STATISTICAL ANALYSIS PLAN for B3SHORT

Administrative information:

Sponsor name	Oslo University Hospital Fredrik Muller, Professor and Head of Department Avdeling for mikrobiologi, Rikshospitalet
Sponsor address	Pb 4950 Nydalen, 0424 Oslo Oslo, Norway
REK no	2017/1334C
Trial title	Shorter recovery time after critical illness
Trial ID	B3SHORT
Trial registration number (clinicaltrials.gov)	NCT04110028

SAP and protocol version:

SAP version and date:	This SAP is version 1.0, dated May 14th, 2024
Protocol version	This document has been written based on information contained in the study protocol version 1.46, dated June 12, 2023

SAP revision history:

Protocol version	SAP version	Section number changed	Description and reason for change	Date changed

SIGNATURE PAGE

PRINCIPAL/COORDINATING INVESTIGATOR:

Arne Søraas, MD, PhD
Avdeling for molekylær mikrobiologi
Oslo University Hospital

Signature

16/05/2024

Date (dd/mmm/yyyy)

TRIAL STATISTICIAN:

Morten Wang Fagerland, PhD
Head of Section for Statistics in Clinical Trials
Oslo Centre for Biostatistics and Epidemiology
Oslo University Hospital

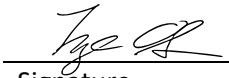
Signature

16/05/2024

Date (dd/mmm/yyyy)

QC STATISTICIAN:

Inge Christoffer Olsen, PhD
Clinical Trial Unit
Oslo University Hospital



Signature

16 May 2024

Date (dd/mmm/yyyy)

TABLE OF CONTENTS

1 Introduction

1.1 Background and rationale

Below follows a short summary from “Background and rationale” from study Protocol (B3Short study: shorter recovery time after critical illness. Protocol ID number 2017/1334C. Version number 1.46 dated June 12, 2023, by Oslo University Hospital, Department of Microbiology).

Please see protocol for list of references.

Background

Patients recovering from acute illness often experience prolonged recovery times, influenced by factors such as age, illness duration, and severity. Impaired mitochondrial energy production due to inflammation-induced depletion of nicotinamide adenine dinucleotide (NAD+) may contribute to prolonged recovery phase.

Therapeutic Information

Nicotinamide riboside (NR) is an orally available, safe NAD+ precursor that has been shown to restore NAD+ levels in humans and prolong life in old mice.

Rationale for the Study

Recovery from acute illness is associated with increased risks and delayed or incomplete recovery is common. No "recovery-improvement drug" currently exists. The study aims to administer NR or placebo for 90 days to patients in the recovery phase.

1.2 Trial Objectives

Primary Objective

The primary objective of this study is to assess the efficacy of dietary supplement NR on shortening recovery time after critical illness.

Secondary Objectives

The secondary objectives of this study are:

- To assess the effect of NR on mortality
- To assess NR adverse events
- Other secondary objectives are described in the protocol but will not feature in the main publication

2 Trial Methods

2.1 Trial Design

The B3SHORT study is designed as a double-blinded, randomized, three-phase, superiority study where three different doses of NR will be compared to placebo for its efficacy of shortening hospital stays after acute illness. The doses, allocation ratios, and sample size in the three phases are given in Table 1.

Table 1: Doses, allocation ratios, and sample size for the B3SHORT study

	Dose of NR	Dose label	Allocation ratio	Number of patients in the NR group	Number of patients in the placebo group
Phase #1	250 mg/day	Low dose	5:3	10	6
Phase #2	500 mg/day	Medium dose	5:3	10	6
Phase #3	1 g/day	Normal dose	10:3	20	6
Total				**40**	**18**

2.2 Randomization

Eligible patients are allocated in a 5:3 (phases 1 and 2) or 10:3 (phase 3) ratio between NR and placebo, using a computer randomization procedure. The randomization is blocked within each phase. Details of block size and allocation sequence generation is provided in a separate document unavailable to those who enroll patients or assign treatment.

