# The BALANCED Anaesthesia Study
# Delirium Substudy

Protocol
December 2014

# Statistical Analysis Plan

Document Date: September 2018

**Status: Draft**

# 1 TITLE

The BALANCED Anaesthesia Study is a prospective, randomised clinical trial of two levels of anaesthetic depth on patient outcome after major surgery.

.

# 2 RATIONALE

Delirium is a common and potentially serious complication for hospitalised elderly patients. It occurs after a wide variety of precipitating factors. Delirium is identified as a major challenge of perioperative care and identification of anaesthesia strategies to minimise its occurrence would have a major health benefit. Delirium is expensive both in terms of personal and economic cost. Moreover, the ageing demographic predicts that the incidence of delirium will increase greatly over the coming years. These poor outcomes lead to increased costs, and greater social and care giver burdens.

A surgical episode offers a unique opportunity in which to study delirium because it is a planned and predictable insult, which is known to result in this adverse outcome. Moreover, the high incidence of postoperative cognitive dysfunction (POCD) following major surgery makes this an appropriate high-risk group in which to investigate protective strategies for delirium.

This statistical analysis plan describes the analysis and reporting of the delirium sub-study of the BALANCED study. The delirium study was undertaken at 8 pre-specified centres within the BALANCED study.

# 3 STUDY DESIGN AND OBJECTIVES

## 3.1 Design

The Balanced study is an international multicentre, prospective, randomised, double blind (subjects, investigators and outcomes assessors), active control, parallel assessment, safety and efficacy study comparing bipsectral index (BIS) target groups 50 and 35. At selected centres additional assessments for delirium and cognitive status were undertaken.

## 3.2 Primary Objective

To compare the post-operative incidence of delirium amongst patients undergoing major surgery between those randomised to a BIS of target 50 (‘light’ anaesthesia) compared with those randomised to a BIS of target 35 (‘deep’ anaesthesia).

Primary endpoint:
The incidence of post-operative delirium, defined as POD present at any of the 10 assessments over the 5 post-operative days. The 3

minute diagnostic interview for Confusion Assessment Method delirium (3D-CAM) or the CAM-ICU is used to determine the presence of delirium. This will be administered preoperatively and twice daily (05:00 -11:00 and 1800 – 2200) for 5 days postoperatively or until discharge. A modified version will be used for patients in ICU (CAM-ICU) who are unable to talk on these days. Delirium assessment will be performed when patients can be aroused sufficiently in order to be assessed for delirium (Richmond Agitation and Sedation Score >-4).

### 3.3 Secondary Objectives

- To quantify and compare the temporal pattern of the incidence of POD amongst patients undergoing major surgery between those randomised to a BIS of target 50 ('light' anaesthesia) compared with those randomised to a BIS of target 35 ('deep' anaesthesia).

- To quantify and compare the severity of POD amongst patients undergoing major surgery between those randomised to a BIS of target 50 ('light' anaesthesia) compared with those randomised to a BIS of target 35 ('deep' anaesthesia).

- To quantify and compare the change in cognitive function at discharge (or 7 days) post-operatively between randomised groups.

- To quantify and compare the change in cognitive function at 3 and 12 months post-operatively between randomised groups.

- To quantify and compare cognitive impairment at discharge (or 7 days), 3 and 12 months post-operatively between randomised groups.

- To test the associations between POD incidence and subsequent MMSE changes or cognitive impairment and AMTS levels at discharge, or 3 and 12 months.

Secondary endpoints:

1. The incidence of postoperative delirium (POD) on each of the 5 post-operative days.
2. The number of postoperative delirium (POD) episodes over the 5 post-operative days (as a measure of delirium severity).
3. Cognitive change as measured by the Mini-Mental State Examination (MMSE) score at discharge (or 7 days).
4. Cognition level as assessed by modified MMSE (Abbreviated Mental Test Score - AMTS) administered at 3 and 12 months postoperatively by telephone.
5. Cognitive impairment defined as a score on the Mini-Mental State Examination (MMSE) at discharge (or 7 days) as <26 or as an AMTS

score of <8 at 3 and 12 months. Calculated for those with a score ≥26 pre-operatively.

### 3.4 Exploratory Objectives

- To identify demographic, co-morbid and perioperative factors (including anaesthesia depth, BIS, MAP and MAC) that are independently associated with of delirium incidence.

- To identify demographic, co-morbid and perioperative factors (including anaesthesia depth, BIS, MAP and MAC) that are independently associated with severity of delirium.

- To identify demographic, co-morbid and perioperative factors (including anaesthesia depth, BIS, MAP and MAC) independently associated with changes in MMSE score and AMTS levels at 3 and 12 months.

