Nasal High-Flow Oxygen Therapy After Cardiac Surgery (NOTACS) Study

Royal Papworth Hospital

Statistical Analysis plan (SAP)

TRIAL FULL TITLE	Nasal high-flow Oxygen Therapy After Cardiac Surgery Study
TRIAL SHORT TITLE	NOTACS
SAP VERSION	4.0
SAP VERSION DATE	17th December 2024
TRIAL STATISTICIAN	Sarah Dawson
PROTOCOL VERSION SAP ASSOCIATED WITH	5.0
TRIAL PRINCIPAL INVESTIGATOR	Dr Andrew Klein
SAP AUTHOR(s)	Sarah Dawson

1 SAP Signatures

I give my approval for the attached SAP (V4.0) dated 17th December 2024:

Chief Investigator

Name: Dr Andrew Klein
Signature:
Date: 17-12-2024

Statistician (author)

Name: Sarah Dawson
Signature: Sarah Dawson
Date: 17-12-2024

Senior statistician (reviewer)

Name: Sofia S. Villar
Signature: Sofia Villar
Date: 17-12-2024

2 Amendments

The following amendments and / or administrative changes have been made to this SAP since the date of preparation.

Amendment Number	Date of Amendment	Version Number	Type of Amendment? (e.g. substantial / non-substantial / administrative change)
1	29/04/2020	0.1	Created
2	25/07/2020	0.2	Substantial
3	17/09/2020	0.3	Non-substantial
4	14/10/2020	0.4	Non-substantial
5	07/12/2020	0.41	Substantial
6	29/01/2021	0.42	Substantial
7	30/09/2022	0.43	Substantial
8	28/10/2022	1.0	Administrative change for approval of v0.43
9	21/12/2022	2.0	Providing a more precise definition of a treatment switch.

10	09/08/2023	3.0	i) Providing a more precise definition of how to calculate DAH score. ii) Providing further clarification on the definition of a treatment switch. iii) Increase in maximum possible sample size (from 1152 to 1280) following protocol change. iv) Update on final agreed sample size following the sample size re-estimation at the interim analysis. v) Provide clarification on how randomised screen-failures will be handled. vi) Provide details of cross-checks that will be done alongside health economics. vii) Clarify that treatment for 16 hours including a 1 hour period off of treatment (so 15+1, not 16+1) is acceptable. viii) Added NIHR to the abbreviations table. viiii) Added a table to the mock final report of summary statistics for DAH90 pre-imputation and post-imputation. x) Added day 30 and day 60 for EQ-5D-5L and Barthel index to the summary statistics listing. xi) Added an additional analysis of the sub-group of patients who required escalation of oxygen therapy (in line with the Protocol defined escalation protocol (Appendix 3)). xii) Added section on funding sources. xiii) Added comment about how we will handle “hospital in the home” in the DAH calculation. xiv) Added section on multiplicity and simplified secondary and sensitivity analyses to reduce number of tests performed. Added clarification that the p-value from multiple imputation will be reported in place of the p-value from the primary analysis if multiple imputation is performed.

11	17/12/2024	4.0	i) Amendments and further details added to the missing data section. ii) Removed reference to Hodges-Lehmann method for calculating confidence interval for primary analysis. Details for the computation of the confidence interval will be detailed and provided where they are needed. iii) Added to 9.4 Missing Data how location data obtained outside of the location diary was handled.

3 Table of Contents

4 Abbreviations and Definitions

Abbreviation	Definition
AE	Adverse Event
ARDS	Acute Respiratory Distress Syndrome
ARISCAT	A risk index for postoperative pulmonary complications
BARTHEL	A measure of performance in activities of daily living
BiPAP	Bilevel Positive Airway Pressure
BMI	Body Mass Index
CABG	Coronary Artery Bypass Graft
CNS	Central Nervous System
CONSORT	Consolidated Standards of Reporting Trials
COPD	Chronic Obstructive Pulmonary Disease
CPAP	Continuous Positive Airway Pressure
CPB	Cardiopulmonary Bypass
CRF	Case Report Form
CRN	Clinical Research Network
CTU	Clinical Trials Unit
DAH	Days Alive and at Home
DAH30	Days Alive and at Home 30 days post-operative
DAH90	Days Alive and at Home 90 days post-operative
DMEC	Data Monitoring & Ethics Committee
ECG	Electrocardiogram
EQ-5D-5L	A quality of life questionnaire
eGFR	Estimated Glomerular Filtration Rate
EUROSCORE	European System for Cardiac Operative Risk Evaluation
Fi02	Fraction of Inspired Oxygen
GP	General Practitioner
ITT	Intention to Treat
HDU	High Dependency unit
HEAP	Health Economics Analysis Plan
HFNT	High Flow Nasal Therapy
IABP	Intra-aortic Balloon Pump
ICER	Incremental cost-effectiveness
ICH	International Council for Harmonisation
ICU	Intensive Care Unit
LOS	Length of Stay
MedDRA	Medical Dictionary for Regulatory Activities
NICE	National Institute for Health and Clinical Excellence

NICOR	National Institute of Cardiovascular Outcomes Research
NIHR	National Institute for Health and Care Research
Physio	Physiotherapy
PI	Principle Investigator
PTUC	Papworth Trials Unit Collaboration
QUALYS	Quality Adjusted Life Years
R&D	Research & Development
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
ROX	Respiratory rate oxygenation
RR	Respiratory Rate
SAP	Statistical Analysis Plan
sCR	Serum Creatinine
SD	Standard Deviation
SOT	Standard Oxygen Therapy
Sp02	Peripheral Capillary Oxygenation Saturation
SUSAR	Serious Unexpected Adverse Reaction
TIA	Transient Ischaemic Attack
TSC	Trial Steering Committee
VF	Ventricular Fibrillation
VT	Ventricular Tachycardia

5 Introduction

The statistical analysis plan (SAP) should be read in conjunction with the NOTACS trial Protocol. This SAP has been developed alongside Protocol version 5.0. This SAP closely follows published guidelines for the content of SAPs in clinical trials (1).

5.1 Preface

Patients undergoing cardiac surgery are at significant risk of postoperative pulmonary complications (PPC) that may lead to prolonged intensive care unit (ICU) and hospital stay and increased mortality (2). The incidence of respiratory complications may be three to four times more common in patients with intrinsic respiratory disease and lower airway obstruction (including asthma or chronic obstructive pulmonary disease (COPD)), or obese patients or current heavy smokers (> 10 pack years) (3). These patients often develop lower respiratory tract infections, with impaired oxygenation/ventilation and prolonged requirement for ventilatory support. They are more likely to require escalation of respiratory support and readmission to ICU during recovery from surgery (4) (5) (6).

High-flow nasal therapy (HFNT) is increasingly used as a non-invasive form of respiratory support (7). It delivers low level, flow-dependent positive airway pressure, and is much better tolerated by patients than alternatives such as continuous positive airway pressure (CPAP) or non-invasive ventilation (8). Patients can talk, eat, drink and walk whilst using HFNT. However, there is equipoise regarding its prophylactic use and effect on important patient-centred outcomes, hence the rationale for this trial. Recent systematic reviews in non-cardiac (9) and cardiothoracic (10) surgery concluded that HFNT could reduce respiratory support and pulmonary complications, and could be safely administered.

5.2 Purpose of the analyses

The analyses will assess the efficacy, safety and cost-effectiveness of HFNT on the outcomes of patients after a cardiac surgery.