Details of the randomization including the final random allocation list are held securely and unavailable to unauthorized trial personnel.

2.3 Sample size

Assuming the time of hospitalization after randomization follows a Weibull distribution with a scale parameter $\lambda = 18$ and shape parameter $\kappa=2.2$ (corresponding to a median of 15 and an IQR of 10 – 21), we need to include 58 patients: 18 in the placebo group, 10 in the 250 mg/day group, 10 in the 500 mg/day group, and 20 in the 1 g/day group) in order to reach approximately 80% power to show a significant dose-response in this study with a maximal reduction in time of hospitalization of 33%. This calculation is based on optimal designs in dose-finding studies according to the MCPMod methodology with three alternative dose-response curves: linear, e-max and exponential (Pinheiro et al., 2014).

2.4 Statistical Framework

The statistical framework will be based on the context of dose-finding, for which there are two main approaches: “regression” and “contrast” (Senn, 2007). In the regression approach, the actual doses used in the trial are viewed as samples from a dose continuum, and the purpose is to estimate the dose response curve. In the contrast approach, interest is on one or (more likely) several contrasts between the doses used in the trial. The former is the most explorative and may be sensitive to modeling choices, small number of doses, and over-fitting, especially for small sample sizes (such as in this trial). The latter is well suited for determining the *minimum effective dose*, in which we compare the different doses to placebo.

Specification of the Primary Objective

The primary objective of this study will be to establish the minimum effective dose of NR on shortening recovery time after critical illness. The primary outcome is length of hospital stay in days (see Section 5.1 for details).

There is only one identified primary analysis in this trial. All other efficacy analyses will be regarded as supportive or exploratory.

Hypothesis Setup

Let μ_1, μ_2, μ_3, and μ_4 denote the mean lengths of hospital stay for placebo, low dose (250 mg/day), medium dose (500 mg/day), and normal dose (1 g/day), respectively.

The statistical scheme to account for multiple testing is based on the *closed multiple test procedure* of Marcus et al. (1976). Under the monotonicity assumption, we define a set of elementary null hypotheses (the many-one comparison) with a preassigned order of performing the tests (Williams, 1971):

Step 1: Test the null hypothesis H_0: $\mu_1 = \mu_4$
Step 2: Test the null hypothesis H_0: $\mu_1 = \mu_3$
Step 3: Test the null hypothesis H_0: $\mu_1 = \mu_2$

The test in Step 1 is carried out with a nominal significance level of $\alpha = 0.05$. If the test is rejected, the test in Step 2 can be carried out, also with $\alpha = 0.05$. If the test in Step 2 is rejected, the test in Step 3 can be carried out (with $\alpha = 0.05$). The tests in steps 2 and 3 are only performed if the preceding null hypothesis is rejected, which – together with the monotonicity assumption – ensures that the significance level is preserved for the entire procedure (Williams, 1971; Bauer, 1991). All tests will be performed two-sided.

Decision Rule

The testing scheme above can result in four situations:

1. The null hypothesis in Step 1 is not rejected. Decision: No effect of NR for the doses in the study.
2. The null hypothesis in Step 2 is not rejected. Decision: The normal dose is the minimum effective dose.
3. The null hypothesis in Step 3 is not rejected. Decision: The medium dose is the minimum effective dose.
4. The null hypothesis in Step 3 is rejected. Decision: The low dose is the minimum effective dose.

Assumption

The dose response is assumed to be monotone, i.e., the effect of low dose is less than or equal to the effect of medium dose, which is less than or equal to the effect of normal dose. Or, consequently, that the mean length of hospital stay decreases monotonous from placebo (μ_1) to normal dose (μ_4):

$$\mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4.$$

Adaptation to Estimation

Hypothesis testing will only be performed for the primary outcome. For all secondary outcomes, estimation of treatment effect with 95% (two-sided) confidence intervals will be carried out instead of hypothesis testing in each of the three steps described above. The overall statistical approach will be the same; however, the criterion for performing steps 2 and 3 is that the 95% confidence interval from the preceding step does not contain the null effect value.

