Clinical Trials Research Unit

(CTRU)

University of Leeds

Statistical Analysis Plan

AMAZE

Version 1.0

15th March 2016

Produced by:

CTRU Statistician Colin Everett

Supervising Statistician Prof Linda Sharples

1 Introduction

1.1 Background

Atrial fibrillation (AF) is the most common disturbance of heart rhythm. With a UK prevalence of 7.2% in patients aged 65 and over and 10.3% in patients aged 75 and over [1] AF has considerable impact on quality of life and NHS resources[2][3]. Treatment of AF and its consequences (anti-arrhythmic & anti-coagulant drugs, hospital monitoring & stroke treatment) are expensive for the NHS and implementation of the recent NICE guidelines (June 2006)[2] on management of AF is estimated to cost £21.86m per year[3]. The NHS devotes 5% of its budget to strokes and 15% of these are associated with AF[1]. Routine anticoagulation is used to reduce the risk of stroke, however this incurs an increased risk of bleeding and the burden of monitoring treatment falls on general practice, anticoagulant clinics and haematology laboratories.

AF ablation devices are a new and costly technology being marketed to treat this condition. Their use is increasing in NHS practice despite the lack of good research evidence to support adoption. Although there are instances of their use as a stand-alone procedure, they are already in use within the NHS as an adjunct procedure for patients having cardiac surgery for other problems.

1.1.1 Existing research

The current basis for treatment and management of AF is dealt with in a UK NICE Guideline (2006)[2] European Guidelines[4] and a Cochrane review[5]. International recommendations on surgical and catheter ablation of AF were published in 2007 jointly by the Heart Rhythm Society, European Heart Rhythm Association and European Cardiac Arrhythmia Society in their Expert Consensus statement[6]. The maze procedure can be performed in two ways:

1. The traditional cut-and-sew technique, known as the Cox-maze with its many modifications, is reliable in restoring sinus rhythm in the majority of patients (references cited in Calkins et al.[6]). Despite being available since 1987, this procedure has signally failed to achieve widespread use. The main reason for this is that it is technically demanding and adds substantially to the operative burden of a heart operation. It is currently in very limited use by a few surgeons in a few centres and tends to be reserved for otherwise fit patients with severely symptomatic AF who are prepared to take the risk of a major intervention to relieve their symptoms.

2. The ablation device maze procedure uses an energy source (heat, cold, radiofrequency or microwave) to replicate the lesion set of the Cox-maze. As a rule, the

procedure is safe, well tolerated and only adds minimal increase in time and burden of the operation.

Common sense suggests that treating AF at the time of cardiac surgery is advantageous to the patient. However the only evidence supporting this comes from 5 small randomised controlled trials of ablation devices as adjuncts to surgery[7][8][9][10][11]. These trials found that SR was restored in 44-94% of treated patients compared to 5-33% of controls. The trials were small and follow-up was short. Success was mostly defined on the basis of a single ECG recording. No trial looked at patient-centred outcomes or cost effectiveness. Despite this lack of robust evidence, an increasing number of patients with AF having open heart surgery are now being offered concomitant ablation maze procedures (cited in Calkins et al.[6]).

The 2007 Expert Consensus Statement on Catheter and Surgical Ablation of Atrial Fibrillation[6] developed by Heart Rhythm Society in partnership with European Heart Rhythm Association and European Cardiac Arrhythmia Society launched a call for high quality prospective multicentre trials to adopt consistent definitions of procedural success in long term assessments of the safety and efficacy of ablation.

The relevance of the Amaze trial is highlighted by the 2010 consensus statement from international cardiothoracic surgeons[12]. This statement emphasises the urgent need for adequately powered and properly designed and reported randomised trials to measure clinically relevant outcomes (e.g. stroke, symptom relief, QoL, long term mortality) and resource use[13]. In other words, dimensions of methodological quality and clinical relevance that are incorporated into the Amaze protocol.

1.1.2 Purpose of Amaze Trial

This trial responds to this call and will inform patients, clinicians and the NHS about the routine adoption of this technology. The study will test the hypothesis that treating AF by incorporating a modified maze procedure (using an ablation device) into elective cardiac surgery will promote a return to SR and improve quality of life as well as being cost-effective from an NHS perspective.

1.2 Study design

The Amaze trial is a pragmatic, multicentre, prospective, double blind, randomised controlled trial to compare clinical, patient-based and cost outcomes for patients with pre-existing AF who undergo routine cardiac surgery either with or without an adjunct device-based ablation procedure.

The trial is double blind to the extent that neither the patient nor the cardiologist who analyses the 4 day ECG, nor the quality of life assessor should be aware which group the patient has been allocated to.

