensive Care Unit (ICU) after surgery, are more likely to develop complications and have a higher risk of dying. Avoiding AF is important.

Some drugs, including beta blockers and amiodarone may help prevent AF if given after surgery. However, these may also lead to complications (such as lung damage). It is therefore important to identify which patients are most likely to benefit from these treatments (i.e., where the benefits outweigh the risks). There are existing tools designed to predict the risk of suffering AF after heart surgery. However, they are unreliable and therefore not used in clinical practice. A modern, reliable risk prediction tool is needed.

The PARADISE study will develop and test new prediction tools to identify which patients are most at risk of developing AF after heart surgery. We will focus our tools on those patients who most commonly develop AF, such as those who have had surgery to repair a valve or blood vessel in their heart.

To do this we will:

- Review the medical literature and assemble a panel of medical experts to create a list of known factors that affect patients' risk of AF after heart surgery
- Use a large UK general practice database (CALIBER) to see whether we can find new risk factors.
- Ask the expert panel to agree a list of known and new risks factors to be included in the prediction tool.
- Develop two new prediction tools using an existing American cardiac surgery database (the Partners research Database). The first will be used before surgery, the second immediately following surgery. Two models are needed as events during surgery may alter the risk of AF.
- Test how reliably our new tools predict which patients suffer AF after surgery, with data from large UK NHS heart centres, one US Hospital (Brigham) and a UK clinical trial (Tight-K).
- We will work with two charities (AF Alliance and StopAfib) to share our results with patients and the wider public.

3 SYNOPSIS

Study Title	Predicting AF after Cardiac Surgery - the PARADISE Score. A Clinical Prediction Rule for Post-operative Atrial Fibrillation in Patients Undergoing Cardiac Surgery		
Internal ref. no. / short title	PARADISE		
Study registration	ClinicalTrials.gov: NCT05255224		
Sponsor	University of Oxford Clinical Trials & Research Governance, Boundary Brook House, Oxford, OX3 7GB, United Kingdom		
Funder	NIHR National Institute of Health Research, Health Technology Assessment Programme (NIHR131227)		
Study Design	Longitudinal cohort study		
Study Participants	Patients who have undergone cardiac surgery		
Sample Size	Retrospective: 38, 000 Prospective: 13, 684		
Planned Study Period	Funded study period: 1st February 2021 to 30th June 2024 Prospective sub-study period: 1st October 2021 to 31st July 2023		
Planned Recruitment period	Retrospective: 1st January 1998 to 31st December 2020 Prospective: 1st October 2021 to 31st July 2023		
	Objectives	Outcome Measures	Timepoint(s)
Primary	**Primary Objective** To develop and externally validate two prognostic models to predict post-operative atrial fibrillation after cardiac surgery using data available in the pre-operative assessment clinic or on admission for surgery (PARADISE-1) and on arrival in the post-operative care unit (PARADISE-2); and compare their performance to other published models	**Primary outcome** Discrimination (c-statistic) and calibration (intercept and slope) in external dataset **Secondary outcomes** Positive and negative predictive values, sensitivity and specificity	7 days after cardiac surgery
Secondary	1. Systematic literature review 2. Analysis of CALIBER database using statistical and machine learning methods	Candidate risk factors for inclusion in new onset atrial fibrillation prognostic models	For pre-operative model, factors up to surgery, for post-operative model factors up to 12 hours post surgery.

	3. Modified Delphi Process 4. Exploring combinations of risk factors in CALIBER and PRD 5. Expert panel – consolidating risk factors to be used in model development		

4 ABBREVIATIONS

ACE	Angiotensin-Converting Enzyme
AF	Atrial Fibrillation
AFACS	Atrial Fibrillation after cardiac surgery
AUROC	Area Under Receiver-Operator Curve
BHC	Barts Heart Centre
BHF	British Heart Foundation
CABG	Coronary Artery Bypass Graft
CAG	Confidentiality Advisory Group
CALIBER	CArdiovascular disease research using LInked Bespoke studies and Electronic health Records
CENTRAL	Cochrane Central Register of Controlled Trials
CHARMS	Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies
CI	Chief Investigator
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CRF	Case Report Form
CRN	Clinical Research Nurse
CTRG	Clinical Trials & Research Governance, University of Oxford
GCP	Good Clinical Practice
UK GDPR	UK General Data Protection Regulation
GP	General Practitioner
HRA	Health Research Authority
ICU	Intensive Care Unit

LASSO	Least Absolute Shrinkage and Selection Operator
LHCH	Liverpool Heart & Chest Hospital
MEDLINE	Medical Literature Analysis and Retrieval System Online
MeSH	Medical Subject Headings
ML	Machine Learning
NHS	National Health Service
NICOR	National Institute for Cardiovascular Outcomes Research
OCTRU	Oxford Clinical Trial Research Unit
OPCS-4	Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures (4th revision)
OpenSIGLE	Open System for Information on Grey Literature in Europe
PROBAST	Prediction model Risk Of Bias ASsessment Tool
PRISMA-P	Preferred Reporting Items for Systematic review and Meta-Analysis Protocols
PROSPERO	International Prospective Register of Systematic Reviews
QUIPS	Quality In Prognosis Studies
PI	Principal Investigator
PIL	Participant/ Patient Information Leaflet
POG	Project Oversight Group
PRD	Partners Research Database
R&D	NHS Trust R&D Department
REC	Research Ethics Committee
RES	Research Ethics Service
SOP	Standard Operating Procedure
SMG	Study Management Group
TIGHT-K	TIGHT-K trial
UK	United Kingdom
US	United States

5 BACKGROUND AND RATIONALE

5.1 Background

Atrial Fibrillation after cardiac surgery (AFACS) is the most common complication following cardiac surgery, with an incidence between 30% and 50%. The risk is higher for patients having either isolated mitral valve or combined mitral valve and coronary artery bypass graft surgery (1–3). Around 35,000 patients undergoing cardiac surgery in the UK every year (4). In the United States, around 500,000

patients undergo cardiac surgery annually (5). AFACS is strongly associated with adverse patient outcomes, longer hospital and ICU stays, increased risk of stroke, increased risk of developing long-term AF (with associated complications and need for anticoagulation), and increased all-cause 30-day and 6-month mortality (6,7,16,8–15). In the US alone, this translates to an additional $10,000–20,000 in hospital costs for each patient who suffers AFACS, resulting in an estimated $1 billion in overall healthcare costs per year (2). Interventions that reduce the incidence of AFACS would have a substantial impact both on patient outcomes and cost.

Even though AFACS can be transient and patients are often discharged from hospital in normal sinus rhythm, patients with new-onset AFACS have a 5-fold increased risk of developing long-term AF (1). A number of risk factors have been identified, including (older age, obesity) (17,18), comorbidities (e.g. prior AF, hypertension, chronic kidney disease, obstructive sleep apnoea) (19), surgical features (e.g. valve surgery, increased aortic cross-clamp time, acute kidney injury) (20,21) and the failure or inability to (re-)introduce Beta-Blockers. However, each of these factors only accounts for part of the overall risk (22). Importantly, even when controlling for these confounding variables are, long-duration AFACS (> 2 days) is independently associated with decreased survival (23), as well as stroke and respiratory complications (24). Current evidence therefore indicates that AFACS itself contributes to poor patient outcomes following cardiac surgery, and that tools to predict, prevent and guide treatment of AFACS are needed.