2.5 Statistical Interim Analyses and Stopping Guidance

There is no interim analysis of efficacy in this study.

The Data Monitoring Committee (DMC) will recommend stopping the trial if there is a safety concern. There is a separate DMC Charter detailing the procedures for the DMC.

2.6 Timing of Final Analysis

The main analysis is planned to after patients have concluded 90 days of treatment and attended the 65 weeks follow-up appointment. All data will be entered, verified and validated before the primary database will be locked, un-blinded and data analysis commenced.

2.7 Timing of Outcome Assessments

The target days and visits windows are defined in the protocol as:

Study visit name (V)	**Target Day**	**Definition (Day window)**
Prior to admission	-14	Two weeks prior to acute admission to hospital
V1. Baseline	Day 1 (Randomization)	Day 1
V3	Day 3	Target day -1 day/+2 days
In-hospital visits if still hospitalized: V7, V14, V21, V28	Day 7, 14, 21, 28	Target day± 3 days
Phone call: V30	Day 30	Target Day ±10 days
V400	Day X	Discharge date ± 0 days
V500	Day X	Withdrawal date ± 0 days
V1000	Day X	Death date ± 0 days
V90	Day 90	Target day + 90 days/-30days.
V365 End of study visit	65 weeks	Target day +/-180 days.

3 Statistical Principles

3.1 Confidence Intervals and *P*-values

All calculated *P*-values will be two-sided and compared to a 5% significance level. If a *P*-value is less than 0.05, the corresponding treatment group difference will be denoted as statistically significant. All efficacy estimates will be presented with two-sided 95% confidence intervals. The statistical approach to account for multiple testing of three doses versus placebo is the closed multiple test procedure (see Section 2.4). As there is only one primary null hypothesis setup to be tested in this trial, there will be no further adjustments for multiplicity.

3.2 Adherence and Protocol Deviations

Adherence to Allocated Treatment

Compliance from inclusion to discharge is defined by administrations recorded in hospital records. Post discharge compliance is determined by self-reported compliance. In addition, results from counting of returned capsules will be reported. Compliance is furthermore assessed based on the percent of subjects who have taken the scheduled number of pills and measurements of NR metabolites that will be available only after unblinding.

Compliance is defined as:

% compliance = (number of days the study drug was taken/ number of days the study medicine was supposed to have been taken.

A patient will be defined as non-compliant if compliance is less than 75% during the hospital stay.

Protocol Deviations

The following are pre-defined major protocol deviations regarded to affect the efficacy of the intervention:

- Entering the trial when the eligibility criteria should have prevented trial entry
- Non-compliance (defined above)
- Diagnosed cancer, or meeting other exclusion criteria, during the treatment period

Protocol deviations are classified prior to unblinding of treatment. The number (and percentage) of patients with major and minor protocol deviations will be summarized by treatment group with details of type of deviation provided. The patients that are included in the Full Analysis Set (see Section 3.4) will be used as the denominator to calculate the percentages. No formal statistical testing will be undertaken.

3.3 Estimands and Intercurrent Events

A *composite variable strategy* will be used to address intercurrent events (ICH E9(R1), 2019). In the B3SHORT trial, the relevant intercurrent event is death. Under the composite variable strategy, patients who die without a measurement of the outcome, and when death is informative of the outcome, will be given a value of the outcome variable that reflects that they are dead. This applies to some of the secondary outcomes. The precise definition of how this will be handled for each outcome will be given in Section 6 (Statistical Analysis Methods) under the Missing Data section.

3.4 Analysis Populations

The Enrolled set will include all patients who have provided informed consent and have been included into the study data base.

The Full Analysis Set (FAS) will be defined as all patients randomly assigned to a treatment group having received at least one dose of study medication.

