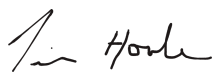
Statistical Analysis Plan (SAP) for the Investigation of Sleep in the Intensive Care Unit (ICU-SLEEP): A Randomized, Double-blind, Parallel-arm, Placebo-controlled Phase 2 Clinical Trial

Trial registration number	NCT03355053
SAP Version	v.6 (12 June 2024)
Protocol Version	IRB ID#: 2017P000090, v.14 (12 July 2022)
SAP Revision History:	**Justification**
v.1 (01 July 2017)	Initial draft
v.2 (05 January 2018)	Update statistical models for the exploratory endpoints
v.3 (15 February 2023)	Update plan for clarity and rigor prior to conducting any trial analyses
v.4 (14 August 2023)	Update plan for clarity and rigor prior to conducting any trial analyses
v.5 (29 November 2023)	Update plan for clarity and rigor prior to conducting any trial analyses
v.6 (12 June 2024)	Update plan for clarity and rigor prior to conducting any trial analyses
SAP Contributors:	
Hao Deng, M.D., M.P.H., Dr.P.H.(c), M.B.A.	Affiliation: Dept. of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital (MGH), Boston, MA, USA Role: Co-Investigator, Statistical design of trial
M. Brandon Westover, M.D., Ph.D.	Affiliation: Dept. of Neurology, Beth Israel Deaconess Medical Center, Boston, MA, USA Role: Principal Investigator, Statistical design of trial

Hao Deng		06 / 14 / 2024
Name of person writing the SAP (Printed)	**Signature of person writing the SAP**	**Date**
Timothy T. Houle		06 / 19 / 2024
Name of senior statistician (Printed)	**Signature of senior statistician**	**Date**
M. Brandon Westover		06 / 14 / 2024
Name of chief investigator (Printed)	**Signature of chief investigator**	**Date**

Introduction

Background and rationale

Sleep deprivation is common and often severe in critically ill patients cared for in intensive care units (ICUs) and is hypothesized to be a modifiable risk factor for delirium, which in turn is hypothesized to be a modifiable risk factor for long-term cognitive disability following recovery from critical illness. Dexmedetomidine (Dex) reduces the incidence of delirium in ICU patients by unknown mechanisms (1-5). The Investigation of Sleep in the Intensive Care Unit (ICU-SLEEP) Trial aims to determine whether Dex reduces delirium by improving sleep, whether a low- and/or very-low dose continuous infusion of Dex increases delirium-free days more, and the relationship between sleep deprivation in the ICU to long-term cognitive outcomes. We followed the guidelines for the content of statistical analysis plans (SAPs) recommended by Gamble et al. (6) to generate this document.

Grant Project Aims

Project Aim 1A: To compare the burden of delirium, as measured by the number of delirium-free days (DFDs), in ICU patients non-ventilated at study enrollment, who are receiving biomimetic sleep induced by Dex, given as a continuous overnight 1) very-low-dose or 2) low-dose infusion vs. 3) usual care and placebo.

Project Aim 1B: To assess whether the 1) very-low-dose continuous overnight infusion of Dex increases DFDs compared to the 2) low-dose continuous overnight infusion.

Project Aim 2A: To determine whether Dex reduces ICU delirium via reducing sleep deprivation, using causal mediation analysis.

Project Aim 2B: To determine the associations between specific components of acute cognitive impairment, seen in sleep deprivation and delirium, with specific measures of sleep deprivation.

Project Aim 3A: To determine whether ICU patients treated with Dex while in the hospital have a lower incidence of long-term cognitive impairment.

Project Aim 3B: To determine whether any differences in long-term cognitive impairment between ICU survivors treated with Dex vs. usual care and placebo are mediated by differences in sleep deprivation.

Study Methods

Trial design

ICU-SLEEP is a single-center, prospective, phase II, double-blind, placebo-controlled, three-arm, parallel-group, mechanistic, randomized clinical trial. The intervention arms of the trial include patients receiving intravenous dexmedetomidine, as either a low-dose (0.3 mcg/kg/hour) or very-low-dose (0.1 mcg/kg/hour) continuous overnight infusion. The control arm includes patients who receive standard ICU care plus normal saline placebo. The trial includes a pre-treatment period, a treatment period (up to 7 days), a post-treatment in-hospital period (up to 7 days), and a long-term follow-up period (12 months).

Randomization and blinding

Patients were randomized into one of the three arms, 1) very-low-dose continuous Dex, 2) low-dose continuous Dex, or 3) usual care and placebo (normal saline). Before trial enrollment begins, the primary biostatistician of the trial (Dr. Houle) will prepare a block randomization list which will be used to dispense study drug (or placebo) by the research pharmacy throughout the trial. The list includes varying, concealed block sizes, in a 1:1:1 proportion (i.e., equal numbers are allocated to each study arm; Dex vs. Placebo = 2:1), using random numbers generated by RStudio statistical software (Posit PBC, Boston, MA) (7). After the research team obtains written informed consent, independent research pharmacists will dispense either one of the two Dex interventions or placebo centrally, according to the computer-generated randomization list.