Currently, there is no widely accepted prediction model that reliably allows clinicians to determine the risk of a patient developing AFACS, despite multiple efforts over the past 15 years to develop one (25–31). Published AFACS prediction models are limited by small sample sizes, a failure to include modern variables and a lack of external prospective validation (26,27,29,32–36). Moreover, our team and others have previously developed AFACS risk prediction models that suffer from limited generalizability, over-simplification, and the fact that they often use information from variables that occur after surgery, or even after onset of AFACS (22,35,37). A recent analysis showed the best area under the receiver operator characteristic curve (AUROC) for any of the published prediction models was 0.68 (95% confidence interval [CI]: 0.67–0.69) (38,39), just marginally better than a model which only included age (0.66, 95% CI: 0.65–0.68) (22). The work emphasised the importance of using representative and detailed multicentre data sets to develop prediction models and prospectively validating these on multiple cohorts.

The lack of effective pre and immediate post-operative prediction models for estimating AFACS risk has prevented the implementation of AF prophylaxis protocols (38). Interventions to prevent AFACS lead to decreased hospital length of stay, lower costs of hospital treatment, and decreased risk of post-operative stroke (40). Pharmacological (e.g., Amiodarone, beta blockers, electrolyte supplementation) and non-pharmacological (e.g., atrial pacing, posterior pericardiotomy, botulinum toxin) therapies provide practical methods to significantly reduce AFACS incidence (2). In selected patients, some peri-operative prophylactic therapies reduce the risk of AFACS (41,42), and have been included in guidelines (43–46). However, these interventions are not effective in all patients and some have a high risk of adverse side effects. Therefore, it is crucial to use clinical information to target interventions to those individuals at highest risk of developing AFACS, who will more likely benefit from prophylaxis. Furthermore, monitoring patients after hospital discharge shows significant numbers of early asymptomatic recurrences. These patients are often at high risk of stroke and may benefit from anticoagulation for stroke prevention.

Improved prediction of AFACS could facilitate targeted monitoring that is more likely to identify AFACS recurrence (1).

This study will develop and validate two different scores as they apply to different situations, in the pre-operative assessment clinic (PARADISE 1) and on arrival in the post-operative care unit (PARADISE 2), have different input variables (the latter including intra-operative and immediately available post-operative variables) and different time horizons for prediction (from weeks before the operation and from immediately post-operative in the post-operative care unit). These two models will have two distinct use cases. For example, PARADISE-1 could be used to start prophylactic medication in the pre-operative assessment clinic, whereas PARADISE 2 will be used to adjust risk on admission to the post-operative care unit. Of the patients who develop AFACS, 70% do so before post-operative day 4 and 94% by post-operative day 7 (8,47). Therefore, both models will be optimised to predict AFACS within the first week of the initial post-operative hospital stay. This ensures we address the unmet need for prediction of AF following cardiac surgery, rather than AF in the context of prolonged hospitalisation or critical illness.

5.2 Aim

The study aims to develop and validate two clinical prediction models to estimate the risk of a patient developing AF in the seven days following cardiac surgery, using data available:

- In the pre-operative assessment clinic or on admission for surgery (PARADISE-1)
- On arrival in the post-operative care unit (PARADISE-2)

6 OBJECTIVES AND OUTCOME MEASURES

Objectives	**Outcome Measures**	**Timepoint(s) of evaluation of this outcome measure (if applicable)**
Primary Objective To develop and externally validate two prognostic models to predict post-operative atrial fibrillation after cardiac surgery using data available in the pre-operative assessment clinic or on admission for surgery (PARADISE-1) and on arrival in the post-operative care unit (PARADISE-2); and compare their performance to other published models	**Primary outcome** Discrimination (c-statistic) and calibration (intercept and slope) in external dataset **Secondary outcomes** Positive and negative predictive values, sensitivity and specificity	7 days after cardiac surgery

Secondary Objectives 1. Systematic literature review 2. Analysis of CALIBER database using statistical and machine learning methods 3. Modified Delphi Process 4. Exploring combinations of risk factors in CALIBER and PRD 5. Expert panel – consolidating risk factors to be used in model development	Candidate risk factors for inclusion in new onset atrial fibrillation prognostic models	7 days after cardiac surgery

7 STUDY DESIGN

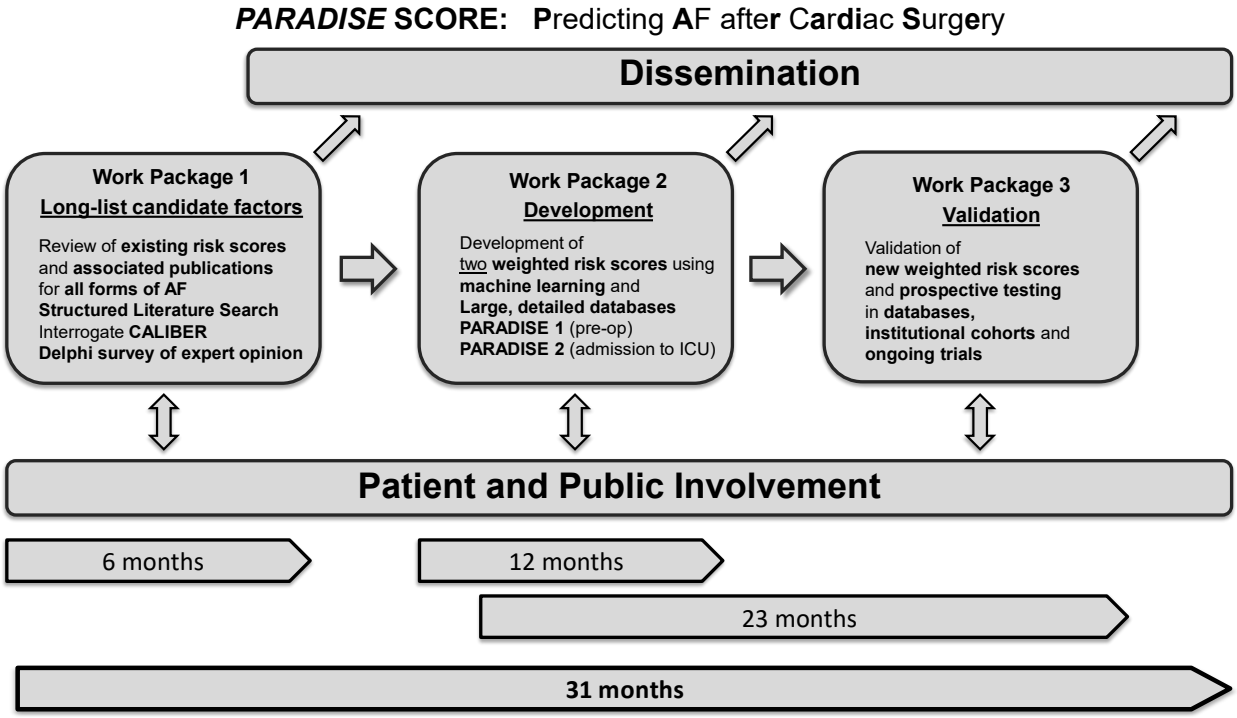
7.1 Overview

This is an international, multi-centre longitudinal cohort study of patients who have undergone cardiac surgery. The study will use retrospective data to develop clinical prediction models to estimate the risk of developing atrial fibrillation in the seven days following surgery. We will develop two models: PARADISE-1 and PARADISE-2. PARADISE-1 will only use patient data collected in pre-operative period (i.e. in the pre-operative assessment clinic or on admission for surgery). PARADISE-2 will include additional data available on arrival in the post-operative care unit. Both models will then be externally validated on prospectively collected data from large UK centres and one UK clinical trial.

7.1.1 Project flowchart

Prediction models will be developed and validated as follows:

1. Identify a long-list of candidate risk factors for AFACS
2. Model development – PARADISE-1 and PARADISE-2
3. Model validation – PARADISE-1 and PARADISE-2



7.2 Setting and data collection

The study will use data collected from patients who have undergone cardiac surgery in specialist centres in the United States and United Kingdom. There are four databases and a prospective trial data set.