The Safety Analysis Set (SAS) will include all patients having received at least one study medication. Subjects who withdraw from the study will be included in the SAS. A list of withdrawn subjects, with the reasons for withdrawal, will be made.

The Per Protocol Analysis Set (PPS) will include all randomised patients meeting the study eligibility criteria, with no major protocol deviations affecting the treatment efficacy, and who are compliant for the study drug.

4 Trial Population

4.1 Screening Data, Eligibility and Recruitment

The total number of screened patients and reasons for not entering the trial will be summarized and tabulated.

A CONSORT flow diagram will be used to summarize the number of patients who were:

- assessed for eligibility at screening (estimate only)
- eligible at screening (estimate only)
- eligible and randomised
- received the randomised allocation
- did not receive the randomised allocation*
- discontinued the intervention
- randomised and included in the primary analysis
- randomised and excluded from the primary analysis*

*reasons will be provided.

4.2 Withdrawal/Follow-up

The status of eligible and randomised patients at trial end will be tabulated by treatment group according to

- Completed intervention and assessments to the primary endpoint (discharge from hospital)
- completed intervention and assessments
- completed assessments but not intervention
- Withdrew (with reason)
- lost to follow-up

4.3 Baseline Patient Characteristics

- The patient demographics and baseline characteristics to be summarized include age in years, gender, ethnicity, body mass index, type of accommodation/level of care before admission, smoking status, level of frailty, Charlston score, in-hospital level of care, type of treatment during hospital stay (pressor, respiratory support, 02 supply), diagnosis, SAPS-score, degree of inflammation (highest CRP), mortality and morbidity (number of new organ failures).

Patient demographics and baseline characteristics will be summarized by randomised treatment arm and overall using descriptive statistics for continuous variables, and number and percentages of patients for categorical variables. There will be no statistical analysis of treatment difference. Any clinical important imbalance between the treatment groups will be noted.

5 Definitions of endpoints and Other Variables of Interest

Primary endpoint
Length of Hospital Stay
The length of hospital stay is defined as the number of days from commencing study treatment to medically fit for discharge from the hospital (to home, long term care facility or rehabilitation center or other facility with a lower care level). The term “Medically fit for discharge” is used instead of “discharge” to correct for extended stays due to logistical challenges (Example: medically fit for discharge, but no availability in long term facilities at that time). Only one patient was not discharged on the same day as she/he was medically fit for discharge). Admission to other hospitals after transfer from the first will be included in the total length of stay. The primary outcome will be the time from start of study treatment to discharge. Patients who die before discharge will be censored at the time of death.

Secondary endpoints (safety)
All-cause Mortality
All-cause mortality is a time to event outcome, with the date of randomization as the time when a patient becomes at risk. If a patient does not die during the study’s follow-up, the observation is censored at the patient’s last automated follow up.

New Infections
New infections is defined as any infection, mild or serious, occurring between the last point of contact to the next i.e. if a patient has an upper airway infection after discharge, but before the 90 day follow up appointment, this will count as one new infection at the 90 day follow up appointment. The outcome will be the number of new infections per patient, which is a discrete numerical variable.

Re-admissions
Re-admissions is defined as number of all-cause hospital admissions, surgical and medical, up until end of study. The outcome will be the number of new admissions per patient, which is a discrete numerical variable.

Other variables of interest
Metabolites
Rise in NR metabolites from inclusion to day 7 (or discharge if discharged prior to day 7) and until day 90.

Table 1: Overview of outcomes to be reported in the main article. Outcomes to be presented in subsequent articles are listed and defined in the protocol but not shown here.