Eligible patients are randomised (in a 1:1 ratio) to receive either their routine cardiac surgery with no additional procedure or their routine cardiac surgery with an additional device-based AF ablation procedure.

1.3 Study aims and objectives

1.3.1 Intermediate Primary objective

To compare two groups for the rate of return to stable SR at 12 months as well as quality-adjusted survival over 2 years, 4-day ECG monitors will be used to assess the predominant rhythm (SR or AF) and the AF load, i.e. the percentage of time that the patient is in AF if the predominant rhythm is SR.

1.3.2 Final Primary Objective

To compare Quality-Adjusted Survival in terms of Quality-Adjusted Life Years over two years between the two groups.

1.3.3 Secondary Objectives

1. To determine whether the adjunct maze procedure improves the rate of return to stable SR at 24 months after surgery.

2. To determine whether the adjunct maze procedure decreases thromboembolic neurological complications (eg. stroke).

3. To determine whether the adjunct maze procedure enables anticoagulant treatment to be withdrawn safely.

4. To determine whether the adjunct maze procedure enables safe reduction or withdrawal of antiarrhythmic medication.

5. To determine whether the adjunct maze procedure is cost effective compared to the routine procedure.

1.4 Sample size and recruitment

1.4.1 Sample size calculation

Sample size calculations are based on both primary endpoints.

Return to SR at 12 months

Published RCTs of ablation as an addition to cardiac surgery have reported rates of return to SR at 12 months [7][8][9][10][11] ranging from 44% to 87% in the trial arms and 5% to 33% in the control arms. If we take a conservative estimate of the difference between the groups (45% vs. 30%) then we would have 80% power to detect this difference with a sample size of 176 in each group, total 352 (2-sided significance 5%). With recruitment of 400 patients this would allow for deaths or loss to follow up at 12 months of approximately 15%.

Clinical effectiveness measured as quality-adjusted survival over 2 years

The emphasis in cost-effectiveness studies is on estimation rather than hypothesis testing so that formal sample size calculations are less important. However, we provide a power calculation based on the effectiveness measure QALY. We could find no studies reporting comparative QALYs in similar patients undergoing ablation and cardiac surgery. From previous studies of patients undergoing angiography for suspected ischaemic heart disease[14] and patients with refractory angina[15] the standard deviation of QALYs over 12 and 18 months is at most 0.3. Over follow up of 2 years the minimum clinically important improvement is considered to be one extra month of quality-adjusted life, or 0.083 QALYs. With a sample of 200 patients per group, total 400, we would have approximately 80% power to detect a difference of 0.083 QALYs, (2-sided significance 5%). If the accepted threshold for cost effectiveness were in the range £20-30,000 per QALY and we could demonstrate a significant increase in QALYs of 0.0833, then the procedure would be cost-effective for an incremental cost of at most £2,500.

1.5 Randomisation

Patients who fulfil the eligibility and have given written informed consent and have sufficient time for discussion and consideration, are randomised (in a 1:1 ratio) to one of two groups to receive either their routine cardiac surgery with no additional procedure or their routine cardiac surgery with an adjunct maze procedure.

Patient allocations are computer generated by the trial statistician and are in random permuted blocks of variable lengths, stratified by surgeon and by planned cardiac procedure (CABG, aortic valve, mitral valve, combined procedure).

1.6 Eligibility

Eligible patients are consecutive elective cardiac surgical patients undergoing major cardiac surgery (such as coronary, valve or combined operations) with a history of paroxysmal, persistent or chronic AF beginning more than 3 months before the date of the operation.

Paroxysmal AF is defined as recurrent AF (> 2 episodes) that terminates spontaneously within 4 days (Expert Consensus Statement on Catheter and Surgical Ablation of Atrial Fibrillation)[6].

Persistent AF is defined as AF which continues beyond 4 days.

Chronic or longstanding AF is persistent AF beyond 1 year.

Inclusion criteria:

age over 18.

elective cardiac surgery planned (Coronary surgery, valve surgery, combined coronary and valve surgery, any other Cardiac surgery requiring cardiopulmonary bypass.)

history of documented atrial fibrillation (chronic, persistent or paroxysmal) beginning more than 3 months before entry into the study.

written informed consent to participation.

Exclusion criteria:

previous cardiac operations.

emergency or salvage cardiac operations.

surgery without cardiopulmonary bypass.

unlikely to be available for follow-up over a two-year period.

deemed not competent to provide consent.

All randomised participants will be included in the final intention-to-treat analysis, except those for whom consent to use data was withdrawn, or written informed consent was not received. Deviations from these criteria will be summarised and reported.

