STATISTICAL ANALYSIS PLAN

Study Protocol Number:	E2006-G000-304
Study Protocol Title:	A Multicenter, Randomized, Double-Blind, Placebo Controlled, Active Comparator, Parallel-Group Study of the Efficacy and Safety of Lemborexant in Subjects 55 Years and Older with Insomnia Disorder
Date:	05Feb2018
Version:	Final Version 3.0

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AE	adverse event
ANCOVA	analysis of covariance
AR	autoregressive covariance matrix
ATC	anatomical therapeutic class
BAI	Beck Anxiety Inventory
BDI-II	Beck Depression Inventory - II
BMI	body mass index
CCMV	complete case missing value
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CSR	clinical study report
eC-SSRS	electronic version of Columbia-Suicide Severity Rating Scale
EOS	end of study
EQ VAS	visual analogue score from EQ-5D-3L questionnaire
FAS	full analysis set
FSS	Fatigue Severity Scale
LEM5	lemborexant 5 mg
LEM10	lemborexant 10mg
ISI	Insomnia Severity Index
LPS	latency to persistent sleep
LS	least squares
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MNAR	missing not at random
PAB	Performance Assessment Battery
PBO	placebo

Abbreviation	Term
PD	pharmacodynamic
PGI-Insomnia	Patient Global Impression - Insomnia
PK	pharmacokinetic
PSG	polysomnography
QTcB	corrected QT interval by Bazett's formula
QTcF	corrected QT interval by Fridericia's formula
SAE	serious adverse event
SAP	statistical analysis plan
SD	Standard deviation
SE	sleep efficiency
SI	Système International
SMQ	Standardized MedDRA Queries
SOC	System Organ Class
sSE	subjective sleep efficiency
sSOL	subjective sleep onset latency
sTST	subjective total sleep time
sWASO	subjective wake after sleep onset
T-BWSQ	Tyrer Benzodiazepine Withdrawal Symptom Questionnaire
TEAE	treatment-emergent adverse event
TEMAV	treatment-emergent markedly abnormal laboratory value
TIB	time in bed
TST	total sleep time
UN	unstructured covariance matrix
WASO	wake after sleep onset
WASO1H	wake after sleep onset in the first half of the night
WASO2H	wake after sleep onset in the second half of the night
WHO DD	World Health Organization Drug Dictionary
ZOL	zolpidem tartrate extended release 6.25 mg (Ambien CR®)

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for Eisai Protocol E2006-G000-304.

This document is prepared based on the final study protocol amendment 4 (dated 05Feb2018). Reader is referred to the study protocol, the case report form (CRF), general CRF completion guidelines for details of study design, conduct and data collection.

1.1 Study Objectives

1.1.1 Primary Objective - US and Non-US

Demonstrate using polysomnography (PSG) that lemborexant (LEM10 and LEM5) is superior to placebo (PBO) on sleep onset as assessed by latency to persistent to sleep (LPS) after the last 2 nights of 1 month of treatment in subjects 55 years and older with insomnia disorder.

1.1.2 Secondary Objectives

Key Secondary Objectives - US Only

- Demonstrate that lemborexant (LEM10 and LEM5) is superior to PBO on sleep maintenance as assessed by sleep efficiency (SE) after the last 2 nights of treatment
- Demonstrate that lemborexant (LEM10 and LEM5) is superior to PBO on sleep maintenance as assessed by WASO after the last 2 nights of treatment
- Demonstrate that lemborexant (LEM10 and LEM5) is superior to zolpidem tartrate extended release 6.25 mg (Ambien CR®; ZOL) on wake after sleep onset in the second half of the night (WASO2H) after the last 2 nights of treatment

Key Secondary Objectives - Non-US Only

- Demonstrate that lemborexant (LEM10 and LEM5) is superior to PBO on sleep maintenance as assessed by SE after the last 2 nights of treatment
- Demonstrate that lemborexant (LEM10 and LEM5) is superior to PBO on wake after sleep onset (WASO) after the last 2 nights of treatment

Additional Secondary Objectives - US and Non-US

- Demonstrate that LEM5 or LEM10 or both LEM5 and LEM10 are superior to ZOL on postural stability in the morning after the first 2 nights of treatment
- Determine whether the efficacy of LEM5 or LEM10 or both LEM5 and LEM10 is superior to that of ZOL on selected PSG variables after the first 2 nights and the last

2 nights of treatment and on selected Sleep Diary variables over the first 7 nights and the last 7 nights of treatment.