6 Study Objectives and Endpoints

6.1 Study Objectives

The primary objectives are to determine if prophylactic use of HFNT (for a minimum of 16 hours after tracheal extubation) is clinically- and cost-effective (in comparison with standard oxygen) up to 90 days after surgery, for adult patients undergoing

cardiac procedures with cardiopulmonary bypass who are at high risk of postoperative pulmonary complications.

The secondary objectives are to determine if prophylactic use of HFNT is able to:

- Reduce mortality, pulmonary complications, intensive care re-admission rate, length of hospital and intensive care stay.
- Reduce incidence of major complications including sepsis, acute kidney injury, myocardial infarction and stroke.
- Reduce readmission to hospital rate.
- Improve oxygenation as measured by the ROX Index.
- Improve patient-centred outcomes as measured using the EQ-5D-5L.
- Reduce patient level of assistance needed with activities of daily living as measured using BARTHEL questionnaire.
- Improve quality of survival as measured using EQ-5D-5L Quality adjusted life years (QALYs)
- Reduce health service and resource use.

This SAP describes the statistical analyses which are planned for the NOTACS trial. A separate HEAP will describe the planned health economic analyses.

6.2 Endpoints

Primary endpoint:

- DAH90 (days alive and at home in the first 90 days after the planned surgery)

Secondary endpoints:

- DAH30 (days alive and at home in the first 30 days after the planned surgery)
- Incidence of adverse events and serious adverse events (including death)
- Incidence of stroke
- Incidence of sepsis
- Incidence of acute kidney injury
- Incidence of myocardial infarction
- Postoperative pulmonary complications
- ICU re-admission rate during index admission
- Total length of ICU stay (days) during index admission
- Total length of hospital stay (days) during index admission
- Re-admission to hospital
- Oxygenation, as measured by ROX Index

- Patient-reported outcomes (EQ- 5D-5L)
- Patient level of assistance needed with Activities of Daily Living (BARTHEL questionnaire)
- Quality of survival (QALYs)
- Health service and resource use

For definition of secondary outcomes see Appendix 1 of the trial Protocol (version 5.0).

Exploratory endpoints:

- Demographic variables, including residence at baseline
- Discharge location following index admission
- Number, duration (days) and location type of periods spent away from baseline residence within 90 days of randomisation
- Time (hours) on randomised therapy
- Compliance

7 Study Methods

7.1 General Study Design and Plan

The trial is an adaptive, multicentre, parallel group, randomised controlled clinical trial with embedded cost-effectiveness analysis, comparing the use of high flow nasal therapy (HFNT) to standard oxygen therapy for a minimum of 16 hours after tracheal extubation, in patients at high risk of respiratory complications following cardiac surgery. Patients will be recruited over 3 years across at least 10 centres in the UK, 8 centres in Australia and 1 centre in New Zealand.

The schedule of data collection is outlined in Table 1.

Table 1 Schedule of events

Visit Number	**Visit 1 Screening**	**Visit 2 Baseline**	**Visit 3 Randomisation**	**Visit 4 Discharge**	**Visit 5 30 Days (+30 days) Post-op**	**Visit 6 90 Days (+90 days) Post-op**
Time Interval of Visit	Prior to Surgical Admission (or after admission if in-house urgent)	Surgery Admission	During or After surgery & prior to extubation	Day of Discharge	30 days (+30 days) Post-op	90 days (+90 days) Post-op
Activity						
Inclusion/Exclusion Criteria	**X**					
Informed Consent		**X**				
Demographics		**X**				
Past Medical History		**X**				
EuroSCORE II & ARISCAT Risk Assessments		**X**				
EQ-5D-5L & BARTHEL Questionnaires		**X**		**X**	**X**	**X**
Participant & Family Resource Use Questionnaires		**X**		**X**	**X**	**X**
Adverse & Serious Adverse Events Assessed (from the point of extubation			**X----------X**			
Inpatient Medication Log (to start from the point of extubation)			**X----------X**			
Inpatient Location Log (to start from the point of extubation)			**X----------X**			
Inpatient Oxygen Therapy Log (to start from the point of extubation)			**X----------X**			
Participant Location and Medication Diary				**X----------X**		
Randomisation/ Initiation of HFNT or Standard Oxygen Therapy			**X**			
ROX Index			**X----------X**			
Record of Respiratory Support Escalation			**X----------X**			
Record of Post-operative Complications			**X----------X**			
Record of Intensive care Length of stay and Re-admissions				**X**		
Record of Hospital Discharge Destination				**X**		
Record of Hospital Length of Stay				**X**		

7.2 Inclusion-Exclusion Criteria and General Study Population

Inclusion criteria:

- Aged 18 years or over.
- Undergoing any elective or urgent first-time or redo cardiac surgery performed on cardiopulmonary bypass.
- Have one or more clinical patient-related risk factor for postoperative pulmonary complications (COPD, asthma, lower respiratory tract infection in last 4 weeks, body mass index $\geq$35 kg/m^2, current (within the last 6 weeks) heavy smokers (> 10 pack years).

Exclusion criteria:

- Requiring home oxygen therapy.
- Deep hypothermic circulatory arrest planned.
- Contraindication to HFNT, e.g. nasal septal defect.
- Requirement for home respiratory support (including: HFNT, CPAP, BiPAP).
- Requiring emergency cardiac surgery defined as surgery required within 24 hours of the decision to operate.
- Patients not fluent in English.

7.3 Randomisation and Blinding

To reduce predictability in the randomisation sequence, stratified random permuted block randomisation will be used to randomise patients with random block sizes of 4 or 6, stratified by centre. The allocation ratio of standard oxygen therapy to HFNT is 1:1. The randomisation service will be hosted by Sealed Envelope.

Due to the nature of the intervention, clinical staff in ICU and on the wards cannot be blinded whilst the patient is receiving randomised therapy. However, a team of research staff at the central clinical trials unit will collect data on outcomes and these staff will be blinded. In addition, the decision to discharge patients from hospital, which affects the primary outcome, will be made by clinicians who are independent of the research team at each site, according to standard protocols. The interim analysis and sample size re-estimation will be done by an independent unblinded statistician so that the trial statisticians can remain blinded until the final analysis in order to preserve type I error rate at 5%.

7.4 Study Variables

The schedule of events given in Table 1outlines the key groups of study variables that will be collected during the NOTACS trial.

Further detail at the level of individual variables, including variable type, has been documented separately by the data management team.

8 Sample Size

The original sample size calculation for the NOTACS trial was based on the primary end-point DAH90. The cost-effectiveness analysis was not considered as part of this calculation.

The minimum target sample size is 850 randomised participants. The adaptive design includes a sample size re-estimation after a pre-specified interim analysis. The sample size will be revisited after a minimum of 300 patients have been randomised and followed up to 90 day and will allow for a maximum sample size increase to 1280 patients (note that this is an increase from the original planned maximum sample size of 1152 in SAP V2.0, see section 9.5.4 for further details). This will provide protection against deviations from original assumptions, which were based on limited data.