2 Outcome

2.1 Primary outcomes

Rate of return to stable SR at 12 months- 4 day ECG monitors are being used to assess the predominant rhythm (SR or AF) and the AF load ie the percentage of time that the patient is in AF if their predominant rhythm is SR.

Clinical effectiveness quality adjusted survival over 2 years

2.2 Secondary outcomes

Clinical endpoints of SR at 24 months after surgery, overall survival and stroke-free survival, incidence of anticoagulant-related haemorrhage.

Health-related QoL measured by the EuroQoL, SF-36 and NYHA.

Resource use and cost-effectiveness of the adjunct maze procedure.

Anticoagulant and antiarrhythmic drug usage.

2.3 Missing data

Data management will focus on the consenting process, participant eligibility, safety, date consistency and test outcomes. Attempts will be made to retrieve missing data on these areas via a thorough data cleaning process.

The levels of missing data and reasons for missingness will be investigated for the consenting process, participant eligibility, safety, date consistency and test outcomes. The quantity of missing data will be monitored by treatment group, and a summary of the number of patients with missing primary endpoint data and the quantity of missing data by treatment group will be reported.

For the intermediate primary endpoint (return to sinus rhythm at 12 months), if a patient withdraws consent or is lost to follow-up for further participation within 12 months, multiple imputation used to impute missing outcome (AF or SR at 12 months) as a function of the baseline heart rhythm, surgeon, surgical procedure and treatment group. Rubin's rules will be used to combine imputed datasets. If more than 5% of patients are withdrawn or lost to follow-up before 12 months, then the outcome will be imputed under the alternative patterns as sensitivity analyses. (See section 5.4.1)

For the final primary endpoint (Quality-Adjusted Survival at 2 years), where a patient is deceased before the end of follow-up, the utility value of 0 will be imputed for all

subsequent assessments. If the response is missing, and the patient is not known to be deceased, the missing value will be imputed using the method of multiple imputation. A “Last Observation Carried Forward” approach will be used as an alternative imputation technique for imputing other (non-death) missing values in a sensitivity analysis.

The primary analysis model will only require the baseline rhythm, surgeon and surgical procedure which are immediately recorded when patients attend the preadmission clinic and has consented to participate or during the period of surgery, so there is little concerns about missing data arising in this model. If any missing values in the covariates are reported, they will be imputed using a function of the known covariates and primary outcomes of interest.

3 Population

3.1 Intention-to-treat Analysis

An intention-to-treat analysis will be the primary method for analysing and summarising the trial data. The intention-to-treat population is defined as all randomised patients, regardless of eligibility, withdrawal, compliance with the protocol, loss to follow-up or actual treatment received. Only patients who have withdrawn their consent for their data to be used in the study, or for whom written informed consent has not been received, will be excluded in this population. These patients will be analysed and summarised according to the intervention they were randomised to receive.

If more than 5% tests or trial conduct constitute a major protocol violation such as cross-over to the other arm or cancelling surgeries, the Complier Average Casual Effect analysis will be considered.

3.2 Quality of life populations

A separate quality of life population will be formed for the analysis of each questionnaire (SF36). Each population will comprise all patients who return an analysable baseline questionnaire, regardless of subsequent questionnaire return: patients without analysable baseline questionnaires will be excluded from the analysis, regardless of subsequent questionnaire return.

3.3 Safety Population

All patients will be included in the safety population if they underwent one of the two procedures. Patients will be included in the arm corresponding to the intervention received. If no intervention was received, then the patient will be summarised separately from the other intervention groups.

4 Data Collection

4.1 Methods

The data is collected on to a web-based system designed and coordinated by the Data Scan and Quality Officer at Papworth Hospital. The Clinical Research Nurse (CRN) at each centre enters the data directly on to the database. Surgical data will be recorded by a designated member of the surgical team either directly or via a paper form. All paper data collection forms are returned to the R&D Unit at Papworth. The Trial Coordinator (TC) are responsible for data monitoring and quality control. The whole process is overseen by the Trial Manager situated in the co-ordinating centre at Papworth Hospital.

4.2 Baseline data collection

We adhere to the ACC/AHA/ECS 2006 Guidelines which recommend that the initial patient description includes demographics, type and duration of AF and the planned cardiac procedure.

The first 4-day ECG recording starts after the patient attends the preadmission clinic and has consented to participate. All other baseline measurements are recorded on the day of admission for surgery. Once these measurements have been taken, the patient is registered with the co-ordinating centre's R&D unit and randomised as described in Section 1.5.