Randomization and blinding of study staff (participant, care provider, investigator, outcomes assessor) for treatment group allocation will be used to minimize bias. Treatment assignments will be concealed from the research team until trial completion. Blinding is further ensured by the fact that dexmedetomidine and placebo cannot be distinguished on the basis of appearance. Unblinding will be allowed only in cases where knowledge of the actual treatment is deemed essential by the treating physicians for further management of the patient. Additionally, if concerns arise regarding participant safety during the Data and Safety Monitoring Board (DSMB) review process, unblinding could be warranted to enable the DSMB to make well-informed decisions based on the available study data (see halting rules). In cases where unblinding becomes necessary, it will be approached cautiously to ensure that the trial's integrity is upheld while adequately addressing the safety concerns identified during the DSMB review.

Power and sample size
Utilizing a proportional odds ordinal logistic regression simulation-based method, a total sample size of N = 450 (150 for placebo vs. 150 + 150 = 300 for two Dex treatment groups) would yield power = 0.8 to detect a difference of 1.15 DFD at a significance level of 0.05. Prior literature (3, 8) suggested that an effect size of DFD = 1.5 is considered to be clinically meaningful, therefore our trial is sufficiently powered.

Framework
This is a superiority trial.

Statistical interim analysis and stopping guidance
No planned interim analysis for efficacy will be conducted. There are no formal stopping rules for the trial.

The DSMB will review safety data and study progress at least semiannually in years 1-4. In the first year, should >200 patients be enrolled, the DSMB will first meet after the 200th enrollment. Additional safety reviews may be requested by the DSMB. Details regarding the DSMB operations are specified in the DSMB charter. Statistical summaries of study safety data will be performed by Drs. Houle and Deng, the study statisticians. Interim analyses of safety data that would prompt temporary suspension of enrollment and/or study intervention use until a full safety review is convened (either routine or ad hoc) include, but are not limited to:

- Determination of unexpected, significant, or unacceptable risks to participants
- Insufficient compliance with protocol requirements
- Data that are not sufficiently complete and/or evaluable

These halting rules apply to study enrollment and administration of the study drug. They apply to all three arms of the study (given the blinding). They do not apply to components of the study outside of the study drug administration period. The study may resume once concerns about safety, protocol compliance, and data quality are addressed, and satisfy the Institutional Review Board (IRB) and DSMB.

Timing
Statistical analysis of the primary endpoint will be performed by Drs. Houle and Deng after this SAP document is finalized and signed by the key personnel listed at the beginning of this SAP. All study outcomes and timepoints are described in detail within the "Outcomes" section.

Statistical Principles
The statistical analyses specified for accomplishing the aforementioned study aims will be performed in accordance with this SAP and will follow the statistical principles described below. Any study outcomes and corresponding analyses not specified in this SAP will be treated as post-hoc analyses and will be analyzed and reported in the Appendix.

Descriptive statistics
We will report the phases of the ICU-SLEEP trial using a Consolidated Standards of Reporting Trials (CONSORT) flow diagram (version 2010) (9). It will demonstrate all stages of our trial, including screening, eligibility, enrollment, randomization and treatment allocation, outcome assessment and follow-up, study withdrawal, as well as analysis. Demographic information, baseline clinical characteristics, and all study outcomes will be summarized and reported for each arm of the trial. Numeric variables will be summarized using either means and standard deviations or medians and 25th/75th percentiles, depending on the data distribution. Categorical variables will be summarized using frequencies and proportions. No hypothesis testing for differences in baseline characteristics among study arms will be planned for assessing the success of randomization per CONSORT recommendations.

Inferential statistics and modeling
For the primary outcome of DFDs, treatment effect sizes will be reported as point estimates with 95% confidence intervals (CI) contrasting the difference among the pooled randomized study groups (pooled Dex-treated group vs. placebo) based on appropriate proportional odds ordinal logistic regression models. Based on prior literature, our primary outcome of DFDs generally follows a non-normal, zero-inflated, and non-standard distribution. We will adopt the proportional odds ordinal logistic regression (POOLR) method for primary analysis. This method is commonly viewed as an analytical equivalence to non-parametric univariate generalized Wilcoxon tests when constructed with no covariate adjustment. Our primary analysis will deploy a multivariable POOLR modeling adjusted by prognostic factors. For all other univariate comparisons

of non-baseline measures among treatment groups, the analysis will be performed using independent sample t-tests and/or Wilcoxon rank-sum tests for numeric variables, and χ^2 tests or Fisher's exact test for categorical variables, as determined to be appropriate based on data distribution. As for correlation metrics, we will use bootstrapped Pearson Correlation or Spearman Rank Correlation as determined to be appropriate. In case the assumptions for parametric statistical tests are violated, options for data transformations, alternative modeling techniques, or nonparametric equivalents will be explored. Prior to inferential analyses, all outcomes and covariates will be visually inspected, and their descriptive statistics will be first examined to determine follow-on corresponding choices of statistical tests and modeling approaches (e.g., non-normal DFD outcomes; median/quantiles; ordinal regressions). No multiple comparison adjustments will be a priori planned for secondary and exploratory outcomes, following the standard practice of analyzing multiple, prospective clinical trial outcomes (10, 11). Statistical analysis results for secondary and exploratory outcomes will be interpreted cautiously.