The primary endpoint (DAH90) typically has a left-skewed bi-modal distribution with a small spike at 0 due to in-hospital deaths. Following the approach of Myles et al (11), patients who die within 90 days of surgery will be assigned a zero DAH score irrespective of whether they spent any time at home during the 90 day follow-up period. This definition is deemed appropriate on the basis that the death rate in the trial population is expected to be low (around 3%, based on pilot data (12) and registry data (13)), most deaths are expected to occur within the initial hospital admission (within a short time of surgery which corresponds to a DAH90 close to 0), the death rate is expected to be comparable between the two treatment arms, and it is not expected that the treatment will impact on death rate. The required sample size was obtained by simulations (100,000 replicates) by first generating length of stay (LOS) using a lognormal distribution. Simulations were used because the endpoint does not have a standard known distribution. Based on the information provided by collaborating hospitals, the parameters of the lognormal distributions in both arms were derived through a pooled weighted average. The variability was calibrated to SD =12.85 in the control arm and SD=3.20 in the treatment arm. The median LOS in the control arm was set to 8 days. We assumed a 3% death rate (based on pilot data (12)

and registry data (13)), and following the approach of Myles (11) we treated any death within the 90 day follow-up period as scoring 90 for LOS regardless of when the death occurred. LOS was truncated at 90 days (the maximum for our follow-up period). Finally DAH90 was computed as 90 minus LOS. The resulting data are bimodal with a spike at 0, as seen with observed data of this type.

A total sample size of 310 has 90% power to detect an increase of 2 days in the median DAH90 using the Mann-Whitney-Wilcoxon test for the analysis. After adjustment for 12% crossover from standard oxygen to HFNT and 25% crossover from HFNT to standard oxygen as well as an extra 5% loss to follow up (equally distributed among arms), the total sample size needed to detect a 2-day increase in DAH90 with 90% power with an intention to treat analysis is 850 patients. Therefore, in the first instance the trial aims to recruit 850 patients.

9 General Considerations

9.1 Timing of Analyses

9.1.1 Interim Analysis

The interim analysis will be performed after a minimum of 300 patients have been randomised and followed-up for 90 days. We expect this analysis will be performed around late 2022, dependent on the rate of recruitment and completion rate of the variable DAH90. If there is greater than 15% missing data for DAH90 (excluding missingness because of death, which is informative missingness) then the interim analysis may be delayed.

The full specification of the variables to be included in the interim analysis is shown in the accompanying mock interim analysis report.

9.1.2 Final Analysis

The final analysis will be implemented after data cleaning and following database lock.

The full specification of the variables to be included in the final analysis is shown in the accompanying mock final analysis report.

9.2 Analysis populations

Each patient will be included or excluded from each of the analysis populations defined below. Where possible this will be carried out prior to unblinding to avoid bias.

9.2.1 Full population

The full population includes all patients recruited to the trial.

In the case of constructing a CONSORT diagram for the trial, in this instance only, the full population will also include those screened for the trial (whether eligible or not).

9.2.2 Safety Population

The safety population includes all subjects entered into the trial from the time of tracheal extubation up to 90 days after surgery (the period for which safety data is being collected). The safety population will be used to provide summary statistics on AEs and SAEs, which will be reported by the treatment arm assigned during randomisation (or by the treatment arm actually received if it is different to that assigned during randomisation). The safety population will be examined at both the interim and final analyses.

For the interim analysis, the safety population will include all subjects entered into the trial from the time of tracheal extubation up to 90 days after surgery (the period for which safety data is being collected) at the time of database lock for the interim analysis, regardless of whether they are formally included in the interim analysis population (outlined in section 9.2.3). This may include patients who were recruited after the 300th patient but who have not yet completed follow-up to 90 days, as well as any patients who have withdrawn from the trial provided they continue to consent to their data being used.

For the final analysis, the safety population will include all subjects entered into the trial from the time of tracheal extubation up to 90 days after surgery (the period for which safety data is being collected) at the time of database lock for the final analysis. This may include patients who have withdrawn from the trial provided they continue to consent to their data being used.

9.2.3 Interim Analysis Population

The interim analysis population will include all patients who have been randomised at the time 300 patients have completed 90 days follow-up. If there is greater than 15% missing data for DAH90 (excluding missingness because of death, which is informative missingness) and the interim analysis is delayed because of that, the interim analysis population may include more than 300 patients.

The trial will recruit from the UK, Australia and New Zealand. It is expected that the majority of the patients recruited by the end of the trial will be from UK centres, with the remaining patients recruited from Australia and New Zealand. On this basis, to ensure that the results of the interim sample size re-estimation reflect the expected proportions of UK, Australian and New Zealand recruits seen at the end of the trial, we will seek to ensure that between 50%-75% of the patients included in the interim analysis are recruited from the UK, with the remainder from Australia and New Zealand. This will ensure that even if differences exist between the UK, Australia and New Zealand, the effect of those differences will not disproportionately bias the interim sample size re-estimation and therefore will not unduly affect the power at the end of the trial.

The interim analysis population is likely to include some patients with partial data (e.g. anyone recruited after the 300th patient who hasn't yet completed their follow up, and anyone recruited before the 300th patient who left the trial before completing follow up).

9.2.4 Intention-to-treat Population

The intention-to-treat (ITT) population is the population that will be used for the majority of the analyses, including the analysis of the primary endpoint and the interim sample size re-estimation. The ITT population includes all subjects who were randomised, regardless of whether they received the treatment randomly allocated to them or completed follow-up. The data will be analysed assuming that the patient received the treatment arm they were randomly allocated to. If a patient dies after being randomised but before extubation occurs, they will be included in the ITT analysis population with a days at home score of zero. If a patient is randomised but subsequently found to be ineligible for the trial after being randomised, they will be included in the ITT analysis population.

9.2.5 Per-Protocol Population

The per-protocol population includes all subjects who adhered to the trial Protocol by receiving the treatment randomly allocated to them for a minimum of 16 hours, no matter whether they completed all of their follow-ups. However, for the purpose of determining treatment compliance, during the 16 hours, up to a total of one hour off treatment is allowed for any required transfers around the hospital and/or physio mobilisation. Any subject who did not receive the treatment randomly allocated to them for at least 16 hours (allowing up to a total of 1 hour off treatment for transfers around the hospital and/or physio mobilisation) will be excluded from the per-protocol population.

Patients who were randomised but subsequently found to be ineligible for the trial after being randomised will not be included in the per-protocol population.

9.2.6 Time-on-Treatment Populations

Two time-on-treatment populations will be defined in a similar way to the per-protocol population, but with alternative minimum treatment periods. These will include all subjects who received the treatment they were randomly allocated to for:

A. At least 8 hours
B. At least 24 hours

For the purpose of determining the time-on-treatment populations, patients will be considered as compliant provided that they receive either 8 hours (time-on-treatment population A) or 24 hours (time-on-treatment population B) of randomised therapy, with a total of one hour off treatment allowed for any required transfers around the hospital and/or physio mobilisation.

9.3 Covariates and Subgroups

Important clinical covariates and sub-groups will be included in secondary analyses. Important covariates of interest include:

- DAH90
- DAH30
- ARISCAT score
- EURROSCORE II

- Gender
- COPD
- Asthma
- Obesity (BMI>35kg/m^2)
- Current smoking status
- Lower respiratory tract infection in last 4 weeks
- Age (≤ or > 80 years)
- First time or re-do surgery
- ROX index
- Extubation timing (≤ or >24 hours after admission to ICU)
- Return to theatre (≤ or > 24 hours of admission to ICU)
- Length of initial ICU stay

Sub-groups of interest include:

- Country
- Centre (UK sites only)
- Patients who require escalation of oxygen therapy (in line with the Protocol defined escalation protocol (Appendix 3))

9.4 Missing Data

Withdrawal rates will be summarised by treatment arm, and the timing of and reason for withdrawals will be reported.