7.2.1 Databases

1. **CALIBER**: The Health Data Research UK CALIBER research platform contains variables extracted from linked electronic health records from primary care, hospital records, social deprivation information and cause-specific mortality data in England. This dataset includes electronic health records from 6.5 million individuals between 1998 and 2016.
2. **PRD**: The Partners Research Database (PRD) contains detailed modern data for over 30,000 individuals undergoing cardiac surgery from 2 centres in the US between 2001-2020. PRD includes pre-operative echocardiographic measurements.
   ***Note**: During the study, we will expand this database to include pre-and peri-operative variables from the clinical data warehouses of the two centres. It will be used for model development.*
3. **Brigham**: The Brigham and Women's CABG Genomics Database is a prospective single centre detailed research database comprising 3000 individuals from 2001-2016. This dataset contains over 1,700 prospectively collected and curated pre, intra, and post-operative variables. It will be used for external US retrospective model validation.
4. **NHS Trusts**: The study will prospectively collect data from the UK NHS Trusts which will allow us to undertake external validation of PARADISE-1/PARADISE-2 with a total of approximately 6000 patients (2000 AFACS events) per annum. The NHS Trusts where data will be extracted from are:
   a. Liverpool Heart & Chest NHS Foundation Trust
   b. Barts Heart Centre (Barts Health NHS Trust)
   c. Oxford University Hospitals NHS Foundation Trust

7.2.2 Trial data sets

1. **Tight K study:** (BHF-funded CS/18/3/34063, ClinicalTrials.gov Identifier: NCT04053816). Tight K is a 22-centre clinical trial that will include 1684 patients, randomised within this non-inferiority trial to investigate whether maintaining serum potassium ≥3.6mEq/L is equivalent to maintaining levels at 4.5- 5.5mEq/L, with a primary endpoint of development of AFACS using precise clinical definitions and 5-day Holter monitoring. The trial is already designed to collect detailed data on AFACS risk factors, minimising any additional burden of PARADISE. It is restricted to patients undergoing isolated coronary artery bypass grafting. This has the strength that detailed data are being collected in a well-defined cohort, but means we need additional prospective data to assess the performance of the PARADISE scores in valve surgery.

7.3 Identifying long-list of risk factors for AFACS

Prior to model development, we will use a diverse range of approaches to identify a comprehensive list of potential risk factors for AFACS:

- Literature – systematic literature review
- Data-driven identification – CALIBER database
- Expert panel – modified Delphi process
- Machine learning methods – exploring combinations of risk factors in the CALIBER and PRD databases
- Expert panel – consolidating risk factors to be used in model development

7.3.1 Literature – systematic literature review

We will perform a systematic review and critical appraisal of pre-operative variables predictive of AFACS. We will include existing prediction models and associated publications, both for AFACS and other AF substrates, such as primary AF (not after intervention), AF in critically ill patients and post-operative AF after non-cardiac surgery.

An initial search of MEDLINE, Embase, and Web of Science using MeSH and free-text search terms, retrieved 3,170 papers. We will expand our search to include: CINAHL, Conference Proceedings Citation Index: Science, OpenSIGLE, the Cochrane Database of Systematic Reviews, and the Cochrane Central Register of Controlled Trials (CENTRAL). We estimate 300-450 full papers will need to be assessed. We will incorporate findings from our published work in general ICU populations (NIHR 17/71/04) and follow guidance by the Cochrane Prognosis Methods Group (48,49). We will snowball to identify any further relevant studies. Searches will be performed by a critical care information specialist, without date or language restrictions. Data will be extracted in duplicate. Authors will be contacted by e-mail where additional information is required.

As most studies identified in our initial search were non-randomised, we are likely to evaluate how tools such as QUIPS for prognostic factors and PROBAST for prediction models will inform our evidence assessment (50,51). The data extraction form will follow the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) guidance (52). We will synthesise available data using a semi-quantitative method previously described by Zaal et al. (53) and adapted by Dettmer et al (54).

We will register and publish the systematic review protocol on PROSPERO following PRISMA-P guidance (55).

7.3.2 Data-driven identification – CALIBER database

We will use patient data within the CALIBER research platform to identify risk factors (comorbidities, medical conditions, and medications) prior to cardiac surgery associated with incident AF.

CALIBER links primary care and hospital records, allowing us to construct a detailed medical history for each patient prior to cardiac surgery. We will use OPCS-4 codes to identify patients who have undergone cardiac surgery. We will identify patients' comorbidities from both primary and secondary care diagnostic coding. Medications will be grouped into broad classes (e.g. ACE inhibitors) using standard ontologies. For medications we will also explore any associations with the duration of exposure to the drug. We will explore univariate relationships between candidate variables and AFACS will be explored using standard statistical Cox models and machine-learning techniques (see Section 10 for further details)

7.3.3 Expert panel – modified Delphi process

We will select a panel of international experts from relevant publications identified in our systematic review. The panel will undertake a modified Delphi process to identify a list of potential risk factors for AFACS.

Stage 1 of the Delphi process will involve panel members suggesting all risk factors they feel may be predictive of AFACS. These candidate variables will be collated and passed back to the panel in Stage 2. The panel will review all variables generated during stage 1, selecting those they judge to be most important. Variables meeting a consensus agreement of 60% will be accepted (56,57); those <20% will be rejected. In stage 3, variables with 20-60% agreement will be presented to the panel, along with their respective percentage agreement, for a final round of voting. Variables that achieve at least 60% agreement will be added to those previously accepted in Stage 2 to form the final list of candidate variables from the Delphi process.

7.3.4 Machine learning methods – exploring combinations of risk factors in CALIBER and PRD

We will combine all candidate variables identified from the systematic review, CALIBER analysis and Delphi into a combined "long-list". Any variables missing from the PRD (to be used for model development) will be extracted from the electronic patient record. We will then apply machine learning methods (see Section 10) to the CALIBER and PRD data sets to derive combinations of risk factors (i.e. interactions) that should be considered for inclusion in model development.

7.3.5 Expert panel – consolidating risk factors to be used in model development

At their final meeting, the expert panel will review the final long-list of candidate variables (and their combinations). Any variables judged to be redundant or likely to represent spurious correlations will be removed.

7.4 Model development – PARADISE-1 and PARADISE-2

Using the PRD data set, we will develop two prognostic models (PARADISE-1 and PARADISE-2) to predict AFACS using the long-list of candidate variables identified above.

PARADISE-1 will only use patient data collected in pre-operative period (i.e. in the pre-operative assessment clinic or on admission for surgery).

PARADISE-2 will include additional data available on arrival in the post-operative care unit.

In developing both models, we will consider standard statistical approaches (e.g. logistic regression) and machine learning methods (e.g. random forests, deep neural networks). Model performance will be assessed using standard metrics (e.g. calibration, discrimination) and for optimism using internal validation. Further details are given in Section 10.

7.5 Model validation – PARADISE-1 and PARADISE-2

Evaluating the prediction models using the same data used to develop them will lead to optimistic estimations of performance. Therefore, we will externally validate (assessing calibration and discrimination) both models using:

- Retrospective data from one US hospital (Brigham)
- Externally validate models on prospective data from UK NHS Trusts
- Externally validate models on prospective data from a clinical trial data set (TIGHT-K)

Retrospective external validation on the Brigham data set will comprise 3000 individuals (≈900 AFACS events, much larger than the minimum recommended sample size of 100 events (58).