Planned subgroup analyses will be performed for both the primary and secondary endpoints. Although they are related, we distinguish subgroup analyses (focused analyses in biological subpopulations) from broader moderation analyses (analysis of treatment response heterogeneity depending on a wide range of covariates). For our subgroup analyses, we limit ourselves to a small set of important biologically/clinically distinct subgroups. All subgroup analyses will be considered "exploratory" /hypothesis-generating. Thus, any statistically significant results of these analyses will be considered preliminary, and future studies should be conducted for corroboration. Nevertheless, to help establish approximate prior probabilities for future confirmatory studies of any positive exploratory results, the numerical statistics and parameters, including all p-values, will be reported with their 95% CI estimates (rather than simply reporting whether the results are "significant" at the 0.05 level).

Confidence intervals and P values

All study tests are two-tailed, and we set the alpha of all study outcomes to 0.05. Effect size estimates of all tests and modeling estimates will be reported as point estimates along with their corresponding 95% CI levels. P-value and 95% CI adjustment may be considered in cases where additional unplanned post-hoc subgroup and exploratory analyses are performed. However, no P-value or 95% CI adjustments will be applied for planned analyses listed in this SAP.

Adherence to the intervention and protocol deviations

Adherence to the intervention will be assessed based on the proportion of study participants in the intervention arms that did not receive Dex treatment (low-dose or very-low-dose) and the proportion of participants in the control arm that received Dex. A descriptive summary table will be reported, including but not limited to the following elements:

- Proportion of participants in the intervention arm that did not receive Dex
- Proportion of participants in the control arm that received Dex
- Proportion of participants that withdrew from the study in all groups
- Total number of participants who received Dex in all groups
- Days between randomization/enrollment and the first Dex treatment

All deviations from the study protocol will be reported. Any deviations that caused safety risks to the study participants will also be considered protocol deviations. The number of participants with protocol deviations will be reported descriptively by the treatment group.

Analysis populations

The analysis population for all predefined outcomes will be based on the modified intent-to-treat (mITT) principle. We will also construct subset datasets for per-protocol and safety analyses based on the principles described below.

- Modified Intention-to-Treat (mITT) dataset: Analysis will be conducted following the mITT principle, meaning all randomized patients will be included for whom at least one efficacy measurement is obtained (12).

- Per-Protocol (PP) dataset: Statistical analysis of the primary endpoint will be performed based on a subset of the participants in the full analysis (mITT) who completed a sufficient number of study assessments that the data are likely to represent the effects of study intervention according to the underlying scientific model. For our PP analysis, we define the population as participants who received all study intervention for all 7 days and completed all of the study visits 1-22 (i.e., completion of the evaluation on the morning after the 7th overnight study drug administration).

- Safety dataset: Analysis will be conducted using the subset of study participants who received at least one dose of the study intervention.

Rigor, transparency, and reproducibility
Code and source data for all non-confidential statistical analyses (i.e., excluding all identifiable patient information) will be made available to reviewers when the study manuscript is submitted for publication. Study-related code and documents will be archived on the Brain Data Science Platform (bdsp.io), and a shareable link will be generated for scientific rigor, reproducibility, and transparency.

Trial Population

Screening data

ICU-SLEEP is a single-center study; hence, screening data (e.g., number of patients assessed for eligibility and corresponding screening notes) will be collected, summarized, and reported for scientific external validity, as applicable. Screening data is a useful indicator of whether our trial sample can represent the entire eligible study population.

Eligibility

Table 1: Study inclusion and exclusion criteria

Inclusion Criteria
1. Admitted or scheduled to be admitted to an MGH medical or surgical ICU (Blake 7 or 12, or Ellison 4)
2. Male or female, aged ≥ 50 years
3. Provision of signed and dated informed consent form (by patient or legally authorized representative (LAR))
4. Stated willingness to comply with all study procedures and availability for the duration of the study
5. Not on mechanical ventilation at the time of enrollment
6. Able to be enrolled before 7PM
7. For females of reproductive potential: pregnancy test is negative
Exclusion Criteria
1. Dementia, as measured by a score of ≥3.3 on the Informant Questionnaire on Cognitive Decline in the Elderly Short Form (IQCODE-SF)
2. Unable to be assessed for delirium (e.g. blindness or deafness)
3. Follow-up would be difficult (e.g. active substance abuse, homelessness)
4. Pregnancy or lactation
5. Known pre-existing neurologic disease or injury with focal neurologic or cognitive deficits
6. Serious cardiac disease (e.g. sick sinus syndrome without a pacemaker, sinus bradycardia, second- or third-degree AV block, congestive heart failure with ejection fraction <30%)
7. Severe liver dysfunction (Child-Pugh class C)
8. Severe renal dysfunction (receiving dialysis)
9. Low likelihood of survival >24 hours
10. Low likelihood of staying in ICU overnight
11. Known allergic reactions to components of dexmedetomidine
12. Patient is receiving or planning to go on dexmedetomidine at the time of enrollment
13. Patient is receiving either of the anticholinergic drugs scopolamine or penehyclidine; or alpha-2-agonist clonidine
14. Concomitant enrollment in another study protocol that may interfere with data acquisition or reliability of measurements
15. Deemed unsuitable for selection by the research team or ICU providers due to any medical, legal, social, language (non-English speaking) or interpersonal issues that would either compromise the study or the routine care of patients

Recruitment
Participants will be recruited from MGH ICU sites of Blake 7, Blake 12, and Ellison 4.