The proportion of missing data will be quantified by treatment group for the variables included in the primary and secondary analyses. Based on the pilot study, the rate of missing data in the primary end-point (DAH90) is expected to be low. Variables with >25% missing data may not be used in statistical regression modelling but will be summarised and reported where appropriate.

The trial was powered on the basis of a 5% loss to follow up rate. Therefore, provided there is no more than 5% missing data in the primary endpoint (DAH90) and no obvious cause for concern over the pattern of missing data, no further investigations into missing data will be carried out. In this scenario we will ignore missing observations and run complete-case analysis.

If the missingness rate in the primary endpoint (DAH90) exceeds 5%, we will take further steps to investigate the type of missingness and how it relates to other variables, such as demographic variables and secondary endpoints. We will create an

indicator variable for DAH90, to show the presence or absence of missing data. We will use logistic regression to investigate whether the missing indicator of DAH90 can be predicted by the other variables. We will also use Little's test (14) for the continuous variables as a further way to assess whether the missingness mechanism is likely to be missing completely at random (MCAR). If the missing indicator for DAH90 can not be predicted by the other variables and if the Little's tests are also non-significant (i.e. $p>0.05$) we will assume that the missingness mechanism for these variables is MCAR. Otherwise we will assume that the missingness mechanism is MAR. However it is important to acknowledge that the trial is not powered for these missing data analyses, so a non-significant result may be due to a lack of power rather than a lack of effect.

If the missing data mechanism is found to be MCAR in all of the missing data mechanism tests, we will ignore missing observations and run complete-case analysis, even if there is greater than 5% missing data for DAH90.

If the missing data mechanism is found not to be MCAR in any of the missing data mechanism tests, as a sensitivity analysis to evaluate the robustness of the primary analysis we will assume the missing data mechanism is MAR and impute the missing data by using multiple imputation by chained equations (15).

DAH90 can be broken into two components:

- y0: the initial hospital stay
- y1: the location diary

As y0 is recorded by hospitals and y1 is recorded by participants, we expect that y0 will be more frequently observed than y1. Further, auxiliary variables (which help impute missing values) for y0 and y1 may be different. Therefore, we will impute y0 and y1 using multiple imputation by chained equations, where we will include relevant auxiliary variables for y0 and y1. These auxiliary variables are those that are good predictors of the variable itself as well as its missingness indicator. It is possible that the auxiliary variables used for the imputation of y0 may differ to those used for the imputation of y1.

A random number seed will be set and recorded within the analysis code for reproducibility purposes. We will impute $m=25$ datasets, however this number may be increased if needed to reduce Monte Carlo variability. For each of the m imputed datasets, ten imputation cycles will be used (16). If convergence is not achieved with ten cycles then the number of cycles may be increased. The imputed values from the last cycle in each imputed data set will be retained and these will form our m imputed data sets. In the imputation model we will use predictive mean matching. The primary

analysis (Mann-Whitney-Wilcoxon test of DAH90 by treatment group) will be repeated for each of the *m* imputed data sets, and the median p-value across these *m* individual analyses will be calculated (17). The median p-value will be provided as the result of this sensitivity analysis.

For derived variables such as DAH90 and DAH30, the presence of missing data in the variables used to calculate these derived variables from may be more difficult to identify. DAH90 and DAH30 rely on participant location diaries to capture all changes in participant location with accurate start and end dates for each change in location. There may be instances where data is missing from participant location diaries despite it appearing that the participant location diary has been completed in full, if for instances a participant fails to record a change of location. This is a potential limitation in this type of end-point. We had initially intended that, where it is clear that a patient diary is incomplete, they will be completed where possible by calling the patient's GP surgery and using hospital discharge summaries for dates of change of living location. However, during the course of data collection, it was noted that additional location information was provided by some patients outside of their location diary (e.g. in the resource questionnaire or during conversation with the follow-up team) that was not in the diary and could not be confirmed by the GP. In some cases, corrections were made to the location diary if there was sufficient information to do so, at the discretion of the chief investigator. When reporting the final primary analysis, and in collaboration with the data team, we will aim to provide our best estimate of the percentage of cases in which this could have happened. Further information regarding death or any additional hospital admission will be collected from the discharge summary and will be cross-referenced with GP records.

As the main focus of the missing data analysis is on the robustness of the primary analysis, if the proportion of missing data for secondary analysis variables is more than 5% but the proportion of missing data for the primary endpoint is less than 5%, no further investigations into the missing data mechanism or imputation will be performed.

The way missing data is handled may be different for health economic purposes, as imputation may be needed even if the proportion of missing data is low to avoid exclusion of patients that partially contribute data. This will be detailed separately in the health economics analysis plan (HEAP).

9.5 Interim Analysis

Several tasks will be performed at the interim-analysis, which can be broken down into three main areas:

- sample size re-estimation
- assessment of safety
- assessment of recruitment, compliance and data completeness

In the event that the DMEC request additional data analyses at the interim analysis stage the trial statistician will be responsible for providing these (via an independent statistician if unblinding is required).

9.5.1 Sample size re-estimation

The assumptions used for the original sample size calculation were based on pilot data (12), registry data (13) and data provided by the largest participating centres. NOTACS is a multicentre trial, using a different primary endpoint to the pilot data. We also found that the sample size calculation was very sensitive to the standard deviation, level of treatment switches and loss to follow up assumed. Therefore NOTACS has been designed as an adaptive trial with an interim sample size re-estimation planned after a minimum of 300 patients complete 90 days post-randomisation follow-up. This sample size adaptation may prevent an underpowered trial if moderate deviations from the assumptions made for the initial sample size calculation are observed.

At the interim sample size re-estimation, the data accumulated so far will be used to re-estimate a number of “nuisance” parameters including:

- standard deviation of DAH90 in the standard-oxygen therapy arm
- standard deviation of DAH90 in the HFNT arm
- treatment switch rate from standard oxygen therapy to HFNT
- treatment switch rate from HFNT to standard oxygen therapy
- overall drop-out rate
- overall death rate

Treatment efficacy will <u>not</u> be assessed at the interim analysis.

A treatment switch occurs when a patient does not complete 16 hours on randomised therapy and spends some time on the non-randomised therapy.

The following are not considered to be treatment switches:

1. Completes 16 hours on randomised therapy and no time on non-randomised therapy.
2. Completes 16 hours on randomised therapy and switches after that.
3. For patients randomised to HFNT, completes at least 15 hours on randomised therapy, with up to a total of one hour on standard oxygen therapy to allow for any required transfers around the hospital and/or physio mobilisation (logistical due to lack of battery pack on the HFNT devices).
4. Does not complete 16 hours on randomised therapy and spends no time on the non-randomised therapy. This would also be a treatment non-compliance.

The trial statistician will prepare the analysis code using a dummy randomisation list to calculate the "nuisance" parameters listed above. An independent statistician will then run the code with the real randomisation list provided by the data manager and output the updated "nuisance" parameter estimates. The original sample size calculation will be repeated with these updated estimates to provide an updated sample size estimate which will be reported to the DMEC.

After the interim analysis, the sample size of the trial will be updated with a maximum increase up to 1280 patients. There are several possible outcomes from the sample size re-estimation which are summarised in Table 2.

Table 2. Recommended sample size from interim sample size re-estimation and course of action

Recommended sample size from interim sample size re-estimation	Course of action
≤850	Continue recruitment to 850 patients
851-1280	Continue recruitment to the new recommended sample size
>1280	Continue recruitment to 1280 patients

The DMEC have the responsibility of agreeing the updated sample size following the interim analysis. The DMEC should report its recommendations to the Chair of the TSC. In exceptional circumstances the DMEC may recommend a course of action not listed in Table 2. Any recommendation outside of those listed in Table 2 should be clearly justified.