4.3 Data collection during and after surgery

Data collection is based on the recommendations of Shemin et al.[13] and includes procedural details-including the lesion set in the experimental group. Data collected after surgery includes: mortality, stroke/thromboembolic events, medications, EuroQoL, health-related quality of life, cardioversion plan if appropriate, 4-day ECG recordings, resource use, adverse events.

Data are collected during surgery, at discharge, 6 weeks after surgery (at a routine service visit), at 6, 12 and 24 months after surgery during out-patient research visits and annually thereafter by telephone follow-up.

4.4 Analysis of ECG recordings

All 4 day continuous ECG recordings will be analysed centrally at Papworth Hospital. Participating centres forward the SD cards from the ECG recorders to Papworth Hospital. Analysis using the proprietary automated software package, together with

manual checking of the recording in its entirety, will be done. Total time spent in sinus rhythm and in AF (AF burden) during the 4 day recording will be calculated, with only those episodes of AF lasting greater than 60 seconds duration included in the analysis. Episodes of atrial flutter will be noted and included in the AF burden.

Occurrences of Atrial Flutter or Atrial Tachycardia ("Organised Atrial Arrhythmia") in patients experiencing AF and Junctional Rhythm in patients reportedly in Sinus Rhythm will be reported.

4.5 ONS tracking

All patients enrolled in this trial (with their consent) are registered with the Office of National Statistics (ONS) Tracking System to allow long term follow up of survival. ONS tracking data is not expected to form part of the primary analysis of outcomes up to 24 months. Instead, this will be used to follow-up patients over a longer period if longer-term follow-up analyses are required.

4.6 Data validation

Data management will focus on the data associated with the consenting process, participant eligibility, safety, date consistency and test outcomes and this section refers to the cleaning of these items. The Data Management Assistant (DMA)/ Trial Co-ordinator (TC) will carry out initial validation of the forms in accordance with the trial-specific Data Management Work Instructions. This will ensure that data is complete, consistent, and up-to-date. The Data Clarification Form (DCF) will be sent to sites to highlight missing data items and queries associated with data collected on CRFs to date. Reasons should be obtained when data is unobtainable.

The database will validate most data in line with validation rules and highlight any issues that need further investigation i.e. with the site. Manual checks on all entered data will be performed prior to the validations being implemented. Data items collected relating to the safety and rights of individual patients are to be highlighted via priority validations and dealt with as a data management priority. Periodic batch validation will also be carried out to detect any data queries that may be missed if case record forms (CRFs) are entered in an order that does not allow real time validation checks to work.

A key data items list drawn up by the Trial Statistician that will include all data items that are required for the analysis of the primary endpoint. All key data items will be checked manually for completeness and accuracy by the DMA/TC, in addition to any automatic checks raised on the database. Data automatically generated through the 24-hour randomisation system will be checked by the Trial Statistician.

The Trial Statistician will also perform checks to identify any missing or inconsistent data and liaise with the Trial-Coordinator to resolve any queries.

The data will be validated and checked using SAS in the following steps:

The data will be read into permanent SAS data sets.

A random sample of 5 patients from each SAS dataset will be checked against the data as seen on the database to ensure that the data transfer has been successful. The names and contents of the variables can be found on the annotated final database specification reports in the Statistician's Trial File.

Data checks will include:-

Eligibility checks

Sequential dates

Checks for unusual and outlying data

Inconsistency in data between forms

Checks for missing data (are there variables which are systematically missing/do specific variables have a large amount of missing data, particularly key outcome data)

Other checks as deemed appropriate

Any inconsistent data will be noted and an e-mail sent to the trial co-ordinator responsible for the study. A copy of this e-mail will be kept in the statistician's trial file. All queries will be resolved and the outcome documented.

5 Data analysis

It is expected that the final analysis of the data will be performed when all patients have completed 24 months of follow-up.

5.1 General calculations

All statistical analyses and reporting will comply with CONSORT guidelines where possible.

Confidence intervals for a single proportion shall be calculated using the Exact method. Confidence intervals for a difference between independent proportions shall be calculated using Exact intervals (Method 8 of [16]).

All percentages will be calculated using the total number of patients within the specified analysis population, percentages will be reported to 1 decimal place. All statistical tests will be 2-sided and performed at the 5% significance level. All analyses will be carried out using SAS.

For summary statistics, the number of non-missing items, the means, standard deviations, medians, upper and lower quartiles and minima and maxima will be summarised to one more decimal place than the data are collected.

5.2 General principles

Multivariable analyses will not be `built' following a model-fitting strategy. Instead, all variables specified for inclusion will be added to the model, and the significance of each factor will be reported. Where one categorical variable has more than one `factor level' then the significance of overall effect of including all factor levels will be tested, rather than those for each individual factor level. For all factor levels, suitable point and interval estimates of effect size will be presented.