Prospective external validation will be performed on four cohorts. Data collected within clinical trials and "real-world" data from patients receiving routine clinical care offer different strengths and weaknesses when validating a prediction models (59). We will therefore undertake prospective external validations in both trial patients (using data from the Tight K study CS/18/3/34063 and PROPHESY studies (BHF pilot PG/17/82/33368) and in a prospective cohort of ≈6000 patients developed from enhanced routine data collection at NHS trusts. Data from the Tight K and PROPHESY studies bring the advantages of multiple centres and study-grade data collection, but in relatively small numbers – we estimate between one half and two thirds of the total intended recruitment will be available for model validation, totalling over 1000 trial patients. Data from the NHS Trusts provide a large "real-world" prospective validation, but will likely contain more missing data (60).

Ethical approvals for enhanced data collection will be obtained prior to analysis.

As part of the external validation, we will assess the need for recalibration and use decision curve analysis to identify appropriate cut-offs for risk stratification (61). Further details are given in Section 10.

We will compare our prediction models to those identified in the systematic review. These will include the CHA_2DS_2VASc score (originally designed to predict stroke in AF) which has also been used to predict AFACS (62,63)

8 PARTICIPANT IDENTIFICATION

8.1 Study Participants

The study will include all adult patients who underwent or are about to undergo cardiac surgery in any of the study databases or clinical trial data sets.

8.2 Inclusion Criteria

- Patients 18 years or over
- Admitted to hospital for any cardiac surgery

8.3 Exclusion Criteria

- Patients who have requested that their data not be used for research (e.g. NHS Opt-out, see Section 9.3.1)

9 PROTOCOL PROCEDURES

9.1 Recruitment

Participants will be recruited from the study databases and trial data sets, as described in section 7.2.

9.2 Screening and Eligibility Assessment

9.3 Informed Consent

Existing ethical approvals are in place for the CALIBER (UK), PRD (US) and Brigham (US) databases.

For the remaining data sets (including the TIGHT-K trial) used for external validation, we will apply for Health Research Approval (HRA), under advice from the Research Ethics Committee and the Confidentiality Advisory Group (CAG) to allow access to confidential medical records without specific written consent (Section 251 support).

9.3.1 Opt-out

For data sets that include UK participants, the study will fully support and implement the NHS National Opt-out. We will use our link to NHS Digital asking them to inform us which patients have opted out and that will be a continuous process from enrolment until the study ends. Records that are already flagged within each site as meeting the Opt-out will not be extracted.

If we are informed by NHS Digital of patients completing the Opt-out after data extraction, these data will be purged and future records relating to these patients will not be extracted.

We will also purge records from the study (in the same manner) if patients contact the study team directly.

We will clearly display on the website links to register for the opt-out, as well as contact details for the study team.

9.4 Enrolment

This is a non-randomised observational study. Enrolment of participants will occur through their data being present within the participating databases or prior enrolment into the Tight-K study.

9.5 Description of study intervention(s), comparators and study procedures (clinical)

This is a non-interventional study.

9.6 Baseline Assessments

Not applicable

9.7 Subsequent Visits

Not applicable

9.8 Sample Handling

No additional samples will be taken during this observational study.

9.9 Early Discontinuation/Withdrawal of Participants

This is an observational study. Participants can request their data to be deleted at any time in accordance with GDPR and the study privacy policy.

9.10 Definition of End of Study

The end of the study will be when all prospective data sets have been collected and used to externally validate both prediction models (PARADISE-1, PARADISE-2).

10 SAFETY REPORTING

This is an observational study, so safety reporting is not applicable.

11 STATISTICS AND ANALYSIS

11.1 Statistical Analysis Plan (SAP)

The statistical aspects of the study are summarised here, with details fully described in a statistical analysis plan that will be available before any analysis takes place.

11.2 Description of the Statistical Methods

11.2.1 Model development

In the PRD, we anticipate that around one third of patients will experience AFACS (10,000 events), so the issue of overfitting is negligible. Using logistic regression, we will select variables using least angle selection and shrinkage operator (LASSO) penalties (64). Continuous predictors exhibiting a nonlinear relationship with AFACS will be modelled using fractional polynomials.

Heterogeneity in model performance will be explored over different hospitals using internal-external cross-validation (65). We will overcome challenges arising from the use of treatments during the study in two ways: Treatments at or before baseline will be handled by including treatments as predictors (66). The impact of any treatments on the absolute risk of AFACS during follow-up will be investigated using inverse probability weighting (59).

In parallel we will investigate machine learning (ML) approaches to developing a prediction model, including methods such as random forests, deep neural networks, and Bayesian Gaussian processes. Such models allow consideration of interactions between available variables providing "latent variables" (complex, non-linear transformations of the original input variables) that may improve prediction performance over the non-transformed input variables. We will prioritise principled, probabilistic methods that permit the incorporation of prior clinical knowledge, such that results are "interpretable", avoiding the "black box" nature of much ML-oriented research in this area (67,68). We will exploit the ability of the complementary modern ML fields of (i) Bayesian non-parametric modelling (which provides a principled means of coping with artefact and measurement noise in the clinical data) and (ii) deep learning, which permits the fusion of large quantities of (potentially time-varying) clinical variables. These models include components ("attention mechanisms") that allow the resulting risk estimations to be interpretable for clinicians. Importantly, these models permit the model to quantify its uncertainty in its score prediction – whereas conventional medical statistical methods always output a score, ML models have the opportunity to quantity probabilistically the certainty with which a score is produced; this offers advantages in applications in which classes are substantially overlapped, as often occurs in complex, real-world clinical problems, where there is imperfect class separation. Accuracy of such models can often be substantially improved by permitting a model not to make a classification if the score is highly uncertain; such approaches are used in many application areas of critical systems to improve acceptability with users.

Additionally, we will investigate the joint construction of phenotypical clustering models (which identify reproducible phenotypes in the physiological data) with predictive models specific to each phenotypical group – this novel approach constructs an ensemble of predictive models that are phenotype-specific, thereby improving predictive accuracy with respect to a single "one-size-fits-all" generic algorithm applied to the entire population.

11.2.2 Model performance and validation

The performance of a prediction model is typically assessed by its calibration and discrimination.

Calibration, which reflects how close the predictions from the model are to the observed outcome frequencies will be assessed graphically, using a calibration plot, plotting observed outcomes against predictions using smoothing techniques. The plot will also be supplemented with results for individuals grouped by similar probabilities (tenths) comparing the mean predicted probability to the mean observed outcome. Calibration will also be quantified by calculating the calibration slope and intercept.

The discrimination of the prediction models will be summarised with the concordance index (equivalent to the Area Under Receiver Operating Characteristic curve) with 95% confidence interval.

An important goal of a prediction model is to classify patients into risk groups. Both models will produce a risk score (probability) for each patient, based on their own predictor values. We will then identify a cut-off value to decide when the risk is high (such that we predict poor outcome) and when it is low (such that we predict a good outcome). The cut-off value (or range of), will be examined using decision curve analysis, which calculates the net benefit of using the model (compared to not using the model). The net benefit of a model is the difference between the proportion of true positives and the proportion of false positives weighted by the odds of the selected threshold for high-risk designation (61).

The internal validity of the final models will also be assessed by the bootstrap re-sampling technique to adjust for over-optimism in the estimation of model performance. ML models will also be additionally validated via k-fold cross-validation. The internal validation will quantify and be used to adjust the performance measures (e.g. discrimination, calibration) for any optimism.

To account for potential differences in case-mix (distribution of patient characteristics and prevalence of the outcome) between the US data used to develop the model and the prospective external validation, we will investigate whether recalibration is needed to update the model to better fit the UK population (60).

Finally, we will compare our prediction models to those identified in the systematic review.

11.3 Sample Size Determination

The model development data set (PRD) contains detailed data for >30,000 individuals undergoing cardiac surgery from 2 centres, 2001-2020. Of these, we expect around one third of patients to have experienced AFACS (10,000 events)

Retrospective external validation on the Brigham data set will comprise 3000 individuals (≈900 AFACS events, much larger than the minimum recommended sample size of 100 events (58).