The sample size re-estimation will be done using an independent statistician to allow the trial statisticians to remain blinded, in order to preserve the type 1 error rate at 5%. Following the interim analysis and review by the DMEC, the updated sample size recommendation will be shared with the trial team. To maintain blinding, no other

results from the interim analysis will be shared with the trial team (including the trial statistician) until after the final database lock.

Sensitivity analysis will also evaluate the impact of accounting for all days alive and at home (i.e. even if patients die before 90 days) on sample size estimation. This will facilitate discussion of the impact of this assumption on the final estimation of DAH90 and QALYs. These analyses will inform how final analyses for the effectiveness and health economics can be aligned in terms of the primary endpoint definition and used to better address the co-primary questions.

9.5.2 Safety

Another important task at the interim analysis is to present a summary of the safety data so that the DMEC can evaluate if there is evidence of treatment safety or treatment harm. The safety data will be listed by MedDRA Preferred Term and further grouped by System Organ Class. The frequencies of the AEs and SAEs will be summarised by treatment group.

Safety will be summarised at both the interim and the final analysis.

9.5.3 Recruitment, compliance and data completeness

Information about recruitment, treatment compliance, and data completeness will be reported at the interim analysis. A summary of patient recruitment data will be presented by centre, as well as by treatment group and time or trial stage where appropriate. Compliance and data completeness will be summarised by treatment group.

Recruitment and treatment compliance will also be summarised in the final analysis.

9.5.4 Sample size update following pre-planned interim analysis

Following SAP V2.0 section 9.5, a pre-planned interim analysis was performed in December 2022, following database soft lock on 1st December 2022. The purpose of the interim analysis was to perform a sample size re-estimation and make assessments around safety, recruitment, compliance and data completeness.

The interim sample size re-estimation was performed after 300 patients had completed 90 days post-randomisation follow-up. During the analyses process it was

confirmed that 241 of the 300 patients included had complete 90-day follow-up data at the time of the interim analysis. The interim sample size re-estimation was performed by an independent statistician to allow the trial statisticians to remain blinded, in order to preserve the type 1 error rate at 5%. This sample size adaptation was pre-planned as part of the adaptive design, to prevent an underpowered trial if moderate deviations from the assumptions made for the initial sample size calculation were observed.

Based on the results of the interim sample size re-estimation, the recommendation of the Data Monitoring and Ethics Committee (DMEC) was to increase the maximum sample size to 1280. Although this is greater than the original planned maximum sample size of 1152, it was agreed that 1280 was feasible with an extension to the recruitment period, and would result in a more robust set of results for this confirmatory trial, maintaining the power at 90%. Therefore, the final sample size has been increased from the original minimum of 850 to 1280 patients. The proposal to increase the sample size to 1280 and an 18 month project extension was agreed by the NIHR HTA in June 2023.

9.6 Data checks

Data checks will follow the recommendations in Kirkwood and Sterne (2003, Ch.38). Outliers in continuous variables will be detected using ranges and plotting distributions by treatment group. Categorical variables will be tabulated by treatment group and unexpected distributions or data points will be further checked. Consistency checks between two or more variables will also be performed, e.g. plotting weight against height. Clear outliers, e.g. $>$ 4 standard deviations from the centre of distribution will be investigated further, and may be removed and regarded as a missing value. In these instances their effect will be investigated through sensitivity analyses. However, a series of potential outliers occurring in a skewed distribution may be considered to hold, and in these cases, where necessary, the data may be transformed by nonlinear function (such as log or squared root).

Some key outcome variables will be independently checked by both health economics and statistics prior to database hard lock and unblinding to ensure that we are handling the data in a comparable way, hence these will be done using pooled data rather than by treatment arm. We will seek to ensure that the values are identical to the fourth decimal place. These key variables will include:

- Mean, standard deviation, median, minimum, maximum and n for DAH90
- Percentage entries for each level within each dimension at each time point for the EQ-5D-5L questionnaires

- Mean, standard deviation, median, minimum, maximum and n for length of hospital say for the index admission

These checks will be an internal process only and will not be included in the statistical report.

9.7 Descriptive Statistics

All relevant variables will be summarised using the following descriptive statistics: for continuous variables, the non-missing sample size, mean, standard deviation, median, maximum and minimum; for categorical variables, the frequency and percentages (based on the non-missing sample size) of observed levels will be reported. In general, data will be listed by treatment group where appropriate. Summary tables will be structured with a column or row for each treatment group and will be annotated with the total population size, including any missing observations, as appropriate.

9.8 Derived variables

9.8.1 Days Alive and at home (DAH)

For the calculation of DAH, we consider “home” as the usual abode of the patient recorded at baseline, which may not be a home in the traditional sense. Escalation of care or time away from the usual abode during the follow-up period will count as a day away from home, as set out in Table 3.

DAH is calculated from the date of index surgery and does not include time spent in hospital prior to this. DAH will be calculated using usual abode recorded at baseline, date of surgery, date of discharge from index hospital admission, discharge location, and location diaries used to record patient location following discharge from the index hospital admission. As patient location diaries are self-reported, they may be subject to biases that are difficult to detect and control for.

Table 3 What counts as a day away from home?

		Escalation of care/time away from "usual abode" during follow-up period										
		Home without support	Home with support	Relative or friend's home without support	Relative or friend's home with support	Residential home without support	Residential home with support	Respite care with or without support⊿	Nursing home with or without support⊿	Hospital with or without support⊿	Other without support	Other with support
Baseline Usual abode	Home without support	N	Y	N	Y	Y	Y	Y	Y	Y	*	*
	Home with support	N	N	N	Y	Y	Y	Y	Y	Y	*	*
	Relative or friend's home without support	N	Y	N	Y	Y	Y	Y	Y	Y	*	*
	Relative or friend's home with support	N	N	N	N	Y	Y	Y	Y	Y	*	*
	Residential home without support	N	N	N	N	N	Y	Y	Y	Y	*	*
	Residential home with support	N	N	N	N	N	N	Y	Y	Y	*	*
	Respite care with or without support⊿	N	N	N	N	N	N	N	Y	Y	*	*
	Nursing home with or without support⊿	N	N	N	N	N	N	N	N	Y	*	*
	Hospital with or without support⊿	N	N	N	N	N	N	N	N	N	*	*
	Other without support	N	*	*	*	*	*	*	*	*	*	*
	Other with support	N	*	*	*	*	*	*	*	*	*	*

* *To be decided case-by-case.*

⊿ *There are some options that could be chosen in the eCRF which are not possible in reality (i.e., locations where support is provided by default). These have been merged (e.g., hospital with or without support)*

For patients that are alive at time T, DAH is derived as follows:

$$DAH_i^{(T)} = \begin{cases} T - \sum_{t=1}^{T} \mathrm{I}(E_{it} = \text{´}Y\text{´}) & \text{if patient } i \text{ is alive at day } T \\ 0 & \text{otherwise} \end{cases}$$

where

- T denotes the length of the period of interest since surgery in days. T equals 30 or 90 in the NOTACS study,
- I(…) denotes the indicator function which equals 1 if the condition is met and 0 otherwise,
- E_{it} denotes the binary escalation assessment of the tth day of the period of interest for patient i according to Table 3 and to the case by case assessment made when needed (with levels 'Y' for escalation and 'N' otherwise).