Withdrawal/Follow-up
Patients may be discontinued from study treatment and assessments for several reasons. These include: voluntary withdrawal by the patient/legally authorized representative; safety reasons as judged by the clinical and/or trial physicians; failure to maintain study eligibility; failure to receive study intervention within 3 consecutive nights following enrollment; failure to receive at least one night of ≥50% (5.5 hours) of study intervention during the dosing window. Patients for which the treatment period is terminated may still undergo post-treatment in-hospital and long-term assessments. Those who receive study intervention and are subsequently withdrawn/discontinued will not be replaced. Patients will be followed in-person for up to 14 days while in-hospital; and contacted by telephone at 3-, 6-, and 12-months ± 2-week window post-enrollment.

Baseline patient characteristics
Baseline characteristics (i.e., those present at admission or otherwise not related to the study intervention), encompassing demographics and laboratory measurements, will be compared between the study groups. Descriptive statistics will be employed to analyze the differences in these characteristics, in accordance with our pre-specified statistical principles. In

line with CONSORT recommendations, no formal hypothesis testing for differences in baseline characteristics among study arms will be planned for the purpose of assessing their comparability. However, exploratory post-hoc between-group tests may be conducted by our study statisticians to identify potential confounders or risk factors that could necessitate adjustment in subsequent analyses such as prognostic modeling.

Table 2: List of baseline patient characteristics

Age
Sex
Race; Ethnicity
Years of education
Informant Questionnaire on Cognitive Decline in the Elderly Short Form (IQCODE-SF)
Charlson Comorbidity Index (CCI)
Instrumental Activities of Daily Living (iADLs)
Primary admission diagnosis (%): Seizure; Neurovascular; Neuro-oncology; Infection; Cardiovascular; Hematology/oncology; Gastrointestinal; Respiratory (failure); Renal failure; Liver/hepatic failure; Metabolic disarray; Pancreatitis; Sepsis; Shock; Trauma; Liver transplant; GI surgery; Elective surgery; Other surgery; Neurology (other); Encephalopathy; Other
Sequential Organ Failure Assessment (SOFA)
Presence of any significant reported cardiac dysfunction (ejection fraction <30%)
Use of sedative or analgesic agents before enrollment (total doses while in ICU) ▪ Benzodiazepine; Propofol; Dexmedetomidine; Opiates; Pharmacologic sleep aids; Antipsychotic medications
Use of sedative or analgesic agents during days 1-7 of enrollment (total doses while in ICU) ▪ Benzodiazepine; Propofol; Dexmedetomidine; Opiates; Pharmacologic sleep aids; Antipsychotic medications
Days with "coma" (Richmond Agitation Sedation Scale (RASS) < -3)
Initial cognitive testing scores (before receiving the first study intervention): ▪ Numeric Rating Scale for Sleep (NRS-Sleep) score for sleep quality on the previous night ▪ Confusion Assessment Method (CAM) / CAM-ICU / CAM-S
Weight
Obstructive sleep apnea (% with diagnosis)
History of depression
Average light and noise levels throughout the ICU course
Extubation status at time of enrollment

Statistical Analysis

Outcomes

Study outcomes and time frames are defined in Table 3.

Table 3: Primary, secondary, and exploratory outcomes

Type	Variable	Description	Timeframe
Primary (Aim: 1A)	Delirium-free days (DFDs)	Sum of the number of days without delirium, defined as any positive CAM or CAM-ICU assessment during the first 14 hospital days in the two Dex arms combined (arms 1 and 2) vs. usual care (arm 3).	First 14 hospital days from start of infusion [or until hospital discharge, whichever occurs first]
Secondary (Aim: 1B)	Delirium-free days (DFDs)	Sum of the number of days without delirium, defined as any positive CAM or CAM-ICU assessment during the first 7 ICU days in the two Dex arms combined (arms 1 and 2) vs. usual care (arm 3).	First 7 ICU days from start of infusion [or until ICU discharge, whichever occurs first]
Secondary (Aim: 2A)	Sleep Quantity-quality (SQ) Score	Sleep composite measure formed by averaging the z-scores for raw measures of sleep quality (TST, SFI, time in N2, time in N3).	First 14 ICU days from start of infusion [or until ICU discharge, whichever occurs first]
Secondary (Aim: 2B)	Acute Cognitive Function (ACF) score	Reliable change index controlling for practice effects (RCI+PE) for a composite of acute measures of cognition. These include daily CAM-S (CAM-Severity) and psychomotor vigilance test (PVT) scores collected in the first 7 ICU days [or until ICU discharge, whichever occurs first]. Composites will be formed by averaging z-scores for CAM-S and PVT scores.	First 14 ICU days from start of infusion [or until ICU discharge, whichever occurs first]
Secondary (Aim: 3A)	Long-term Cognitive Function (LCF) score	Composite average of z-scores from long-term cognitive outcome measures of the different components.	3-, 6-, and 12 months post-enrollment
Exploratory (Aim: 2B)	Spearman correlation coefficient rho between SQ and ACF	Correlation between sleep quality on previous night (SQ) with acute cognitive function, as measured by the Acute Cognitive Function (ACF) score.	First 14 ICU days from start of infusion [or until ICU discharge, whichever occurs first]

*Definitions for composite variables SQ, ACF, and LCF can be found in the "Secondary outcomes/endpoints" section

Primary outcome/endpoints

In-hospital Delirium-free days (IH-DFDs) are calculated as follows: For each patient, delirium is assessed twice daily via the Confusion Assessment Method (CAM) and/or Confusion Assessment Method for the ICU (CAM-ICU) for days 1-14 within the hospital. Each assessment yields a binary determination (1 = delirious/CAM+, 0 = non-delirious/CAM-). Days with coma are counted together with delirium. If a patient has any positive delirium assessments on any given day of the assessment days, they are considered to have had delirium during these days. DFDs are calculated using the sum of the number of days without delirium, defined as any positive CAM or CAM-ICU assessment during the first 14 hospital days. CAM in all cases where it is possible to administer; otherwise, CAM-ICU is used.