Level	Outcome	Timeframe	Type
Primary	Length of hospital stay (to medical fit for discharge)	65 weeks	Time to event
Secondary	Change in grip strength	90 days	Continuous (change from baseline)
	Change in grip strength	7 days	Continuous (change from baseline)
	All-cause mortality	-	Time to event
	Number of new infections per patient	65 weeks	Discrete numerical
	Number of re-admissions per patient	65 weeks	Discrete numerical

6 Statistical Analysis Methods

The overall statistical approach is presented in Section 2.4 and involves 1-3 steps of hypothesis testing and/or estimation of treatment effect. In step 1, only the patients in the normal dose group of NR will be used (in addition to the patients in the placebo group); in Step 2, only the patients in the medium dose group will be used; and in step 3, only the patients in the low dose group will be used. This approach will be followed for the primary and all secondary outcomes. The statistical methods for steps 1-3 will be presented for each type of outcome in the following subsections.

As an exploratory analysis, we will also perform an overall vs. placebo analysis, where the low, medium, and normal dose groups are combined and compared with placebo.

6.1 Primary Outcome

The primary outcome is length of hospital stay (time to discharge), measured in number of days, and the primary analysis of the primary outcome will be performed on the full analysis set (FAS). A secondary (sensitivity) analysis of the primary outcome will be performed on the per protocol set (PPS).

Statistical Method for Steps 1-3

The comparison of the length of hospital stay between the NR and placebo groups – in each of the three steps – will be carried out with a Weibull survival model, with time as the number of days of hospital stay and treatment (NR vs placebo) as explanatory variable. The estimated treatment effect will be the hazard ratio (HR) for treatment with a 95% confidence interval.

Presentation of Results

The estimated treatment effect (HR) with a 95% confidence interval will be presented, together with the *P*-value for the null hypothesis of a zero-treatment difference (HR = 1).

The observed mean and standard deviation for placebo and each dose will be presented in a figure, along with the observed values (length of hospital stay) for each individual, as a scatter with connected lines-plot. Figure 1 shows an example with simulated data.

A plot of the Kaplan-Meier survivor function (for time to discharge) of each dose and placebo will be presented.

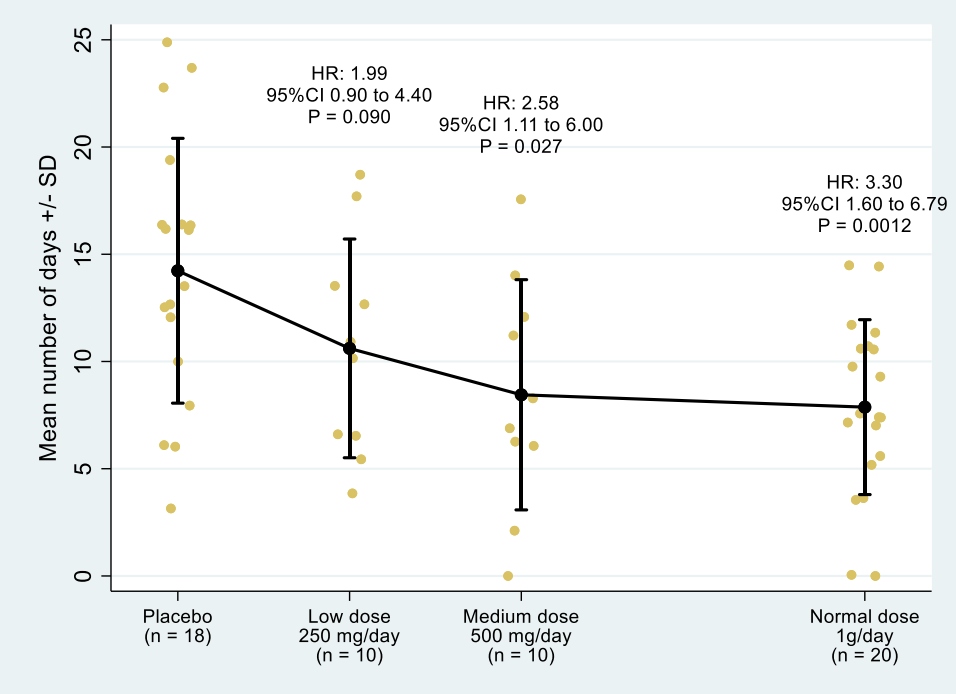


Figure 1: Simulated data

Missing Data

We will have complete data on the length of hospital stay, except for the following situation:
A patient dies before being discharged from the hospital. In this case, the patient will be censored at the time of death. Imputation of the length of hospital stay will be performed as a sensitivity analyses (see item 3 below).