For example, for a patient whose usual abode at baseline is recorded as "Home without support", $\sum_{t=1}^{T} \mathrm{I}(E_{it} = \text{´}Y\text{´})$ would include days spent at:

- Home with support
- Relatives home with support
- Residential home without support
- Residential home with support
- Respite care with or without support
- Nursing home with or without support
- Hospital with or without support
- Other without support (*to be decided case-by-case*)
- Other with support (*to be decided case-by-case*)

but would not include days spent at:

- Home without support
- Relatives home without support.

For those patients whose baseline usual abode is recorded as "Other without support" or "Other with support", we will use at least two blinded members of the trial team to decide whether a subsequent change in location should be counted as a day away from "home"/escalation of care (unless moving to "Home without support"). Similarly, if a patient records "Other without support" or "Other with support" as their discharge location or in the patient location diary at any point during the 90 day follow-up period, we will use at least two blinded members of the trial team to decide whether this should be counted as a day away from "home"/escalation of care. If consensus is not reached, an additional blinded member of the trial team will be consulted.

Patients who receive "hospital in the home" during their index hospital admission will be treated as if they were in hospital for that period of time.

Patients that are not alive at time T will be imputed with a DAH of 0 (11).
Where a participant changes location (e.g. hospital to home, or home to hospital) their location for that day is based on where they spent the night. Specifically, if a patient changes location on day 30 or day 90 it will be incorporated in the DAH score based on where the patient spent the night on day 30 or 90.

9.8.2 Time spent on intervention and Compliance

Time spent on standard oxygen therapy or HFNT will be calculated for each patient using the oxygen log CRF.

The date and time started and the date and time stopped are captured within the oxygen log CRF. Changes to oxygen therapy settings are also recorded with start and stop dates and times.

In straightforward cases, the time spent on either standard oxygen therapy or HFNT can be derived simply by the modulus of the time difference in minutes between the date and time treatment was started and stopped.

For patients with multiple start and stop dates and times, or for patients where the intervention settings were changed, the time spent on intervention can be derived by taking the modulus of the time difference in minutes between the date and time treatment was started and stopped for each individual entry on the oxygen log CRF, and summing across these for each patient to get a total time in minutes.

For the purpose of determining treatment compliance, during the 16 hours, up to a total of one hour off treatment is allowed for any required transfers around the hospital and/or physio mobilisation.

To allow for patients who change from standard oxygen therapy to HFNT or vice versa, time on standard oxygen and time on HFNT will both be calculated separately for each patient.

9.8.3 Age

Age will be defined as age at baseline. Age will be calculated in years, as the difference between the date the baseline CRF was completed and the date of birth. Age will be rounded down to whole years (e.g. 50.65 years will be rounded down to 50 years).

9.8.4 ROX Index

ROX Index can be used to predict HFNT outcome. Data to calculate ROX Index will be collected at 2, 6, 12, 24 and 48 hours post extubation. It is calculated by the ratio of SpO_2/FiO_2 (%) divided by respiratory rate (breaths per minute).

9.8.5 EuroSCORE

The EuroSCORE II can be used to predict mortality at 30 days after surgery. Care data will be collected to calculate the EuroSCORE II. This will be recorded in the database as a single, pre-calculated percentage with no further derivation needed.

9.8.6 ARISCAT

The ARISCAT score can be used to predict the risk of in hospital pulmonary complications after surgery. Data will be collected to calculate the ARISCAT score. ARISCAT score will be recorded in the database as 7 individual components, from which we will calculate the ARISCAT score using the following scoring system (18) (19):

ARISCAT component	Level	Score
Age	≤50	0
	51-80	+3
	>80	+16
Pre-operative SpO2	≥96%	0
	91-95%	+8
	≤90%	+24
Respiratory infection in the last month	No	0
	Yes	+17
Preoperative anemia (Hgb ≤10 g/dL)	No	0
	Yes	+11
Surgical incision	Peripheral	0
	Upper abdominal	+15
	Intrathoracic	+24
Duration of surgery	<2 hours	0
	2-3 hours	+16
	>3 hours	+23
Emergency procedure	No	0
	Yes	+8

A patient receives one score for each of the seven ARISCAT components. The overall ARISCAT score is calculated by summing the scores for the seven ARISCAT components for each patient.

ARISCAT score can also be categorised into three risk categories such that:

- Low risk: ARISCAT score <26
- Intermediate risk: ARISCAT score 26-44
- High risk: ARISCAT score >44

9.8.7 EQ-5D-5L and QALYs

The mapping utility value of EQ-5D-5L will be created through crosswalk index value sets (20) or new EQ-5D-5L value sets should these become available, and QALYs will be calculated at 30 days and 90 days after surgery.

9.8.8 BMI

Body mass index (BMI) will be calculated using the following formula:

$$BMI = \frac{m}{h^2}$$

Where m = mass (in kilograms)
h = height (in meters)

9.8.9 Procedure length

Procedure length will be calculated as the time of procedure end minus the time of procedure start, and will be calculated in minutes.

9.8.10 Time with cross-clamp on

Time with cross-clamp on will be calculated as the time of cross-clamp coming off minus the time of cross-clamp going on, and will be calculated in minutes.

9.8.11 Time in inpatient locations

For the index admission, the time spent in a number of inpatient locations (ward, critical care unit, recovery unit, etc) will be calculated as the difference in minutes between the date and time a patient left a location minus the date and time a patient entered a location.

For example, if a patient entered the critical care unit at 3:15pm on 1st January 2021 and left the critical care unit at 7:20am on 2nd January 2021, their time spent on the critical unit would be 965 minutes (or 16 hours and 5 minutes). If the same patient then arrived on a ward at 7.30am on 2nd January 2021 and left the ward at 6:45pm

on 4th January 2021, their time spent on the ward would be 3555 minutes (or 2 days 11 hours and 15 minutes).

9.9 Protocol Deviations

There are two analysis populations relating to Protocol non-adherence. The intention-to-treat analysis population includes all subjects who were randomised, regardless of whether they received the treatment randomly allocated to them or completed follow-up. The data will be analysed assuming that the patient received the treatment arm they were randomly allocated. The per-protocol population includes all subjects who adhered to the trial Protocol by receiving the treatment randomly allocated to them for a minimum of 16 hours no matter whether they complete all of their follow-ups. Any subjects who did not receive the treatment randomly allocated to them will be excluded from the per-protocol population.

We do not expect other types of Protocol deviation that could impact the statistical analyses. However, if such an event happens, we will amend the SAP to include methods for handling these where necessary.

9.10 Demographic and Baseline Variables

Baseline data will be collected following consent. This will include basic demographic data (age, sex, residential status, etc), past medical history, as well as quality of life (EQ-5D-5L), activity of daily living (BARTHEL), health service and resource use questions, the EuroSCORE II and the ARISCAT score.

10 Efficacy Analysis

10.1 Primary Efficacy Analysis

The primary outcome is Days Alive and at Home within the first 90 days after surgery (DAH90). The calculation of Days Alive and at Home is specified in the Derived variables section (section 9.8). For the primary analysis, DAH90 will be treated as 0 for any patient that dies in the period between randomisation and the 90 day follow-up (11).

The primary analysis will be on the basis of intention to treat (ITT). The effects of adherence, attrition, and likely sources of bias on the primary analysis will be evaluated using per-protocol and sensitivity analyses.

Due to the skewed nature of DAH scores, a Mann-Whitney-Wilcoxon test will be used for the primary efficacy analysis at the end of the trial. Contrasts between treatment groups for the primary outcome (DAH90) will be used to evaluate the difference in the median DAH90 at a 5% significant level. 95% confidence intervals giving a range of plausible effects will be reported.