If any analysis requires the use of simulation and / or re-sampling methods, the initial `seed' value for the random number generation will be 0471346543. The same seed will be used at the start of every such analysis.

5.3 Baseline data and surgery data

Patient baseline data and surgery data as recorded on the baseline assessment or during surgery will be tabulated using frequencies and summary statistics by treatment group, for each randomising centre and in total, for the intention-to-treat population (and safety populations if appropriate). No statistical testing will be carried out on these data.

5.4 Analysis of Primary Endpoints

Quality-Adjusted Survival

For the final primary outcome, QALYs will be estimated from serial measurements of the EQ-5D for each patient up to 2 years using interpolation. The Health Economics analysis will be given in Section 6.

At baseline, on discharge and at 6 weeks, 6 months, 12 months and 24 months post-surgery all patients complete the EuroQol questionnaire, including the EQ-5D. The social tariff for the EQ-5D, as estimated by Dolan et al. [17] will be applied to each patient's self-reported classification in order to calculate utility values. Using actual rather than nominal times of assessment, and assuming a linear change in values between time points, patient-specific utility curves up to 24-months post randomisation will be calculated. A value of zero will be applied at the date of death for those patients who died.

The QALYs experienced by each patient to 24-months post randomization are calculated as the area under their utility curve to 24 months or time of death, whichever occurs first, where the true test dates rather than nominal test dates will be utilized in plotting the utility curve. In order to adjust for differences in baseline utilities a linear regression will be fitted to the utilities post treatment, with baseline utility and treatment group as explanatory variables. The linear regression will also include a random effect for surgeon if it is feasible to fit, and yields a positive variance component for the surgeon effect. Adjusted treatment effects will be taken from the treatment group coefficient of this regression. For patients who do not complete all EuroQoL measurements and are censored the methods of Willan and Lin[18] will be used to estimate mean QALYs and costs. The adequacy of model fit for the linear regression model at each timepoint will be assessed by examining distributions of standardised residuals, association with the predicted values, as well as identifying influential observations by referring to leverage statistics.

The differences in Quality-adjusted Survival will be presented. A confidence interval for the true difference will be formed using a non-parametric bootstrap resampling approach: [19]

A simple random sample with replacement will be drawn from the full analysis dataset of the same size as the full dataset. (ie some patients may be drawn more than once)

The difference in QALYs between the two treatment groups will be estimated for this bootstrap sample.

Steps (1) and (2) will be repeated 1000 times.

The 95% confidence interval will be formed as the interval between the 2.5% and the 97.5% percentile of the differences computed in these bootstrap samples.

Return to SR

The intermediate primary outcome, whether patient returns to SR at 12 months, will be summarized according to the group to which they were randomised. The comparison between the heart rhythm (AF or SR) at the baseline and 12 months will be also summarized.

For the primary endpoint analysis, the rate in SR for Routine treatment group will be compared to that for Routine+Maze treatment group using a binary logistic regression model, which will include surgeon (as a random effect), surgical procedure and baseline heart rhythm as fixed effects. The odds ratio for the rate of return to SR at 12 months of Routine treatment group against Routine+Maze treatment group will be reported with 95% confidence interval and p-value for the data seen, under the null hypothesis that the rates of return to SR at 12 months are no different between the pairs of groups.

The adequacy of the logistic regression models for the primary endpoint will be assessed by examining the following statistics and relevant graphical summaries:

Pearson Residuals/Deviance (Half-normal plots)

Leverage values.

Cook's Distance.

Cross validation probabilities (the probability of a particular observation, conditional on the remaining observations).

L-statistics (the influence of an observation on the difference in deviance due to fitting an the treatment effect).

Actual percentage time in AF across the 4 days of monitoring at baseline and at 12 months, i.e. the percentage of time that the patient is in AF if their predominant rhythm is SR, is reported by treatment group. Based on the interim report, the percentage time

in AF is almost dichotomized at 0% and 100%. Therefore, only summary statistics are reported in this case.

5.4.1 Sensitivity Analysis

For the final primary endpoint (Quality-adjusted Survival), we will impute as zero any EQ5D utility value that is missing due to the patient death. Any remaining missing values will be imputed using multiple imputation. For the sensitivity analysis, a Last Observation Carried Forward approach will be used to impute missing (nonzero) utility values at 24 months, and any missing intermediate utility values. Alternative imputation techniques will be considered. We will additionally consider a sensitivity analysis using SF6D-derived utility values and other valuation methods as appropriate.