Secondary outcomes/endpoints

ICU-Delirium-free days (ICU-DFDs) are calculated similarly to the primary endpoint of IH-DFDs. The difference is that ICU-DFDs are calculated using the sum of the number of days without delirium, defined as any positive CAM or CAM-ICU assessment, during the first 7 ICU days (instead of the first 14 hospital days).

Sleep quantity-quality (SQ) score: The measures of sleep for this study are based on the sleep staging data, from which we will extract the following measures of sleep quantity and quality. Each measure is computed once per 11-hour sleep staging period during the ICU stay: TST (total sleep time): time spent in any sleep stage (N1, N2, N3, R) between 8PM to 7AM; SFI (sleep fragmentation index): # arousals and awakenings/shifts to N1 divided by TST; Stage distribution: time spent in each sleep stage (expressed as a % of TST and in minutes). Sleep staging will be done from sleep physiology recordings (respiration and heart rate variability), using recently developed automated methods (13). By combining our measures of sleep quality (TST, SFI, time in N2, time in N3), a sleep composite measure can be formed. Composites will be created by z-normalizing raw scores and then averaging z-scores of component items (14). The composite measures reduce bias due to floor and ceiling effects.

Acute Cognitive Function (ACF) score: For each short-term cognitive outcome measure, following standard practice in neuropsychology (15), a reliable change index controlling for practice effects (RCI+PEscore) will be computed to minimize the practice effects of multiple assessments and to account for uncontrolled variability associated with time. RCIs will be expressed as z-scores. Each subject will have z-scores associated with performances on individual measures (CAM-S; and PVT scores), as well as a combined z-score. Cognitive impairment will be defined by either a combined z-score of -1.96, or at least two single-test z-scores of -1.96.

Long-term Cognitive Function (LCF) score: By combining our measures of long-term cognitive performance, a cognitive composite measure will be formed. Composites will be created by z-normalizing raw scores and then averaging z-scores of component items (14). The composite measures reduce floor and ceiling effects.

Analysis methods and model assumptions

Primary analysis

The primary endpoint (IH-DFDs) will be analyzed using the proportional odds ordinal logistic regression method. For primary analysis, the multivariable ordinal model aims to examine the association of levels of IH-DFDs (all available integer levels ranged from 0-14) as the dependent variable and the study group (i.e., pooled Dex vs. Placebo) as the independent variable. We expect that the randomization procedures will result in an adequate balance of baseline characteristics. To account for factors influencing the outcome other than the random treatment assignment process, the multivariable POOLR model will be performed by adjusting a predefined list of baseline variables (Table 2), including age, SOFA score, number of days in the ICU at enrollment, number of days in the hospital at enrollment, delirium status on the day of enrollment (Y/N), and IQCODE (cognitive impairment, Y/N, defined as a score of ≥3.3), which can be modified/updated based on findings from prior analysis of covariates and baseline characteristics. Results of the primary endpoint analysis will be presented as point estimates of Odds Ratios (ORs) with corresponding 95% CIs and P-values.

Secondary analyses

Similar to the primary endpoint, we will analyze the ICU-DFDs using the multivariable POOLR approach having the levels of ICU-DFDs (ranging from 0-7) as the dependent variable. The covariate-adjustment model construct will be performed similarly to the primary analysis.

For other secondary and numeric outcomes of SQ, ACF, and LCF, we will utilize linear mixed-effects models setting patient identifiers as random intercepts to address the repeated nature of these outcomes, which are collected at different timepoints. The independent variable is the group variable (all Dex arms vs. Placebo), and the dependent variables are SQ, ACF, and LCF numeric scores, respectively. Multivariable models may also be performed with appropriate confounding adjustments, starting with the same or a similar variable list used for the primary analysis.

Planned subgroup/sensitivity/exploratory analyses

To assess the correlation between sleep quality, as measured by the SQ score, with the Acute Cognitive Function (ACF) score, we will employ a linear mixed effects model with the patient identifier as a random intercept. The independent variable is the daily ACF score, and the dependent variables are the SQ score on the same day and the SQ from the previous night as a covariate. Pairwise comparisons among the Dex treatment groups will be performed to compare all primary and secondary outcomes among two treatment groups (i.e., low-dose vs. very-low-dose).

Analyses will be performed to test for differential treatment effects conditional on a priori selected variables. These variables were identified from prior literature (16-18), suggesting that they place individuals at altered risk for delirium, or based on potentially clinically important biological differences including age, sex, race, obstructive sleep apnea, opiate and antipsychotic drug administration, baseline severity of illness (SOFA, CCI) and cognitive impairment (IQCODE-SF scores). To examine differential treatment effects (i.e., moderation analyses), treatment effects will be isolated for each subgroup/moderator using a corresponding ordinal regression or mixed model by introducing an additional interaction term for examination (e.g., subgroup/moderator x treatment group). If a statistically significant interaction is observed, we will further quantify the effect size attributed to subgroups or effect modifiers using appropriate metrics and estimation methods.