The primary analysis will be unadjusted for baseline variables (secondary analyses will be performed with adjustment for baseline variables for comparison).

The statistical analysis will be reported according to CONSORT extension guidelines for reporting of adaptive trials (21).

A review statistician will independently reproduce the final primary efficacy analysis, with the trial statistician's analysis code available to them if desired.

10.2 Secondary Analyses

10.2.1 DAH90 by country

As a secondary analysis, we will report the difference in medians with corresponding 95% confidence intervals for DAH90 between treatment groups for each individual country.

10.2.2 DAH90 by centre

As a secondary analysis, we will report the difference in medians with corresponding 95% confidence intervals for DAH90 between treatment groups for each individual centre, focusing on UK sites only.

10.2.3 DAH30

It is expected that the secondary outcome of days alive and at home up to 30 days (DAH30) will have similar distributional characteristics to that of DAH90. As a

secondary analysis, we will report the difference in medians with corresponding 95% confidence intervals for DAH30 between treatment groups.

10.2.4 Length of ICU stay and ROX index

As a secondary analysis, we will report the difference in medians with corresponding 95% confidence intervals between treatment groups the following variables:

- Length of initial ICU stay
- ROX index at 2 hours
- ROX index at 6 hours
- ROX index at 12 hours
- ROX index at 24 hours
- ROX index at 48 hours

10.2.5 Regression models for DAH90 and DAH30

Median (τ=0.50) quantile regression models, allowing estimation of effect sizes and adjustment for baseline covariates, will be used for the following analyses:

Outcome variables	Variables	Levels
DAH90	Treatment group	Standard oxygen [reference group], HFNT
DAH30	Treatment group	Standard oxygen [reference group], HFNT
DAH90	Treatment group	Standard oxygen [reference group], HFNT
	ARISCAT risk category	Low [reference group], intermediate, high
	EUROSCORE II	-
	Gender	Female [reference group], Male
	COPD	No [reference group], Yes
	Asthma	No [reference group], Yes
	Obesity	No [reference group], Yes
	Current smoker	No [reference group], Yes
	Lower respiratory tract infection in last 4 weeks	No [reference group], Yes
	Age	-
	First time or re-do surgery	First time [reference group], Redo
	Country	UK [reference group], Australia, New Zealand

Quantile regression was chosen due to the heavily skewed nature of the DAH endpoints and due to the expectation of two spikes close to 0 and 90/30 (depending on whether the endpoint under consideration is DAH90 or DAH30).

10.2.6 Binary outcomes (extubation timing and return to theatre)

Contingency tables will be used to assess the relationship between treatment group and the following binary outcome variables:

- Extubation timing (≤ or >24 hours after admission to ICU)
- Return to theatre (≤ or > 24 hours of admission to ICU)

10.2.7 ARISCAT score

The ARISCAT score can be used to predict the risk of in hospital pulmonary complications after surgery. We will therefore create a scatterplot of DAH90 against ARISCAT score, split by treatment group.

10.3 Sensitivity Analyses

10.3.1 Alternative population definitions

The primary efficacy analysis will be repeated with a number of alternative population definitions to examine the robustness of the results from the primary efficacy analysis. These alternative analyses will include:

- Using the per-protocol population instead of the intention-to-treat population
- Using each of the two time-on-treatment populations instead of the intention-to-treat population
- Including only the sub-group of patients who required an escalation of oxygen therapy (in line with the Protocol defined escalation protocol (Appendix 3)).

Due to multiplicity, we will only report the difference in medians with corresponding 95% confidence intervals between treatment groups for these alternative population definitions.

10.3.2 Alternative definition of DAH90 for patients that die during follow-up

In the primary efficacy analysis DAH90 will be treated as 0 for any patient that dies during the trial. We will perform a sensitivity analysis for the primary efficacy analysis to assess the impact of assigning a DAH90 score of zero to patients that die at any time within the 90 day follow-up period by relaxing this rule and replacing these zero values by the observed DAH90 value for these patients. In this sensitivity analysis, DAH90 will be calculated as per section 9.8.1 but DAH90 will not be automatically

scored as 0 for patients who die. Due to multiplicity, we will only report the difference in medians with corresponding 95% confidence intervals between treatment groups for this analysis.

10.3.3 Missing data

Missing data sensitivity analyses will be conducted in accordance with section 9.4.

10.4 Descriptive analyses

10.4.1 Summary statistics

Descriptive statistics (see Section 9.7) will be reported at the final analysis for a range of variables, split by treatment arm. These will include:

- Missing data for primary and secondary analysis variables
- Baseline demographic variables
- Baseline ethnicity, separately for UK, Australia and New Zealand
- Chosen follow-up method
- Patient-related risk factor for postoperative pulmonary complications
- Baseline residence and support requirements
- Past medical history
- Baseline surgical intervention plan
- Baseline EUROSCORE II and ASA risk rating
- Baseline ARISCAT score
- Pre-operative blood results
- Pre-admission medications
- Surgical procedure
- Oxygen log
- Post-operative blood gas on intervention
- Post-operative blood gas post intervention
- Post-operative blood results post extubation
- Post-extubation ROX index
- Return to theatre
- Diagnostic tests completed after tracheal extubation to discharge
- 30 day resource use
- 90 day resource use
- Discharge Participant and Family Resource Use questionnaire
- DAH30

- DAH90
- Duration (days) of hospital stay for index admission following surgery
- Time spent at each inpatient location for index admission
- Inpatient medication
- Medication after discharge from index admission
- Discharge destination
- Health professional visits after tracheal extubation
- Hospital re-admission following discharge
- Duration (days) of all hospital stays (including both index admission and any re-admissions)
- Frequency and duration (days) of non-hospital stays away from home following discharge from index hospital admission
- Baseline EQ-5D-5L questionnaire
- Discharge EQ-5D-5L questionnaire
- Day 30 EQ-5D-5L questionnaire
- Day 90 EQ-5D-5L questionnaire
- Baseline Barthel Index questionnaire
- Discharge Barthel Index questionnaire
- Day 30 Barthel Index questionnaire
- Day 60 Barthel Index questionnaire

The full specification of the variables to be included in the summary statistics is shown in the accompanying mock final analysis report.

10.4.2 Recruitment

A CONSORT diagram will show the flow of patients through the trial, from recruitment through to treatment allocation and the three follow-up time-points (discharge, 30 days and 90 days).

Descriptive statistics (see Section 9.7) of recruitment variables will be reported at the final analysis, split by treatment arm. These will include:

- Inclusion criteria
- Exclusion criteria
- Eligibility
- Number of patients at each study site by study stage
- Number of patients completing 90 day follow-up by country
- Withdrawals

The full specification of the variables to be included in the summary statistics of recruitment variables is shown in the accompanying mock final analysis report.

10.4.3 Compliance

Descriptive statistics (see Section 9.7) of compliance variables will be reported at the final analysis, split by treatment arm where appropriate. These will include:

- Summary of treatment compliance
- Reasons for non-compliance
- Summary of compliance for the time-on-treatment populations

The full specification of the variables to be included in the summary statistics of compliance variables is shown in the accompanying mock final analysis report.

10.5 Multiplicity

If missing data is below the threshold set out in section 9.4, or if missing data is above the threshold set out in section 9.4 but we do not proceed with multiple imputation, then we will only report a p-value for our primary analysis.