If missing, the intermediate primary endpoint (heart rhythm) will be estimated by multiple imputation technique. If more than 5% of patients are withdrawn or lost to follow-up before the 12 months is reported, then the following methods will be used as sensitivity analyses to estimate the primary endpoint:

1. a `death=AF, censored=AF' strategy: If a patient is withdrawn or lost to follow-up within the 12 months, he will be assumed to be in AF.

2. a `death=AF, censored=OMIT' strategy: If a patient dies for any cause within 12 months, he will be assumed to be in AF at 12 months. If a patient is withdrawn or lost to follow-up but is alive within 12 months, his record will not be included into the sensitivity analysis.

5.5 Subgroup Analysis

Subgroup analysis will include those patients for whom measurements are available.

The first objective of subgroup analysis is within the whole dataset, to account for potential variation in the treatment effects between

patients with paroxysmal and non-paroxysmal AF: Paroxysmal AF vs. Persistent Chronic or longstanding AF.

individual centres (as a random effect to allow for heterogeneities in small centres

different cardiac surgical procedures

different surgeons.

different lesion sets.

The different lesion sets are to be defined as follows:

minimal LA lesion set: pulmonary vein isolation only ± LA appendage line

more extensive LA lesion set excluding mitral annulus

more extensive LA only lesion set including mitral annulus

minimal LA lesion set + RA lesion set

more extensive LA lesion set excluding mitral annulus + RA lesion set

more extensive LA lesion set including mitral annulus + RA lesion set

In the event that patients are too sparsely-distributed across the 6 categories, the categories will be combined into 4, by combining (i) with (ii) and combining (iv) with (v). If this is still too sparse to facilitate comparison, then the lesion set subgroup will be reduced to a comparison of category (vi) to all other groups.

For the Quality-Adjusted Survival Endpoint, a linear regression model will be fitted to the Area Under the Utility Curve, with baseline EQ5D score, surgeon, surgical procedure, treatment group, subgroup variable and the subgroup-by-treatment interaction variable. The treatment modifying effect will be reported with a 95% confidence interval. The Area under the utility curve will be appropriately transformed prior to performing subgroup analyses, and the results back-transformed where necessary.

For the Return to SR Endpoint, a logistic regression model will be fitted to heart rhythm at 12 months with the baseline heart rhythm, surgeon, surgical procedure, treatment group, subgroup variable of interest and its interaction term with treatment group. The odds ratio of the interaction term between treatment and subgroup variable will be reported with 95% confidence interval and p-value for the data seen, under the null hypothesis that there is no difference in the treatment effects on patients within different subgroup of interest.

The second objective is within the Routine+Maze treatment group, to account for variation in the treatment effects between

different ablation devices

completeness of lesion sets – both on a continuous scale and categorised as 0-4, 5-9, 10+.

In each subgroup analysis listed above, the regression model will include only patients in the Routine+Maze treatment group. The interaction effect will then be tested in the same manner as for the previous interaction effects. The odds ratio or parameter estimate of the subgroup variable will be reported with 95% confidence interval and p-value for the data seen, under the null hypothesis that in the Routine+Maze treatment group, there is no difference in the treatment effects on patients within different subgroup of interest.

5.6 Key Secondary Endpoint Analysis

The key secondary endpoint, the rate to return stable SR at 24 months will be analysed in a similar way to the rate of return to SR at 12 month, but will include only those patients for whose measurements are available. The rate in SR at 24 months for Routine treatment group will be compared to that for Routine+Maze treatment group using a binary logistic regression model including baseline heart rhythm, surgeon, surgical procedure and treatment group. The odds ratio for the rate of return to SR at 24 months of Routine treatment group relative to the Routine+Maze group will be reported with 95% confidence interval and p-value for the data seen, under the null hypothesis that the rates of return to SR at 24 months are no different between the pairs of groups.

5.7 SF36

The SF-36 consists of eight scaled scores (vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social functioning and mental health), which are the weighted sums of the questions in their section. Each scale is directly transformed into a 0-100 scale and a higher score represents a better health for that domain. Standardised physical and mental health scores are then calculated which for a general UK population are expected to be approximately normally distributed with mean 50 and standard deviation 10. [20]

The SF-36 questionnaire will be administered at baseline, 6 months, 12 months and 24 months. Then the scores for the summary measures of SF36, (Physical score and mental health score) will be given at each timepoint.

SF-36 component summary scores will be analysed using a linear regression model, adjusting for time point, treatment group, time by treatment group interaction, baseline SF-36 scores (all modelled as fixed effects) and allowing random intercepts for patients.