Data from patients within trials and “real-world” data from patients receiving routine clinical care offer different strengths and weaknesses when validating a prediction rule (59). We will therefore undertake prospective external validations in both trial patients (using data from the ongoing British Heart Foundation (BHF) funded Tight K study CS/18/3/34063 and in a prospective cohort of ≈6000 patients developed from enhanced routine data collection at NHS Trusts.

11.4 Analysis populations

Our analysis populations will include all eligible participants in each study data set, as described above.

11.5 Decision points

Not applicable

11.6 Stopping rules

Not applicable

11.7 The Level of Statistical Significance

The level of statistical significance will be set at $p<0.05$.

11.8 Procedure for Accounting for Missing, Unused, and Spurious Data.

To avoid excluding patients when developing and validating our models, we will use multiple imputation to impute missing values, under a missing at random assumption. Identifying the true underlying missing data mechanism from the available is rarely possible. Assumptions need to be made on the plausible mechanism, and approaches needed to be used. Under a missing completely at random mechanism.

(MCAR), the missingness mechanism does not depend on unobserved (unseen) data. Carrying out a complete case analysis will produce unbiased estimates (but with a loss precision if full data are observed). Under the missing at random (MAR) assumption, the missingness after conditioning on the observed data does not depend on the unobserved (unseen) data. Under this approach, we can apply approaches such as multiple imputation, by fitting a joint model to the observed data and impute the missing data, taking account of the uncertainty in the estimated parameters of this joint model. We feel this, MAR, approach makes a less strong and more realistic assumption than the MCAR approach. The MAR imputation model will include all variables considered for the multivariable model building, the outcome and any auxiliary variables that will help explain the missingness. The assumption of a missing not at random (MNAR) approach whilst not implausible is considerably more complex to investigate – there is a dearth of research investigating MNAR in the context of prediction model research. We will nevertheless explore whether the MAR assumption holds by comparing the imputed values (after accounting for the observed values) and the missing values to identify if there are any systematic differences to suggest a MNAR assumption.

Data from the NHS Trusts provide a large "real-world" prospective validation, but is likely to contain more missingness (60).

11.9 Procedures for Reporting any Deviation(s) from the Original Statistical Plan

Any deviation from the statistical analysis plan will be described and justified in all study publications.

12 DATA MANAGEMENT

The plan for the data management of the study is outlined below. There is not a separate Data Management document in use for the study.

12.1 Source Data

The study will use data collected from patients who have undergone cardiac surgery in the specialist centres in the United States and United Kingdom. There are four databases and one prospective trial data set. All data will be anonymised and assigned a pseudonymous "study key" prior to analysis.

Creation of the NHS Trust data set requires accessing and linking confidential health care records. Access to these data will be subject to approval from a REC and Section 251 support from the CAG.

All pseudonymised data transferred to the Sponsor for analysis will be held within a secure "Data Safe Haven", which is owned and maintained by the Critical Care Research Group (University of Oxford). The "Data Safe Haven" is a computing environment which has been designed to store and analyse large datasets in a manner that is safe and secure. It conforms to NHS Digital Security Toolkit and Cyber Essentials Plus accreditation. The environment is designed to prevent patient level data leaving this environment.

12.1.1 Databases

1. **CALIBER**
2. **PRD**
3. **Brigham**
4. **NHS Trusts**

12.1.2 Trial data sets

1. **Tight K study:** (BHF-funded CS/18/3/34063, ClinicalTrials.gov Identifier: NCT04053816, REC: 19/LO/1064).

12.2 Access to Data

Creation of the NHS Trusts data set requires accessing and linking confidential health care records. Where variables (e.g. ECHO findings) have not been transferred into a structured data source, research nurses from the study team will manually extract them from participants' electronic records. All members of the study team involved in this process will undergo appropriate local governance training.

Pseudonymised data set will be accessed by authorised members of the study team at the University of Oxford for analysis purposes. At the University of Oxford, access will be via the Data Safe Haven subject to compliance with local information governance policies.

12.3 Data Recording and Record Keeping

All study records will be electronic – an overview of their generation is detailed above. All records will be subject to quality assurance polices both at the University and research group level. These are designed to guarantee the accuracy and validity of the study data.

The participants will be identified by a unique study number (pseudonymous study key).

We will use the CALIBER and PRD data sources to validate known (and potentially identify novel) risk factors for AF after cardiac surgery. A complete list of risk factors will not be known a priori, as it depends on the results of the systematic review and Delphi process. However, likely required data will be contained within following categories:

- Demographics (e.g., age, sex, ethnicity)
- Diagnostic coding (e.g., known chronic conditions)
- Limited pre- and post-operative medications (e.g., beta-blockers, vasoactive drugs)
- Limited surgical procedures and intraoperative findings (e.g., type of cardiac surgery, duration of bypass, blood transfusion)
- laboratory blood tests (including blood gases)
- vital signs (e.g., heart rate, blood pressure)
- Limited Echocardiogram findings
- Limited electrocardiogram (ECG) findings

Excluding age, sex and ethnicity we will not be accessing data on protected characteristics (e.g., religion, civil status, sexual orientation).

The expert panel will agree a final list of risk factors, which will be used to develop the two prediction models.

The models will be validated in one US data set (Brigham), one clinical trial data set (Tight-K) and prospective data collected from UK cardiac surgery centres (NHS Trusts). Only data from the above categories included in the candidate risk factor list agreed by the expert panel will be extracted from the UK cardiac centres along with the required outcome data for model assessment (including in-hospital

mortality and episodes of atrial fibrillation). Tight-K is an ongoing clinical trial with its own approvals. Data for the NHS Trust data sets will be acquired from the individual Trusts patient records in collaboration with the respective organisations.

13 QUALITY ASSURANCE PROCEDURES

All research team members will be trained in Information Governance, data protection and confidentiality. The study may be monitored, or audited in accordance with the current approved protocol, GCP, relevant regulations and standard operating procedures.

13.1 Risk assessment

This is a retrospective observational study, where researchers will not interact directly with patients or intervene in their care. We will undertake and document a formal risk assessment of the project.

This study requires access to confidential patient records. Eligible patient records will be identified by the participating organisations. Directly identifiable data will only be required for record linkage and will not be available to the Sponsor.

To mitigate the risk of reidentification of participants and the risk of data loss we will undertake the following:

- All records will be accessed and de-identified at each participating site, using a dedicated computer that will conform to NHS information security standards.
- Only pseudonymous personal data will be transferred via secure/encrypted protocols to the coordinating centre (Critical Care Research Group (CCRG), Nuffield Department of Clinical Neurosciences, Oxford University).
- Only pseudonymous personal data will be held by the CCRG.
- Pseudonymous personal data will be held inside the Sponsors “Data Safe Haven” which conforms to the same NHS standards of information security and cyber security.

13.2 Study monitoring

All research team members will be fully trained in Information Governance, data protection and confidentiality. The study may be monitored, or audited in accordance with the current approved protocol, GCP, relevant regulations and standard operating procedures.

13.3 Study Committees

13.3.1 Study management group (SMG)

The study management group will consist of the Chief Investigator and the named investigators listed under Key Contacts, and individuals directly funded by the project. The study management group will be primarily responsible for the running and conduct of the study. They will be responsible for ensuring that standard operating procedures are followed and that regulations are adhered to. Where appropriate public patient involvement will be gained in any changes or amendments that are needed during the study.

13.3.2 Project Oversight Group (POG)

The project oversight group, is made up of the CI (Peter Watkinson) and co-lead investigator (Ben O'Brien), two PPI members, the Founder & CEO of the Atrial Fibrillation Association, a Consultant Cardiologist, and Professor of Cardiovascular Medicine, from the East and North Hertfordshire NHS Trust and University of Hertfordshire, a Professor of Stroke and Older People's Care at the University of Central Lancashire, a Professor of Evidence Synthesis at the University of York; the Director of Nursing & Midwifery for the National Institute for Health Research (NIHR) and a Professor of Medical Statistics, from the University of Leicester.