If missing data is above the threshold set out in section 9.4 and we proceed with multiple imputation then we will instead report the median p-value of the imputation method deemed acceptable.

We will not report p-values for other secondary, exploratory and sensitivity analyses.

11 Safety Analyses

All safety analyses will be performed on the safety population. Adverse events (AEs) and serious adverse events (SAEs) will be summarised separately.

Data on AEs and SAEs will be collected from the time of tracheal extubation to discharge. From discharge up to 90 days after surgery only data on SAEs will be collected.

11.1 Adverse Events

AE data will be listed by MedDRA Preferred Term and further grouped by System Organ Class. The frequencies of the AEs will be summarised by treatment groups.

11.2 Deaths, Serious Adverse Events and other Significant Adverse Events

The SAE data will be analysed and reported in the same manner as the AE data. Additionally, death, stroke, sepsis, myocardial infarction and acute kidney injury are serious adverse events of special interest. Therefore the mortality rate and incidence rates of stroke, sepsis, myocardial infarction and acute kidney injury will be estimated for each treatment group separately to the main SAE summary table.

12 Figures

Each figure will have a title and where appropriate a legend explaining any abbreviations and figure specific detail. Scales will be clearly labelled and kept the same across plots when appropriate.

13 Reporting Conventions

P-values $\geq$0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001". The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

14 Technical Details

This SAP is based on version 5.0 of the trial Protocol.

The statistical analyses will be carried out using R (www.r-project.org). Other major statistical software may be used where appropriate. The version number of any software used for the analysis will be recorded.

Data and analysis files will be saved to a secure, backed-up network drive and will be transferred to Royal Papworth Hospital at the end of the trial following standard operating procedures set out by Papworth Trials Unit Collaboration.

15 Summary of Changes to the Protocol

If there are any changes to the Protocol that impact the statistical analyses, this SAP will be updated and version-tracked accordingly.

16 Trial Funding

UK: National Institute for Health Research (NIHR Health Technology Assessment). Unique Award Identifier NIHR128351.

Australia: Medical Research Future Fund, Australia (APP2006100).

New Zealand: Green Lane Research and Educational Fund (21/23/4159).

17 References

1. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. Carrol Gamble, Ashma Krishan, Deborah Stocken, Steff Lewis, Edmund Juszczak, Caroline Doré, Paula R. Williamson, Douglas G. Altman, Alan Montgomery, Pilar Lim, Jesse Berlin, Stephen Senn, Simon Day, Yolanda Barbachano, Elizabeth Loder. s.l. : JAMA, 2017, Vols. 318(23), 2337-2343.

2. Predictors and early and late outcomes of respiratory failure in contemporary cardiac surgery. Filsoufi F, Rahmanian PB, G CJ, Chikwe J, Adams DH. 133, s.l. : CHEST, 2008. 713-21.

3. The use of continuous positive airway pressure by face mask and thoracic epidural analgesia after lung transplantation. Gothenburg Lung Transplant Group. Westerlind A, Nilsson F, Ricksten S. 13(3), s.l. : J Cardiothorac Vasc Anesth, 1999. 249-252.

4. High-flow nasal cannula: impact on oxygenation and ventilation in an acute lung injury model. Frizzola M, Miller T, Rodriguez M. 46(1), s.l. : Pedoatr Pulmonol, 2011. 67-74.

5. The effects of flowon airway pressure during nasal high-flow oxygen therapy. Parke R, Eccleston M, McGuinness S. 56(8), s.l. : Respir Care, 2011. 1151-1155.

6. Nasal high-flowtherapy delivers low level positive airway pressure. Parke R, McGuinness S, Eccleston M. 103(6), s.l. : Br J Anaesth, 2009. 886-890.

7. Clinical Usage of High-flow Oxygenation in Postcardiac Surgery Patients. Jabbari A, Alijanpour E, Tabasi S. 22(1), s.l. : Ann Card Anaesth, 2019. 107-108.

8. Assessment of humidification performance: still no easy method! Cuquemelle E, Lellouche F. 58(9), s.l. : Respir Care, 2013. 1559-1561.

9. The effect of high-flow nasal oxygen therapy on postoperative pulmonary complications and hospital length of stay in postoperative patients: A systematic review and meta-analysis. Zhonghua L, Chang W, Meng S, Xue M, Xie J, Xu J, et al. s.l. : Journal of Intensive Care Medicine, 2018.

10. Effect of high-flow nasal cannula oxygen therapy vs conventional oxygen on adult postcardiothoracic operation: A meta-analysis. X, Wu. 97(14), s.l. : Medicine (Baltimore), 2018.

11. Validation of days at home as an outcome measure after surgery: a prospective cohort study in Australia. Paul S Myles, Mark A Shulman, Stephane Heritier, Sophie Wallace, David R McIlroy, Stuart McCluskey, Isabella Sillar, Andrew Forbes. 7, s.l. : BMJ Open, 2017.

12. The effect of high-flow nasal oxygen on hospital length of stay in cardiac surgical patients: a randomised controlled trial. Zochios V, Collier T, Blaunszun G, Butchart A, Earwaker M, Jones N, et al. 73, s.l. : Anaesthesia, 2018. 1478-1488.

13. Programme NCA. National Audit Cardiac Surgery Audit 2014-2017 Summary report. [Online] 2018. Available at: https://www.nicor.org.uk/wp-content/uploads/2018/11/Adult-Cardiac-Surgery-Summary-Report-2014-17.pdf.

14. A Test of Missing Completely at Random for Multivariate Data with Missing Values. Little, R.J.A. 83, s.l. : Journal of the American Statistical Association, 1988. 1198-1202.

15. Multiple imputation by chained equations: what is it and how does it work? Azur MJ, Stuart EA, Frangakis C, Leaf PJ. s.l. : Int J Methods Psychiatr Res, 2011, Vols. 20(1):40-49.

16. Multiple imputation using chained equations: Issues and guidance for practice. Ian R. White, Patrick Royston, Angela M. Wood. 4, s.l. : Statist. Med, 2010, Vol. 30. 377–399.

17. Methods for significance testing of categorical covariates in logistic regression models after multiple imputation: power and applicability analysis. Eekhout, I., van de Wiel, M.A. & Heymans, M.W. 17, s.l. : BMC Med Res Methodol, 2017. 129.

18. Prediction of Postoperative Pulmonary Complications in a Population-Based Surgical Cohort. Jaume Canet, Lluís Gallart, Carmen Gomar, Guillem Paluzie, Jordi Valle, Jordi Castillo, Sergi Sabate, Valentín Mazo, Zahara Briones, Joaquín Sanchis. 113, s.l. : Anesthesiology, 2010. 1338 –1350.

19. Prospective external validation of a predictive score for postoperative pulmonary complications. Mazo V, Sabaté S, Canet J, Gallart L, de Abreu MG, Belda J, Langeron O, Hoeft A, Pelosi P. s.l. : Anesthesiology, 2014, Vol. 121(2). 219–231.

20. Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L values sets. al, Van Hout B et. 15, s.l. : Value in Health, 2012. 708–715.

21. The Adaptive designs CONSORT Extension (ACE) statement: a checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design. Munyaradzi Dimairo, Philip Pallmann, James Wason, Susan Todd, Thomas Jaki, Steven A Julious, Adrian P Mander, Christopher J Weir, Franz Koenig, Marc K Walton, Jon P Nicholl, Elizabeth Coates, Katie Biggs, Toshimitsu Hamasaki. 2020, BMJ, p. 369.