- To test the association between any POD diagnosis and the clinical outcomes, any cardiovascular event (MI, cardiac arrest, stroke, PE) and any infection during 1 year follow-up.

## 4 GENERAL ANALYSIS DEFINITIONS

### 4.1 Treatment Allocation

Particpants are randomised in a 1:1 ratio to a BIS of target 50 (‘light’ anaesthesia) or a BIS of target 35 (‘deep’ anaesthesia), with randomisation stratified by country.

### 4.2 Sample size calculation

Chan *et al.* showed that delirium incidence was reduced from 24.1% to 15.6% with the use of BIS-guided anaesthesia versus routine care in a broadly similar patient group to the BALANCED Study cohort, with the exception that 83% were ASA 1 or 2. Using the data from Chan et al. we hypothesise that the incidence of delirium will be reduced from 25% in the deep anaesthesia group (BIS target 35) to 15% in the light anaesthesia group (BIS target 50). A power analysis using this reduction from 25% to 15%, with power of 80% and two-tailed alpha = 0.05, indicates that approximately 270 patients are required in each group (total of 540 patients). We therefore plan to recruit approximately n=600 patients to allow for all-cause loss-to follow-up.

### 4.3 Participant populations

Approximately 600 participants from 8 sites in Australia, Hong Kong, USA and China will be enrolled in the study.

#### 4.3.1 Intention-to-treat (ITT) population

All randomised participants participating in the sub-study who do not have a diagnosis of delirium pre-operatively will be included in the ITT population with treatment allocation based on randomised treatment. The primary and secondary endpoints will be analysed using the full analysis set population i.e. the intention-to-treat population which includes all randomised patients, without a pre-operative diagnosis of delirium undergoing induction of general anaesthesia for surgery.

#### 4.3.2 Per-protocol population

The per-protocol (PP) population is defined as all randomised participants in the sub-study who meet all the inclusion/exclusion criteria for BALANCED with BIS group classified according to the actual median BIS value achieved irrespective of randomisation. Participants will be allocated to the BIS=50 group if the achieved median BIS is between 45 and 55 inclusive, and to the BIS=35 group if the achieved median BIS is between 30 and 40 inclusive. Participants who are not within these ranges will be excluded from these analyses. The primary and secondary endpoints will be compared between the two groups defined as per-protocol.

## *5 DEMOGRAPHIC AND BASELINE CHARACTERISTICS*

Participant demographic and baseline clinical characteristics will be summarised descriptively by randomised group. Summaries will include frequencies and percentages for categorical data and means, medians, ranges, standard deviations and inter-quartile ranges as appropriate for continuous data. These baseline summaries will include the following variables:

- Age
- Weight
- Sex
- Ethnicity
- ASA physical status

- Charlson score
- WHODAS score
- Cancer diagnosis
- Emergency/elective surgery
- Surgery type
- Haemoglobin
- Albumin
- Creatinine
- Surgical duration
- Pre-operative MMSE

These data will be examined for any imbalances between randomised treatments. No formal hypothesis testing will be conducted on these baseline variables.

# 6 PARTICIPANT DISPOSITION

A flowchart will be produced showing the flow of participants throughout the study. The numbers of participants;

- Randomised from the relevant centres
- Having each pre-operative and post-operative delirium assessment Surgical duration
- Having each pre-operative and post-operative MMSE assessment Surgical duration
- Lost to follow-up
- Dying during the 1 year follow-up will be summarised.

# 7 EFFICACY ANALYSES

## 7.1 Primary endpoint analyses

The primary endpoint analysis will compare the primary endpoint, diagnosis of delirium at any post-operative assessment, in the intention-to-treat population, using a Mantel-Haenszel Chi-square test, stratified by study-centre groups. The pooled estimate of the odds ratio and 95% confidence interval from this analysis will be used to summarise the statistical comparison. Analyses will be undertaken on both the Intention-to-Treat population and the Per-Protocol populations.

## 7.2 Secondary endpoint analyses

The analysis of the secondary endpoint the incidence of POD at each assessment time will be undertaken as supporting analyses of the primary

endpoint. These analyses will compare POD between randomised groups at each post-operative assessment, in the intention-to-treat population, using a Mantel-Haenszel Chi-square test, stratified by study-centre groups. The pooled estimate of the odds ratio and 95% confidence interval from this analysis will be used to summarise the statistical comparison. Analyses will be undertaken on both the Intention-to-Treat population and the Per-Protocol populations.