Sensitivity Analyses

The following sensitivity analyses will be performed on the primary outcome:

1. As described above but on the per protocol set (PPS)
2. As described above but with Cox proportional hazard regression instead of Weibull regression
3. As described above but with Cox proportional hazard regression and imputation (instead of censoring) of the length of hospital stay with the maximum length of hospital stay (regardless of treatment group) for patients who die before discharge

Subgroup Analyses

There will be no subgroup analyses.

6.2 Secondary Continuous Outcomes (Change from Baseline)

Secondary continuous outcomes with change from baseline include grip strength at day 90 and grip strength at day 7. The analyses will be performed on the FAS.

Statistical Method for Steps 1-3

A linear mixed model will be fitted to the three measurements of grip strength (baseline, day 7, and day 90) in each step. The model will contain fixed effects for treatment (NR vs placebo), time point (day 7 vs baseline & day 90 vs baseline), and treatment x time point interaction and a random intercept. Based on the fitted model, we can estimate change from baseline to day 7 and change from baseline to day 90 in each group and the between-group difference in changes from baseline for both time points.

Presentation of Results

We will present the observed mean values and standard deviation for each group at each time point (baseline, day 7, and day 90). We also present the estimated change from baseline to day 7 and the estimated change from baseline to day 90 in each group with 95% confidence intervals, and the estimated between-group differences in changes from baseline to day 7 and to day 90 (the treatment difference), also with 95% confidence intervals.

Missing Data

Missing data on grip strength at either day 7 or day 90 will be handled by the linear mixed model, under the assumption of missing at random. There is no missing data on grip strength at baseline, so all patients will be included in the model.

6.3 Secondary Time to Event Outcomes

All-cause mortality is the only secondary time to event outcome. The analysis will be performed on the FAS.

Statistical Method for Steps 1-3

In each step, a Cox proportional hazard regression model with treatment (NR vs placebo) as explanatory variable for the outcome will be fitted. The time at risk is defined in Section 5.

Presentation of Results

The estimated HR for treatment with a 95% confidence interval will be reported. A plot of the Kaplan-Meier survivor function (for time to all-cause mortality) of each dose and placebo will be presented.

Missing Data

There will be no missing data on mortality during the study's follow-up time.

6.4 Secondary Discrete Numerical Outcomes

The secondary discrete numerical outcomes are number of new infections per patient and number of re-admissions per patient. Both outcomes are measured as a count from 0 to the maximal number of events (new infections or re-admissions). We expect the maximal numbers of events to be at most 10. The analyses will be performed on the FAS.

Statistical Method for Steps 1-3

The mean number of events per patient will be compared between the NR and placebo groups – in each of the three steps – with a two-sample T test with adjustment for unequal variances (Fagerland et al., 2011).

Presentation of Results

The observed number and percentages in each treatment group in each category (0, 1, 2, …, max number of events) will be presented together with the observed mean values and standard deviations in each treatment group and the estimated treatment difference with a 95% confidence interval.

Missing Data

If a patient has died, we will impute the missing data with the 80th percentile for the variable of interest (based on all patients regardless of treatment group). In cases where the 80th percentile= 0, the maximum number of events will be used.

6.5 Assumption Checks and Alternative Analyses

Assumption checks will be performed for the primary and all secondary outcomes according to the statistical methods used to analyzed them.