14 PROTOCOL DEVIATIONS

Any deviations from the protocol will be documented in a protocol deviation form and filed in the study master file.

15 SERIOUS BREACHES

A "serious breach" is a breach of the protocol or of the conditions or principles of Good Clinical Practice which is likely to affect to a significant degree –

(a) the safety or physical or mental integrity of the trial subjects; or

(b) the scientific value of the research.

This is a non-interventional study and will not interact directly with its participants, only their data. The risk of a serious breach is therefore extremely low. However, the study will process directly identifiable patient data at the NHS Trusts to create an anonymised data for model validation. This present potential for a data breach. In the event that a data breach is suspected, the study team will follow local hospital policies relating to GDPR.

The Sponsor will also be contacted within 1 working day. In collaboration with the CI, the serious breach will be reviewed by the Sponsor and, if appropriate, the Sponsor will report it to the approving REC/CAG committees and the relevant NHS host organisation within seven calendar days.

16 ETHICAL AND REGULATORY CONSIDERATIONS

16.1 Declaration of Helsinki

The Investigator will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki.

16.2 Guidelines for Good Clinical Practice

The Investigator will ensure that this study is conducted in accordance with relevant regulations and with Good Clinical Practice.

16.3 Approvals

Following Sponsor approval, the protocol, and any associated documentation will be submitted to an appropriate Research Ethics Committee (REC), CAG and HRA (where required) and host institutions for written approval.

The Investigator will submit and, where necessary, obtain approval from the above parties for all substantial amendments to the original approved documents.

16.4 Other Ethical Considerations

The CALIBER research platform has existing research ethics approval (09/H0810/16) and ECC approval (ECC 2-06(b)/2009 CALIBER dataset). The PRD has existing research ethics approval (Ref:2016P001986) from the Mass General Brigham Institutional Review Board.

16.5 Reporting

The CI shall submit once a year throughout the study, or on request, an Annual Progress report to the REC Committee, CAG, HRA (where required) host organisation, Sponsor, and funder (where required). In addition, an End of Study notification and final report will be submitted to the same parties.

16.6 Transparency in Research

Not applicable as the research is non-interventional.

16.7 Participant Confidentiality

The study will comply with the UK General Data Protection Regulation (UK GDPR) and Data Protection Act 2018, which require data to be de-identified as soon as it is practical to do so. The processing of the personal data of participants will be minimised by making use of a unique participant study number only on all study documents and any electronic database(s). All documents will be stored securely and only accessible by study staff and authorised personnel. The study staff will safeguard the privacy of participants' personal data.

16.8 Expenses and Benefits

No payments or any other benefits will be provided to participants.

17 FINANCE AND INSURANCE

17.1 Funding

The study is funded by the NIHR National Institute of Health Research, Health Technology Assessment Programme HTA Project: NIHR131227.

17.2 Insurance

The University has a specialist insurance policy in place which would operate in the event of any participant suffering harm as a result of their involvement in the research (Newline Underwriting Management Ltd, at Lloyd's of London).

17.3 Contractual arrangements

Appropriate contractual arrangements will be put in place with all third parties.

18 PUBLICATION POLICY

The Investigators will be involved in reviewing drafts of the manuscripts, abstracts, press releases and any other publications arising from the study. Authors will acknowledge that the study was funded by the NIHR National Institute of Health Research, Health Technology Assessment Programme HTA Project: NIHR131227. Authorship will be determined in accordance with the ICMJE guidelines and other contributors will be acknowledged.

19. DEVELOPMENT OF A NEW PRODUCT/ PROCESS OR THE GENERATION OF INTELLECTUAL PROPERTY

Ownership of IP generated by employees of the University of Oxford vests in the University. The University will ensure appropriate arrangements are in place as regards any new IP arising from the study.

19 ARCHIVING

All pseudonymised data collected as part of this study will be stored securely within the CCRG Data Safe Haven for a minimum of five years, in keeping with the MRC Retention Framework for Research Data and Records. The CCRG Data Safe Haven is a secure environment run by the Critical Care Research Group (University of Oxford), which conforms to Data Security and Protection Toolkit standards. Data will not be used or released from this environment and it will be deleted at the end of its retention period.

20 REFERENCES

1. Lowres N, Mulcahy G, Jin K, Gallagher R, Neubeck L, Freedman B. Incidence of postoperative atrial fibrillation recurrence in patients discharged in sinus rhythm after cardiac surgery: A systematic review and meta-analysis. Interact Cardiovasc Thorac Surg. 2018;26(3):504–11.

2. Greenberg JW, Lancaster TS, Schuessler RB, Melby SJ. Postoperative atrial fibrillation following cardiac surgery: a persistent complication. Eur J Cardio-Thoracic Surg [Internet]. 2017 Oct 1;52(4):665–72. Available from: http://academic.oup.com/ejcts/article/52/4/665/3091250/Postoperative-atrial-fibrillation-following

3. Yadava M, Hughey AB, Crawford TC. Postoperative Atrial Fibrillation. Cardiol Clin [Internet]. 2014 Nov;32(4):627–36. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0733865114000563

4. Society for Cardiothoracic Surgery in Great Britain and Northern Ireland Blue Book. 2013;

5. Badhwar V, Rankin JS, Thourani VH, D'Agostino RS, Habib RH, Shahian DM, et al. The Society of

Thoracic Surgeons Adult Cardiac Surgery Database: 2018 Update on Research: Outcomes Analysis, Quality Improvement, and Patient Safety. Ann Thorac Surg. 2018;106(1):8–13.

6. Almassi GH, Wagner TH, Carr B, Hattler B, Collins JF, Quin JA, et al. Postoperative atrial fibrillation impacts on costs and one-year clinical outcomes: The veterans affairs randomized on/off bypass trial. Ann Thorac Surg. 2015;99(1):109–14.

7. Aranki S, Shaw D, Adams D, Rizzo R, Couper G, VanderVliet M, et al. Predictors of Atrial Fibrillation After Coronary Artery Surgery Current Trends and Impact on Hospital Resources. Circulation. 1996;94:390–7.

8. Gialdini G, Nearing K, Bhave PD, Bonuccelli U, Iadecola C, Healey JS, et al. Perioperative Atrial Fibrillation and the Long-term Risk of Ischemic Stroke. JAMA [Internet]. 2014 Aug 13;312(6):616. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2014.9143

9. Horwich P, Buth KJ, Légaré JF. New onset postoperative atrial fibrillation is associated with a long-term risk for stroke and death following cardiac surgery. J Card Surg. 2013;28(1):8–13.

10. Kim K, Yang P-S, Jang E, Yu HT, Kim T-H, Uhm J-S, et al. Long-Term Impact of Newly Diagnosed Atrial Fibrillation During Critical Care. Chest [Internet]. 2019 Sep;156(3):518–28. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0012369219308761

11. Konstantino Y, Zelnik Yovel D, Friger MD, Sahar G, Knyazer B, Amit G. Postoperative Atrial Fibrillation Following Coronary Artery Bypass Graft Surgery Predicts Long-Term Atrial Fibrillation and Stroke. Isr Med Assoc J. 2016;18(12):744–8.

12. Lapar DJ, Speir AM, Crosby IK, Fonner E, Brown M, Rich JB, et al. Postoperative atrial fibrillation significantly increases mortality, hospital readmission, and hospital costs. Ann Thorac Surg. 2014;98(2):527–33.