5.8 Additional Secondary Endpoint Analyses

To investigate whether the adjunct maze procedure decreases thromboembolic neurological complications (e.g. stroke), the patients who have suffered a stroke will be summarized within 12 months of surgery and the overall proportion of stroke events will be calculated by treatment group, using the total number of patients participating in the trial as the denominator. The absolute differences between the proportions for Routine treatment group and Routine+Maze group will be tested by Fisher's exact test, and reported along with 95% confidence intervals for differences in proportions.

The number of recruited patients who use anti-arrhythmic drugs will be tabulated by time points (at baseline, discharge, 6 weeks, 6 months, 12 months and 24 months) and drug categories (Sotalol, Amiadarone and Flecainade). Logistic Regression for the outcome of each patient (1=have one or more drugs during time period t, 0=have no drug during time period t) will be fitted, including drug category, time period using drug, baseline drug usage and treatment group as independent variables. The odds ratio for using anti-arrhythmic drug in the Routine treatment group relative to the Routine+Maze group will be reported with 95% confidence interval and p-value for the data seen, under the null hypothesis that the usage of anti-arrhythmic drugs is no different between the groups.

In the similar way, the number of recruited patients who use anti-coagulant drugs will be tabulated by time point (at baseline, discharge, 6 weeks, 6 months, 12 months and 24 months) and drug categories (Warfarin, Sinthrome and other anticoagulants). The logistic regression for the outcome of each patient (1=have one or more drugs during time period t, 0=have no drug during time period t) will be fitted, including the drug category, time period using drug, baseline drug usage and treatment group. The odds ratio of using anti-coagulant drug in the Routine treatment group relative to the Routine+Maze group will be reported with 95% confidence interval and p-value for the data seen, under the null hypothesis that the usage of anticoagulants drug is no different between the pairs of groups.

The occurrence of atrial flutter and atrial tachycardia ("organised atrial arrhythmia" - OAA) and junctional rhythm (JR) will be summarised by arm. An exploratory analysis will look at the relation between the completeness of the lesion set and the occurrences of OAA and JR.

5.9 Safety Analysis

A listing of total number of adverse events in each category as well as the deaths from any cause will be presented, and summarised by treatment group, corresponding to the intervention received. Events will be summarised according to whether they meet the

criteria of Serious Adverse Events, and whether they are thought to be related to the procedure.

Adverse Events are not planned to be categorised. However, a number of pre-specified Adverse Event categories have been specified.

6 Economic Analysis

Health Economic Analysis will be performed by the Health Economic Analysis team. A separate analysis plan has been written, and should be referred to for a description of the planned analyses.

References

[1] F. R. Hobbs, D. Fitzmaurice, J. Mant, E. Murray, S. Jowett, S. Bryan, J. Raftery, M. Davies, and G. Lip, “A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. the safe study,” Health Technology Assessment, vol. 9, no. 40, p. 93pp, 2005.

[2] “Atrial fibrillation: the management of atrial fibrillation,” NICE clinical guideline, no. 36, pp. 1–48, 2006.

[3] “Atrial fibrillation: the management of atrial fibrillation. costing report. implementing nice guidance in england,” NICE clinical guideline, no. 36, pp. 1–39, 2006.

[4] J. Dunning, T. Treasure, M. Versteegh, and S. A. Nashef, “Guidelines on the prevention and management of de novo atrial fibrillation after cardiac and thoracic surgery,” European journal of cardio-thoracic surgery, vol. 30, no. 6, pp. 852–872, 2006.

[5] E. Crystal, M. S. Garfinkle, S. Connolly, T. Ginger, K. Sleik, and S. Yusuf, “Interventions for preventing post-operative atrial fibrillation in patients undergoing heart surgery,” Cochrane Database Syst Rev, vol. 4, 2004.

[6] H. Calkins, J. Brugada, D. L. Packer, R. Cappato, S.A. Chen, H. J. Crijns, R. J. Damiano, D. W. Davies, D. E. Haines, M. Haissaguerre, et al., “Hrs/ehra/ecas expert consensus statement on catheter and surgical ablation of atrial fibrillation: Recommendations for personnel, policy, procedures and follow-up a report of the heart rhythm society (hrs) task force on catheter and surgical ablation of atrial fibrillation developed in partnership with the European heart rhythm association (ehra) and the European cardiac arrhythmia society (ecas); in collaboration with the American college of cardiology (acc), American heart association (aha), and the society of thoracic surgeons (sts). endorsed and approved by the governing bodies of the American college of cardiology, the American heart association, the European cardiac arrhythmia society, the European heart rhythm association, the society of thoracic surgeons, and the heart rhythm society.,” Europace, vol. 9, no. 6, pp. 335–379, 2007.