Planned sensitivity analyses include unadjusted (i.e., crude) POOLR modeling for both primary and secondary endpoints without being adjusted by the prognostic variables. Additionally, we plan to perform sensitivity analyses of missing data handling under different assumptions (more details are explained in the "Missing data" section). We will also perform sensitivity analysis to examine the impact of baseline imbalance on the interpretation of treatment effects, if it emerges, and vary the assumptions or statistical methods used to account for these imbalances. Because our trial involves repeatedly measured outcomes such as LCF, additional sensitivity analyses may be performed to test the robustness of the results by varying different assumptions of the potential time-related intercurrent events or their modeling techniques.

Missing data

Missing data will be primarily statistically imputed using Multiple Imputations by Chain Equations (MICE) as described below. For models in which MICE cannot be implemented, we will use the model-based single imputation method or complete-case analysis as alternatives.

Missing outcomes

For the primary and secondary analysis of DFDs, we anticipate that up to 20% of formal delirium assessments may be missing because the patient is unavailable at the times the study assessor visits. We will use a variety of techniques to minimize missed assessments and instruct study team members to make an extra attempt to avoid missing data from >1 consecutive day. To accommodate the mITT principle, we will examine hospital records of diagnoses and treatments for missing assessment days to fill in missing outcomes, imputing delirium for any positive assessment, diagnosis, or treatment indicated for delirium.

- A patient being unavailable for assessment might be indicative of delirium in either direction (e.g., delirious patients could have more off-unit studies, spend more time in their rooms, refuse assessment more or less often, etc.). To address this, we will implement a set of predefined, rule-based imputation approaches and subsequent sensitivity analyses, contingent upon specific predetermined criteria/events.
- For the primary analysis, the outcome is framed as the cumulative days (up to day 14 or the point of hospital discharge) during which a patient remains both alive and free of delirium. Should patients exhibit no completed delirium assessments (for instance, due to death before regaining consciousness), they won't meet the mITT parameters and will thus be excluded from the primary analysis. Assuming survival up to day 14 or hospital discharge, or in the absence of any confirmed death within this 14-day window (given our presumption that patients discharged are generally in improved health and less likely to succumb quickly), the approach to address missing delirium data is as follows: 1) Code any missing evaluation between two negative assessments as non-delirious; 2) Code any missing evaluation adjacent to a positive assessment as delirious; 3) Evaluations post-discharge are marked as non-delirious.

- In instances where patients pass away prior to the 14-day mark or hospital discharge, the strategy for imputing missing outcomes includes: 1) Evaluation at the death date or subsequent evaluations are coded as delirious; 2) Code any missing evaluation prior to death, yet between two negative assessments, as non-delirious; 3) Code any missing evaluation or evaluations adjacent to a positive assessment as delirious, while the date of death is considered as positive.
- We will perform two sensitivity analyses. Sensitivity analysis #1 will interpret both the primary and secondary DFD results as zero (a conservative approach for validation). Sensitivity #2 will be conditional on the residual missing outcome rate post-rule-based imputation. Should this rate exceed 5%, this analysis will leverage advanced imputation techniques, such as MICE or random forests (e.g., the MissForest R package). Conversely, with a missing rate of 5% or lower, we will employ a complete-case analysis approach. For other non-DFD secondary outcome assessments (such as LCF), we will report for each outcome an analysis using only the complete data and separately for partially completed assessments with multiple imputations using the contemporary measured items.

Missing predictors

The core of our primary and secondary analysis of DFDs is an as-randomized comparison with covariate adjustment for baseline characteristics. For such adjusted models, the inclusion of these covariates is used to reduce the statistical variance of the estimates for enhanced study power. We will disclose an analysis employing only the fully completed data, and an additional one for partial datasets utilizing multiple imputations based on the concurrent measurements if substantial missing data is found (≥15%).

Safety data

Adverse events (AEs) will be coded using MedDRA (Medical Dictionary for Regulatory Activities). Each AE will be counted once only for a given participant. The severity, frequency, and relationship of AEs to study intervention will be presented by System Organ Class (SOC) and preferred term groupings, as specified in MedDRA. Information to be reported about each AE will include start date, stop date, duration, severity, relationship, expectedness, and outcome. Adverse events leading to premature discontinuation from the study intervention and serious treatment-emergent AEs will also be presented. Safety endpoints will be analyzed as summary statistics during treatment, using the statistical principles discussed in the descriptive statistics section.

Software details

RStudio and R version 4.0 (or above) will be used for statistical analyses (7). PASS (Power Analysis and Sample Size) software (19) was used to perform the power analysis for sample size determination.

Conclusion

This manuscript serves as the formal statistical analysis plan for the ICU-SLEEP Trial. All analyses will be performed as specified in this SAP. All amendments to this SAP have been reported.