13. Chung MK. Atrial fibrillation after cardiac surgery. In: Atrial Fibrillation. 2004. p. 279–337.

14. Saxena A, Dinh D, Dimitriou J, Reid C, Smith J, Shardey G, et al. Preoperative atrial fibrillation is an independent risk factor for mid-term mortality after concomitant aortic valve replacement and coronary artery bypass graft surgery. Interact Cardiovasc Thorac Surg. 2013;16(4):488–94.

15. Shen J, Lall S, Zheng V, Buckley P, Damiano RJ, Schuessler RB. The persistent problem of new-onset postoperative atrial fibrillation: A single-institution experience over two decades. J Thorac Cardiovasc Surg. 2011;141(2):559–70.

16. Wang TKM, Ramanathan T, Choi DH-M, Gamble G, Ruygrok P. Preoperative atrial fibrillation predicts mortality and morbidity after aortic valve replacement. Interact Cardiovasc Thorac Surg [Internet]. 2014 Aug;19(2):218–22. Available from: https://academic.oup.com/icvts/article-lookup/doi/10.1093/icvts/ivu128

17. Phan K, Khuong JN, Xu J, Kanagaratnam A, Yan TD. Obesity and postoperative atrial fibrillation in patients undergoing cardiac surgery: Systematic review and meta-analysis. Int J Cardiol [Internet]. 2016 Aug;217:49–57. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0167527316308981

18. Wong CX, Sullivan T, Sun MT, Mahajan R, Pathak RK, Middeldorp M, et al. Obesity and the Risk of Incident, Post-Operative, and Post-Ablation Atrial Fibrillation. JACC Clin Electrophysiol [Internet]. 2015 Jun;1(3):139–52. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2405500X15000997

19. Qaddoura A, Kabali C, Drew D, van Oosten EM, Michael KA, Redfearn DP, et al. Obstructive Sleep Apnea as a Predictor of Atrial Fibrillation After Coronary Artery Bypass Grafting: A Systematic Review and Meta-analysis. Can J Cardiol. 2014;30(12):1516–22.

20. Cole OM, Tosif S, Shaw M, Lip GYH. Acute Kidney Injury and Postoperative Atrial Fibrillation In Patients Undergoing Cardiac Surgery. J Cardiothorac Vasc Anesth. 2020;000:1–8.

21. Indja B, Woldendorp K, Vallely MP, Grieve SM. New Onset Atrial Fibrillation Following Transcatheter and Surgical Aortic Valve Replacement: A Systematic Review and Meta-Analysis. Hear Lung Circ. 2020;(April):1–12.

22. Pollock BD, Filardo G, da Graca B, Phan TK, Ailawadi G, Thourani V, et al. Predicting New-Onset Post-Coronary Artery Bypass Graft Atrial Fibrillation With Existing Risk Scores. Ann Thorac Surg. 2018;105(1):115–21.

23. Sigurdsson MI, Longford NT, Heydarpour M, Saddic L, Chang TW, Fox AA, et al. Duration of Postoperative Atrial Fibrillation After Cardiac Surgery Is Associated With Worsened Long-Term Survival. Ann Thorac Surg. 2016;102(6):2018–26.

24. Megens MR, Churilov L, Thijs V. New-Onset Atrial Fibrillation After Coronary Artery Bypass Graft and Long-Term Risk of Stroke: A Meta-Analysis. J Am Heart Assoc. 2017;6(12).

25. Amar D, Shi W, Hogue CW, Zhang H, Passman RS, Thomas B, et al. Clinical prediction rule for atrial fibrillation after coronary artery bypass grafting. J Am Coll Cardiol. 2004;44(6):1248–53.

26. Magee MJ, Herbert MA, Dewey TM, Edgerton JR, Ryan WH, Prince S, et al. Atrial Fibrillation After Coronary Artery Bypass Grafting Surgery: Development of a Predictive Risk Algorithm. Ann Thorac Surg. 2007;83(5):1707–12.

27. Mathew JP, Fontes ML, Tudor IC, Ramsay J, Duke P, Mazer CD, et al. A Multicenter Risk Index for Atrial Fibrillation after Cardiac Surgery. J Am Med Assoc. 2004;291(14):1720–9.

28. Mahoney EM, Thompson TD, Veledar E, Williams J, Weintraub WS. Cost-effectiveness of targeting patients undergoing cardiac surgery for therapy with intravenous amiodarone to prevent atrial fibrillation. J Am Coll Cardiol. 2002;40(4):737–45.

29. Mariscalco G, Biancari F, Zanobini M, Cottini M, Piffaretti G, Saccocci M, et al. Bedside tool for predicting the risk of postoperative atrial fibrillation after cardiac surgery: the POAF score. J Am Heart Assoc. 2014;3(2).

30. Thorén E, Hellgren L, Jidéus L, Ståhle E. Prediction of postoperative atrial fibrillation in a large coronary artery bypass grafting cohort. Interact Cardiovasc Thorac Surg. 2012;14(5):588–93.

31. Zaman AG, Archbold RA, Helft G, Paul EA, Curzen NP, Mills PG. Atrial fibrillation after coronary artery bypass surgery: A model for preoperative risk stratification. Circulation. 2000;101(12):1403–8.

32. El-Chami MF, Kilgo PD, Elfstrom KM, Halkos M, Thourani V, Lattouf OM, et al. Prediction of new onset atrial fibrillation after cardiac revascularization surgery. Am J Cardiol. 2012;110(5):649–54.

33. Helgadottir S, Sigurdsson MI, Ingvarsdottir IL, Arnar DO, Gudbjartsson T. Atrial fibrillation following cardiac surgery: risk analysis and long-term survival. J Cardiothorac Surg. 2012;7(1):1–7.

34. Borde D, Gandhe U, Hargave N, Mathew M, Joshi S, Pandey K. Prediction of postoperative atrial fibrillation after coronary artery bypass grafting surgery: Is CHA 2 DS 2 -VASc score useful? Ann Card Anaesth [Internet]. 2014;17(3):182. Available from:

http://www.annals.in/text.asp?2014/17/3/182/135841

35. Kolek MJ, Muehlschlegel JD, Bush WS, Parvez B, Murray KT, Stein CM, et al. Genetic and Clinical Risk Prediction Model for Postoperative Atrial Fibrillation. Circ Arrhythmia Electrophysiol [Internet]. 2015 Feb;8(1):25–31. Available from: https://www.ahajournals.org/doi/10.1161/CIRCEP.114.002300

36. Tran DTT, Perry JJ, Dupuis JY, Elmestekawy E, Wells GA. Predicting New-Onset Postoperative Atrial Fibrillation in Cardiac Surgery Patients. J Cardiothorac Vasc Anesth. 2015;29(5):1117–26.

37. Cameron MJ, Tran DTT, Abboud J, Newton EK, Rashidian H, Dupuis J-Y. Prospective External Validation of Three Preoperative Risk Scores for Prediction of New Onset Atrial Fibrillation After Cardiac Surgery. Anesth Analg [Internet]. 2018 Jan;126(1):33–8. Available from: http://journals.lww.com/00000539-201801000-00013

38. Muehlschlegel JD, Burrage PS, Ngai JY, Prutkin JM, Huang C-C, Xu X, et al. Society of Cardiovascular Anesthesiologists/European Association of Cardiothoracic Anaesthetists Practice Advisory for the Management of Perioperative Atrial Fibrillation in Patients Undergoing Cardiac Surgery. Anesth Analg [Internet]. 2019 Jan;128(1):33–42. Available from: http://journals.lww.com/00000539-201901000-00011

39. O'Brien B, Burrage PS, Ngai JY, Prutkin JM, Huang CC, Xu X, et al. Society of Cardiovascular Anesthesiologists/European Association of Cardiothoracic Anaesthetists Practice Advisory for the Management of Perioperative Atrial Fibrillation in Patients Undergoing Cardiac Surgery. J Cardiothorac Vasc Anesth. 2019;33(1):12–26.