Weibull and Cox Proportional Hazard Models

The Weibull and Cox proportional hazard regression models assumes that the hazard ratio is constant over time. This will be checked by plotting -log(log(survival)) curves for each treatment against log(analysis time) and assessing if the curves are parallel. If the proportional hazard assumption is deemed to be violated, alternative parametric survival models will be fitted instead. The following survival distributions will be considered: exponential, Gompertz, and lognormal. The goodness-of-fit of each model will be assessed with plots of Cox–Snell residuals against the estimated cumulative hazard function of the residuals, and the one with the best fit will be the chosen model.

Linear Mixed Models

For linear mixed models, we will compare observed and model-predicted outcomes. If we detect a poor-fitting model, we will refit the model with a log-transformation of the outcomes. We do not expect other model-fit problems, due to the simplicity of the model (only a random intercept and a few fixed effects); however, in the unlikely event that we will encounter numerical problems with fitting the model, we will remove the two time point variables and substitute it with a single (linear) variable for time.

Two-sample T test

The discrete numerical outcomes (number of new infections per patient and number of re-admissions per patient) are limited to only a few possible values (with no great maximum), and as

such, there is no concern about the distribution of the outcomes, except for the situation where almost all patients have the same number and there is not enough variation in the data to get a meaningful result from the statistical methods. In that case, the outcome will be presented with the table of observed number and percentages in each treatment group in each category only.

7 Safety Analyses

General safety evaluations will be based on the incidence, intensity, and type of AEs, and clinically significant changes in the patient's physical examination findings, vital signs, clinical laboratory results and serum drug concentrations. Safety variables will be tabulated and presented for all patients in the safety set.

7.1 Adverse Events

Adverse events were registered at each point of contact with the patient. Additionally, some adverse events were recorded after a participant made contact or if the investigator otherwise was made aware of any adverse event. Both adverse events reported by the patient and those recorded in hospital records, are registered.

Adverse events were recorded according to an event reporting system (Common Terminology Criteria for Adverse Events (CTCAE version 5.0)). The investigator graded each AE accordingly using the levels mild, moderate severe, life threatening or leading to death. For tabulations, only treatment emerging adverse events (TEAEs) will be presented. TEAEs are defined as AEs with a start date on or after date of first randomised treatment. Any AEs prior to treatment will be listed but not tabulated.

The number (%) of subjects with any TEAEs, with 1, 2 or > 3 TEAEs, with treatment related TEAEs, with treatment emerging serious AEs (TESAE) and TEAEs leading to study drug withdrawal will be summarized by treatment group. The number of events and number (%) of subjects with adverse events by system organ class (SOC) and preferred term (PT) will be summarized by treatment group, overall, for severe AEs and for AEs leading to study discontinuation.

7.2 Clinical Laboratory Parameters

Safety clinical laboratory parameters were collected and assessed, but only used to identify adverse events. Clinical laboratory parameters will be summarized by treatment group and time point.

7.3 Vital Signs

Changes in vital signs (including systolic and diastolic blood pressure [mmHg], heart rate [beats per minute], peripheral 02 saturation (%), respiration frequency (numerical value), temperature (in Celsius degrees) and weight [kg]) will be summarized by treatment group and time point during admission.

8 Statistical Software

All statistical analyses will be done in Stata SE version 17 (StataCorp LLC, College Station, TX, USA).

9 References

Bauer P. Multiple testing in clinical trials. *Statistics in Medicine* 1991; 10:871-890.

Fagerland MW, Sandvik L, Mowinckel P (2011) Parametric methods outperformed non-parametric methods in comparisons of discrete numerical variables. *BMC Medical Research Methodology*, 11:44.

ICH E9(R1) Addendum on estimands and sensitivity analysis in clinical trials. *International council for harmonisation of technical requirements for pharmaceuticals for human use*; 2019.

Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika* 1976; 63:655-660.

Pinheiro J, Bornkamp B, Glimm E, Bretz F. Model-based dose finding under model uncertainty using general parametric models. *Statistics in Medicine* 2014; 33:1646-1661.

Senn S. *Statistical Issues in Drug Development*. John Wiley & Sons, 2007.

Williams DA. A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics* 1971; 27:103-117.