Trial status

Recruitment status: Closed to enrollment
Recruitment start date: 05/29/2018
Recruitment completion date: 03/30/2022
Long-term Follow-up completion date: 03/31/2023
Overall trial status: Closed (Data Analysis Only)

Abbreviations

ACF: Acute Cognitive Function; ACME: Average Causal Mediation Effect; ADE: Average Direct Effect; ADLs: Activities of Daily Living; AE: Adverse Event; ARR: Absolute Risk Reduction; CAM: Confusion Assessment Method; CAM-ICU: Confusion Assessment Method for the Intensive Care Unit; CAM-S: Confusion Assessment Method – Severity; CCI: Charlson Comorbidity Index; CDAC (Clinical Data Animation Center); CI: Confidence interval; CMA: Causal Mediation Analysis; CONSORT: Consolidated Standards of Reporting Trials; Dex: Dexmedetomidine; DFDs: Delirium-free Days; DSMB: Data Safety and Monitoring Board**;** iADLs: Instrumental ADLs; ICU: Intensive Care Unit; IRB: Institutional Review Board; IQCODE-SF: Informant Questionnaire on Cognitive Decline in the Elderly Short Form; ICU-DFDs: ICU-Delirium-free days; IH-DFDs: In-hospital Delirium-free days; LAR: Legally Authorized Representative; LCF: Long-term Cognitive Function; MedDRA: Medical Dictionary for Regulatory Activities; MGH (Massachusetts General Hospital); MICE: Multivariate Imputation by Chained Equations; mITT: modified Intention-To-Treat; NRS-Sleep: Numeric Rating Scale for Sleep; OR: Odds Ratio; PASS: Power Analysis and Sample Size; PP: Per-Protocol; POOLR: Proportional Odds

Ordinal Logistic Regression; RASS: Richmond Agitation Sedation Scale; RCI: Reliable Change Index; RERI: Relative Excess Risk due to Interaction; SAE: Serious Adverse Event; SAP: Statistical Analysis Plan; SD: Sleep Deprivation; SFI: Sleep Fragmentation Index; SOC: System Organ Class; SOFA: Sequential Organ Failure Assessment; SQ: Sleep Quantity-quality; TST: Total Sleep Time

Acknowledgments
Not applicable.

Authors' contributions
HD, OA, BTT, TTH, RT, and MBW have contributed significantly to this statistical analysis plan. TTH is the senior statistician of the study. MBW is the principal investigator of the study. All authors have read and approved the final SAP version.

Funding
The ICU-SLEEP trial was supported by a grant awarded by the National Institute of Neurological Disorders and Stroke (NINDS) (R01NS102190). Funding agencies were not involved in the writing of this statistical analysis plan.

Availability of data and materials
Updates and results of the trial will be available to the public at ClinicalTrials.gov. Following study completion, the de-identified, archived data will be transmitted to and stored at the Brain Data Science Platform (bdsp.io), for use by other researchers, including those outside of the study.

Declarations
Ethical approval and consent to participate
The ICU-SLEEP Trial has been approved by the Partners Healthcare Institutional Review Board (IRB) (Reference Number 2017P000090). Written informed consent was obtained from all study participants.

Consent for publication
Not applicable.

Competing interests
OA has a provisional patent application describing the use of alpha-2 agonists for promoting N3 sleep.

Addendum

In a separate effort, we will perform additional exploratory analyses using causal mediation analysis methods. Proposed analyses include but are not limited to the items listed below. More details and complete analytical methods will be reported separately.

Type	Variable	Description	Timeframe
Exploratory (Aim: 2A)	Average Causal Mediation Effect (ACME)	Causal effect of sleep deprivation (SD) on IH-DFDs and ICU-DFDs, where SD is quantified by the sleep quantity-quality (SQ) score.	First 14 ICU days from start of infusion [or until ICU discharge, whichever occurs first]
Exploratory (Aim: 3B)	Average Causal Mediation Effect (ACME)	Causal effect of sleep deprivation (SD) on both acute and long-term cognitive outcomes, where SD is quantified by the SQ score, and cognitive function is quantified by the LCF score.	First 14 ICU days from start of infusion [or until ICU discharge, whichever occurs first]; 3-, 6-, and 12 months post-enrollment

Causal mediation analysis (CMA) will be used to estimate the Average Causal Mediation Effect (ACME) of sleep deprivation (SD), measured by the SQ score, on in-hospital/ICU delirium incidence, quantified as DFDs, and acute (short-term) and long-term cognitive outcome. We will also measure the average direct effect (ADE) of Dex on delirium, and of delirium burden on both short-term and long-term cognitive outcomes. The analysis will rely on the R mediation package (20).