[7] T. Deneke, K. Khargi, P. Grewe, A. Laczkovics, S. Von Dryander, T. Lawo, K.M. Müller, and B. Lemke, “Efficacy of an additional maze procedure using cooled-tip radiofrequency ablation in patients with chronic atrial fibrillation and mitral valve disease. a randomized, prospective trial,” European Heart Journal, vol. 23, no. 7, pp. 558–566, 2002.

[8] A. Schuetz, C. Schulze, K. Sarvanakis, H. Mair, H. Plazer, E. Kilger, B. Reichart, and S. Wildhirt, “Surgical treatment of permanent atrial fibrillation using microwave energy ablation: a prospective randomized clinical trial,” European Journal of Cardio-Thoracic Surgery, vol. 24, no. 4, pp. 475–480, 2003.

[9] G. Doukas, N. J. Samani, C. Alexiou, M. Oc, D. T. Chin, P. G. Stafford, L. L. Ng, and T. J. Spyt, “Left atrial radiofrequency ablation during mitral valve surgery for continuous atrial fibrillation: a randomized controlled trial,” Jama, vol. 294, no. 18, pp. 2323–2329, 2005.

[10] C. A. Abreu Filho, L. A. Lisboa, L. A. Dallan, G. S. Spina, M. Grinberg, M. Scanavacca, E. A. Sosa, J. A. F. Ramires, and S. A. Oliveira, “Effectiveness of the maze procedure using cooled-tip radiofrequency ablation in patients with permanent atrial fibrillation and rheumatic mitral valve

disease,” Circulation, vol. 112, no. 9 suppl, pp. I–20, 2005.

[11] B. Akpinar, I. Sanisoglu, M. Guden, E. Sagbas, B. Caynak, and Z. Bayramoglu, “Combined off-pump coronary artery bypass grafting surgery and ablative therapy for atrial fibrillation: early and mid-term results,” The Annals of thoracic surgery, vol. 81, no. 4, pp. 1332–1337, 2006.

[12] N. Ad, D. C. Cheng, J. Martin, E. E. Berglin, B.C. Chang, G. Doukas, J. S. Gammie, T. Nitta, R. K. Wolf, and J. D. Puskas, “Surgical ablation for atrial fibrillation in cardiac surgery: a consensus statement of the international society of minimally invasive cardiothoracic surgery (ismics) 2009,” Innovations: Technology and Techniques in Cardiothoracic and Vascular Surgery, vol. 5, no. 2, pp. 74–83, 2010.

[13] R. J. Shemin, J. L. Cox, A. M. Gillinov, E. H. Blackstone, and C. R. Bridges, “Guidelines for reporting data and outcomes for the surgical treatment of atrial fibrillation,” The Annals of thoracic surgery, vol. 83, no. 3, pp. 1225–1230, 2007.

[14] V. Fuster, L. E. Rydén, D. S. Cannom, H. J. Crijns, A. B. Curtis, K. A. Ellenbogen, J. L. Halperin, J.Y. Le Heuzey, G. N. Kay, J. E. Lowe, et al., “Acc/aha/esc 2006 guidelines for the management of patients with atrial fibrillation: full text a report of the American college of cardiology/American heart association task force on practice guidelines and the European society

of cardiology committee for practice guidelines (writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation) developed in collaboration with the European heart rhythm association and the heart rhythm society,” Europace, vol. 8, no. 9, pp. 651–745, 2006.

[15] L. Sharples, V. Hughes, A. Crean, M. Dyer, M. Buxton, K. Goldsmith, and D. Stone, Cost-effectiveness of functional cardiac testing in the diagnosis and management of coronary artery disease: a randomised controlled trial. The CECaT trial. Gray Pub., 2007.

[16] R. G. Newcombe, “Interval estimation for the difference between independent proportions: comparison of eleven methods,” Statistics in medicine, vol. 17, no. 8, pp. 873–890, 1998.

[17] P. Dolan, C. Gudex, P. Kind, and A. Williams, A social tariff for EuroQol: results from a UK general population survey. Centre for Health Economics University of York, UK, 1995.

[18] A. Willan, D. Lin, R. Cook, and E. Chen, “Using inverse-weighting in cost-effectiveness analysis with censored data,” Statistical methods in medical research, vol. 11, no. 6, pp. 539–551, 2002.

[19] B. Efron and R. J. Tibshirani, An introduction to the bootstrap, vol. 57. CRC press, 1994.

[20] Brazier J E, Harper R, Jones N M, O'Cathain A, Thomas K J, Usherwood T, Westlake L. Validating The SF-36 Health Survey Questionnaire: New Outcome Measure For Primary Care. British Medical Journal 1992; 305: 160-4.SF-36