The analysis of the secondary endpoint the severity of POD, in the intention-to-treat population, will be undertaken using a general linear model with study-centre and randomized groups as fixed factors in the model. The mean difference between randomised groups derived from the model will be presented with the 95% confidence interval. If the distrubution of the severity does not meet the assumptions for this parametric analysis then the severity maybe square root transofmered prior to analysis. Analyses will be undertaken on both the Intention-to-Treat population and the Per-Protocol populations.

The analysis of the secondary endpoint the change in MMSE score at discharge, in the intention-to-treat population, will be undertaken using a general linear model with randomised group and study-centre as fixed factors in the analysis. The mean difference in the changes between randomised groups derived from the model will be presented with the 95% confidence interval. Analyses will be undertaken on both the Intention-to-Treat population and the Per-Protocol populations.

The analysis of the secondary endpoint the AMTS level at 3 and 12 months, in the intention-to-treat population, will be undertaken using a general linear model with randomised group and study-centre as fixed factors in the analysis. The mean difference in the levels between randomised groups derived from the model will be presented with the 95% confidence interval. Analyses will be undertaken on both the Intention-to-Treat population and the Per-Protocol populations.

The analysis of the secondary endpoint cognitive impairment at discharge (MMSE < 26) or at 3 and 12 months (AMTS<8) will be undertaken, in the intention-to-treat population, using a Mantel-Haenszel Chi-square test, stratified by study-centre groups. The pooled estimate of the odds ratio and 95% confidence interval from this analysis will be used to summarise the statistical comparison. Analyses will be undertaken on both the Intention-to-Treat population and the Per-Protocol populations, with all those identified as having cognitive impairment pre-operatively (MMSE < 26) excluded from the analyses.

The associations between POD incidence and subsequent MMSE changes at discharge and the AMTS levels at 3 and 12 months will be tested using a general linear model which will include randomised group, centre-stratum and

POD diagnosis as fixed factors and the MMSE change as the dependent variable.

### *7.3 Exploratory analyses*

The univariate associations between demographic, co-morbid and perioperative factors (including anaesthesia depth, BIS, MAP and MAC) and POD will be firstly undertaken using univariate logistic regression analyses. Factors identified from these analyses as showing some association with POD ($p<0.15$) will then be entered into multivariate forward and backward logistic regression models to identify a model containing those measures which are independently associated with POD. Adjusted Odds ratios and 95% confidence intervals will be deived from these models.

The univariate associations between demographic, co-morbid and perioperative factors (including anaesthesia depth, BIS, MAP and MAC) and POD severity will be firstly undertaken using 1-way ANOVA and Pearson's correlation coefficients. Factors identified from these analyses as showing some association with POD severity ($p<0.15$) will then be entered into multivariate forward and backward linear regression models to identify a model containing those measures which are independently associated with POD severity. Adjusted means and regression coefficients with 95% confidence intervals will be derived from these models. If the distribution of POD severity does meet the assumptions for these analyses then the data may be square root transformed prior to the analyses.

The univariate associations between demographic, co-morbid and perioperative factors (including anaesthesia depth, BIS, MAP and MAC) and MMSE changes and AMTS levels will be firstly undertaken using 1-way ANOVA and Pearson's correlation coefficients. Factors identified from these analyses as showing some association with MMSE and AMTS ($p<0.15$) will then be entered into multivariate forward and backward linear regression models to identify a model containing those measures which are independently associated with MMSE and AMTS. Adjusted means and regression coefficients with 95% confidence intervals will be derived from these models.

The association between any POD diagnosis and the clinical outcomes, any cardiovascular event (MI, cardiac arrest, stroke, PE) and any infection during 1 year follow-up will firstly be tested using a chi-square test. Following this multivariate models will be developed for these clinical outcomes using demographic, co-morbid and perioperative factors (including anaesthesia depth, BIS, MAP and MAC), to which a POD diagnosis will be added.

## 8 MISSING DATA

In the event that a patient withdraws (discontinues the study or is lost to follow-up) from the study, information on their survival and secondary endpoint status at the time they are withdrawn will be included in the relevant endpoint analysis. All available data will be used as appropriate to the objectives and analyses outlined in the analysis plan, there will be no imputation of missing data.

## 9 PROCEDURE FOR AMENDMENTS TO STATISTICAL PLAN

It is intended that all statistical analyses specified in this protocol will be performed. However, it is conceivable that some scheduled analyses may not be performed. In addition, study observations or analysis results may suggest the need for additional statistical analyses of the collected study data. Any revisions to this document prior to database lock will be made in the form of an amendment to the Statistical Analysis Plan. Any deviations or additional analyses that are performed will be summarised in the form of an addendum to the Statistical Analysis Plan. In either case, deviations (subtractions or additions) from the planned statistical analysis will be fully described in the final clinical study report.