References

1. Maldonado JR, Wysong A, van der Starre PJA, Block T, Miller C, Reitz BA. Dexmedetomidine and the reduction of postoperative delirium after cardiac surgery. Psychosomatics. 2009 Jun;50(3):206–17.
2. Pandharipande PP, Sanders RD, Girard TD, McGrane S, Thompson JL, Shintani AK, et al. Effect of dexmedetomidine versus lorazepam on outcome in patients with sepsis: an a priori-designed analysis of the MENDS randomized controlled trial. Crit Care. 2010;14(2):R38.
3. Su X, Meng ZT, Wu XH, Cui F, Li HL, Wang DX, et al. Dexmedetomidine for prevention of delirium in elderly patients after non-cardiac surgery: a randomised, double-blind, placebo-controlled trial. Lancet. 2016 Oct 15;388(10054):1893–902.
4. Skrobik Y, Duprey MS, Hill NS, Devlin JW. Low-Dose Nocturnal Dexmedetomidine Prevents ICU Delirium. A Randomized, Placebo-controlled Trial. Am J Respir Crit Care Med. 2018 May;197(9):1147–56.
5. Djaiani G, Silverton N, Fedorko L, Carroll J, Styra R, Rao V, et al. Dexmedetomidine versus Propofol Sedation Reduces Delirium after Cardiac Surgery: A Randomized Controlled Trial. Anesthesiology. 2016 Feb;124(2):362–8.
6. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, Williamson PR, Altman DG, Montgomery A, Lim P, Berlin J, Senn S, Day S, Barbachano Y, Loder E. Guidelines for the content of statistical analysis plans in clinical trials. JAMA. 2017;318(23):2337–43
7. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. R Foundation for Statistical Computing. *http://www.R-project.org/*.
8. Hughes CG, Mailloux PT, Devlin JW, Swan JT, Sanders RD, Anzueto A, Jackson JC, Hoskins AS, Pun BT, Orun OM, Raman R. Dexmedetomidine or propofol for sedation in mechanically ventilated adults with sepsis. New England Journal of Medicine. 2021 Apr 15;384(15):1424-36.
9. Schulz KF, Altman DG, Moher D. Consort 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:.
10. Schulz KF, Grimes DA. Multiplicity in randomised trials i: endpoints and treatments. Lancet. 2005;365(9470):1591–5.
11. Li G, Taljaard M, Van den Heuvel ER, Levine MA, Cook DJ, Wells GA, Devereaux PJ, Thabane L. An introduction to multiplicity issues in clinical trials: the what, why, when and how. Int J Epidemiol. 2016;46(2):746–55.
12. Petrovsky N, Ettinger U, Hill A, et al. Sleep deprivation disrupts prepulse inhibition and induces psychosis-like symptoms in healthy humans. The Journal of neuroscience: the official journal of the Society for Neuroscience 2014;34:9134-9140.
13. Sun H, Ganglberger W, Panneerselvam E, Leone MJ, Quadri SA, Goparaju B, et al. Sleep staging from electrocardiography and respiration with deep learning. Sleep. 2020 Jul 13;43(7):zsz306.
14. Banks S, Dinges DF. Behavioral and physiological consequences of sleep restriction. J Clin Sleep Med. 2007 Aug 15;3(5):519–28.
15. Duff K. Evidence-Based Indicators of Neuropsychological Change in the Individual Patient: Relevant Concepts and Methods. Archives of Clinical Neuropsychology. 2012;27(3):248-61. doi: 10.1093/arclin/acr120. PubMed PMID: PMC3499091.
16. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet 2014;383:911-922.
17. Nadler JW, Evans JL, Fang E, et al. A randomised trial of peri-operative positive airway pressure for postoperative delirium in patients at risk for obstructive sleep apnoea after regional anaesthesia with sedation or general anaesthesia for joint arthroplasty. Anaesthesia 2017;72:729-736.
18. Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. The New England Journal of Medicine 2013;369:1306-1316.
19. PASS 15 Power Analysis and Sample Size Software (2017). NCSS, LLC. Kaysville, Utah, USA, ncss.com/software/pass.
20. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. 2014.

Dropbox Sign

Audit trail

Title	SAP_ICU-SLEEP_06.12.24_Signed
File name	SAP_ICU-SLEEP_06....lean_Unsigned.pdf
Document ID	9e8ece11394003432d5895ea3cd3265edbfea496
Audit trail date format	MM / DD / YYYY
Status	● Signed

Document History

SENT	**06 / 14 / 2024** 15:27:33 UTC	Sent for signature to Hao Deng (hdeng1@mgh.harvard.edu), Timothy Houle (thoule1@mgh.harvard.edu) and M Brandon (bwestove@bidmc.harvard.edu) from rtesh1@bidmc.harvard.edu IP: 98.118.61.7
VIEWED	**06 / 14 / 2024** 15:36:33 UTC	Viewed by M Brandon (bwestove@bidmc.harvard.edu) IP: 65.242.64.66
SIGNED	**06 / 14 / 2024** 15:36:44 UTC	Signed by M Brandon (bwestove@bidmc.harvard.edu) IP: 65.242.64.66
VIEWED	**06 / 14 / 2024** 18:06:15 UTC	Viewed by Hao Deng (hdeng1@mgh.harvard.edu) IP: 73.68.53.44
SIGNED	**06 / 14 / 2024** 18:06:32 UTC	Signed by Hao Deng (hdeng1@mgh.harvard.edu) IP: 73.68.53.44

Dropbox Sign

Audit trail

Title	SAP_ICU-SLEEP_06.12.24_Signed
File name	SAP_ICU-SLEEP_06....lean_Unsigned.pdf
Document ID	9e8ece11394003432d5895ea3cd3265edbfea496
Audit trail date format	MM / DD / YYYY
Status	Signed

Document History

VIEWED	**06 / 19 / 2024** 17:38:39 UTC	Viewed by Timothy Houle (thoule1@mgh.harvard.edu) IP: 24.34.85.202
SIGNED	**06 / 19 / 2024** 17:38:56 UTC	Signed by Timothy Houle (thoule1@mgh.harvard.edu) IP: 24.34.85.202
COMPLETED	**06 / 19 / 2024** 17:38:56 UTC	The document has been completed.