40. Arsenault K, Yusuf A, Crystal E, Healey J, Morillo C, Nair G, et al. Interventions for Preventing Post-Operative Atrial Fibrillation in Patients Undergoing Heart Surgery. Cochrane Database Syst Rev. 2013;1:CD003611.

41. CHATTERJEE S, SARDAR P, MUKHERJEE D, LICHSTEIN E, AIKAT S. Timing and Route of Amiodarone for Prevention of Postoperative Atrial Fibrillation after Cardiac Surgery: A Network Regression Meta-analysis. Pacing Clin Electrophysiol [Internet]. 2013 Aug;36(8):1017–23. Available from: http://doi.wiley.com/10.1111/pace.12140

42. Burgess DC, Kilborn MJ, Keech AC. Interventions for prevention of post-operative atrial fibrillation and its complications after cardiac surgery: A meta-analysis. Eur Heart J. 2006;27(23):2846–57.

43. Gillis AM, Verma A, Talajic M, Nattel S, Dorian P. Canadian Cardiovascular Society atrial fibrillation guidelines 2010: Rate and rhythm management. Can J Cardiol. 2011;27(1):47–59.

44. Hillis LD, Smith PK, Anderson JL, Bittl JA, Bridges CR, Byrne JG, et al. 2011 ACCF/AHA guideline for coronary artery bypass graft surgery: Executive summary. J Am Coll Cardiol. 2011;58(24):2584–614.

45. Jones C, Pollit V, Fitzmaurice D, Cowan C. The management of atrial fibrillation: summary of updated NICE guidance. BMJ [Internet]. 2014 Jun 19;348(jun19 1):g3655–g3655. Available from: https://www.bmj.com/lookup/doi/10.1136/bmj.g3655

46. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: A report of the American College of cardiology/American heart association task force on practice guidelines and the heart rhythm society. Circulation. 2014;130(23):e199–267.

47. Echahidi N, Pibarot P, O'Hara G, Mathieu P. Mechanisms, Prevention, and Treatment of Atrial

Fibrillation After Cardiac Surgery. J Am Coll Cardiol. 2008;51(8):793–801.

48. Bedford JP, Harford M, Petrinic T, Young JD, Watkinson PJ. Risk factors for new-onset atrial fibrillation on the general adult ICU: A systematic review. J Crit Care [Internet]. 2019 Oct;53:169–75. Available from: https://doi.org/10.1016/j.jcrc.2019.06.015

49. Riley RD, Moons KGM, Snell KIE, Ensor J, Hooft L, Altman DG, et al. A guide to systematic review and meta-analysis of prognostic factor studies. BMJ [Internet]. 2019 Jan 30;364:k4597. Available from: https://www.bmj.com/lookup/doi/10.1136/bmj.k4597

50. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med [Internet]. 2019 Jan 1;170(1):51. Available from: http://annals.org/article.aspx?doi=10.7326/M18-1376

51. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280–6.

52. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Med. 2014;11(10).

53. Zaal IJ, Devlin JW, Peelen LM, Slooter AJC. A Systematic Review of Risk Factors for Delirium in the ICU*. Crit Care Med. 2014/09/25. 2015;43(1):40–7.

54. Dettmer MR, Damuth E, Zarbiv S, Mitchell JA, Bartock JL, Trzeciak S. Prognostic Factors for Long-Term Mortality in Critically Ill Patients Treated with Prolonged Mechanical Ventilation: A Systematic Review. Crit Care Med. 2017;45(1):69–74.

55. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1–9.

56. Hunter DJ, Arden N, Conaghan PG, Eckstein F, Gold G, Grainger A, et al. Definition of osteoarthritis on MRI: results of a Delphi exercise. Osteoarthr Cartil [Internet]. 2011 Aug;19(8):963–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1063458411001464

57. Jackson KA, Glyn-Jones S, Batt ME, Arden NK, Newton JL. Assessing risk factors for early hip osteoarthritis in activity-related hip pain: a Delphi study. BMJ Open [Internet]. 2015 Sep 29;5(9):e007609. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2015-007609

58. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: A resampling study. Stat Med. 2016;35(2):214–26.

59. Pajouheshnia R, Peelen LM, Moons KGM, Reitsma JB, Groenwold RHH. Accounting for treatment use when validating a prognostic model: A simulation study. BMC Med Res Methodol. 2017;17(1):1–12.

60. Fanaroff AC, Steffel J, Alexander JH, Lip GYH, Califf RM, Lopes RD. Stroke prevention in atrial fibrillation: Re-defining “real-world data” within the broader data universe. Eur Heart J. 2018;39(32):2932-2941b.

61. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. Med Decis Mak [Internet]. 2006 Nov 5;26(6):565–74. Available from: http://journals.sagepub.com/doi/10.1177/0272989X06295361

62. Burgos LM, Seoane L, Parodi JB, Espinoza J, Galizia Brito V, Benzadón M, et al. Postoperative atrial fibrillation is associated with higher scores on predictive indices. J Thorac Cardiovasc Surg. 2019;157(6):2279–86.

63. Chen Y lin, Zeng M, Liu Y, Xu Y, Bai Y, Cao L, et al. CHA2DS2-VASc Score for Identifying Patients at High Risk of Postoperative Atrial Fibrillation After Cardiac Surgery: A Meta-analysis. Ann Thorac Surg. 2020;109(4):1210–6.

64. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B. 1996;58:265–88.

65. Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: Opportunities and challenges. BMJ. 2016;353:27–30.

66. Groenwold RHH, Moons KGM, Pajouheshnia R, Altman DG, Collins GS, Debray TPA, et al. Explicit inclusion of treatment in prognostic modeling was recommended in observational and randomized settings. J Clin Epidemiol. 2016;78:90–100.

67. Shamout FE, Zhu T, Sharma P, Watkinson PJ, Clifton DA. Deep Interpretable Early Warning System for the Detection of Clinical Deterioration. IEEE J Biomed Heal Informatics [Internet]. 2019;PP(Li):1–1. Available from: https://ieeexplore.ieee.org/document/8844833/

68. Yang Y, Walker TM, Walker AS, Wilson DJ, Peto TEA, Crook DW, et al. DeepAMR for predicting co-occurrent resistance of Mycobacterium tuberculosis. Bioinformatics. 2019;35(18):3240–9.

21 APPENDIX A: AMENDMENT HISTORY

Amendment No.	Protocol Version No.	Date issued	Author(s) of changes	Details of Changes made
	1	01/06/2021	Peter Watkinson and study team	N/A
	2	03/08/2021	Peter Watkinson and study team	Changed study start date to 1/10/21; Clarification of primary outcome; Addition of secondary outcomes measure to primary objective; Revision of Plain English Summary; Details of model variables added to Section 12.3; Clarification of Project Oversight group in section 13.3.2
	3	04/03/2022	Peter Watkinson and study team	Addition of clinical trials.gov registration; Specific references to The Liverpool Heart & Chest Hospital (LHCH) and Barts Heart Centre (BHC) have been replaced with “NHS Trusts” to

				cover additional study sites
	4	14/06/2022	Peter Watkinson and study team	Study period separated into “Funded” and “Prospective” to clarify NIHR funding period.
	5	03/07/2023	Peter Watkinson and study team	Funded study end date changed to 30 June 2024 in Section 3, as agreed with the funder (National Institute for Health and Care Research, NIHR). The CAG reference number has been corrected to 21/CAG/0097 in the study documentation.

List details of all protocol amendments here whenever a new version of the protocol is produced.

Protocol amendments must be submitted to the Sponsor for approval prior to submission to the REC committee and HRA (where required).