- Confirm the efficacy of LEM5 and LEM10 compared to PBO on sleep as measured by PSG after the first 2 and last 2 nights of treatment and as measured by Sleep Diary over the first 7 and last 7 nights of treatment
- Evaluate the proportions of sleep onset and sleep maintenance responders to LEM5 and LEM10 and determine whether they are superior to that of ZOL and PBO as defined by response on PSG LPS and WASO and Sleep Diary subjective sleep onset latency (sSOL) and subjective wake after sleep onset (sWASO)
- Evaluate the safety and tolerability of lemborexant
- Determine whether the efficacy of LEM5 or LEM10 or both LEM5 and LEM10 is superior to that of ZOL and PBO on daytime functioning as assessed by the Insomnia Severity Index (ISI) and Fatigue Severity Scale (FSS) at the end of treatment
- Determine whether the safety of LEM5 or LEM10 or both LEM5 and LEM10 is superior to that of ZOL and PBO as assessed by cognitive performance in the morning after the first 2 nights of treatment

1.1.3 Exploratory Objectives - US and Non-US

- Explore the effects of LEM5, LEM10, ZOL and PBO on:
    - Subjective quality of sleep
    - Postural stability in the morning after the last 2 nights of treatment
    - Cognitive performance after the last 2 nights of treatment
    - Rebound insomnia in the 2 weeks following 30 days of treatment
    - Subjective ratings of morning sleepiness during and following completion of treatment
    - Sleep architecture parameters and other PSG variables
    - Health outcomes on the Patient Global Impression - Insomnia (PGI-Insomnia) and EQ-5D-3L
    - Withdrawal symptoms after completion of treatment
- Summarize plasma concentrations of lemborexant and its metabolites M4, M9, and M10
- Conduct population pharmacokinetic (PK) modeling for lemborexant
- Explore PK/pharmacodynamic (PK/PD) relationships between lemborexant concentrations and selected efficacy and safety variables

1.2 Overall Study Design and Plan

E2006-G000-304 is a multicenter, randomized, double-blind, placebo-controlled, active comparator (ZOL), parallel-group study of 2 dose levels of lemborexant for 30 nights in approximately 950 subjects 55 years or older with insomnia disorder. Subjects will be males 65 years or older or females 55 years or older. Approximately 60% of the subjects will be age 65 years or older.

The study will have 2 phases: The Prerandomization Phase and the Randomization Phase. The Prerandomization Phase will comprise 3 periods that will last up to a maximum of 28 days: a Screening Period, a Run-in Period, and a Baseline Period. The Randomization Phase will comprise a Treatment Period during which subjects are treated for 30 nights, and a minimum 14-day Follow-up Period before an End of Study (EOS) Visit.

An interim analysis is planned to be conducted after approximately 50% of subjects (approximately 475 subjects) have been randomized and either completed Day 31 assessments or discontinued from the study. This interim analysis will be conducted for administrative reasons as detailed in the separate Interim Analysis Charter.

The study design is illustrated in Figure 1.

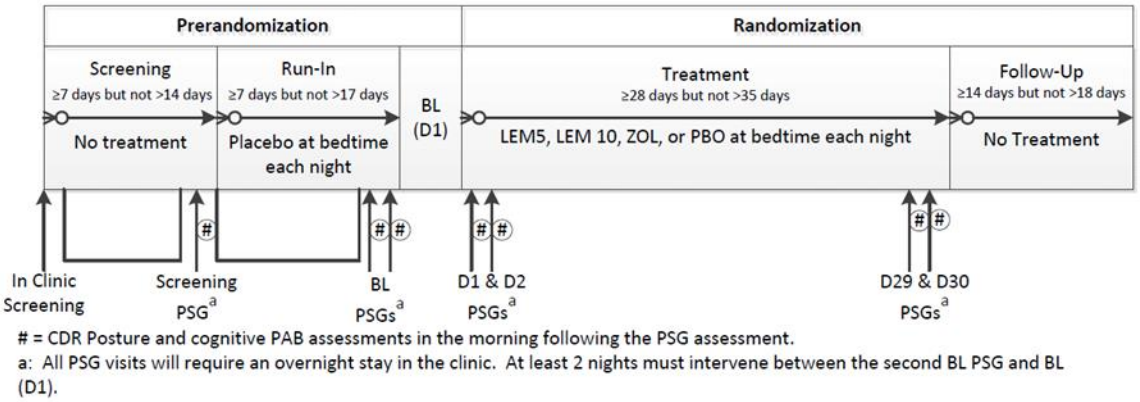

Figure 1 Study Design

"D" refers to the study day.
BL = baseline, EOS = End of Study, LEM5 = lemborexant 5 mg, LEM10 = lemborexant 10 mg, PAB = performance assessment battery, PBO = placebo, PSG = polysomnography, ZOL = zolpidem tartrate extended release 6.25 mg.

2 DETERMINATION OF SAMPLE SIZE

The sample size was estimated for each comparison of LEM10 vs. PBO and LEM5 vs. PBO with respect to the mean change from baseline of LPS at Month 1, on the basis of a two-sided t-test at the 0.05 α-level for each treatment comparison.

On the basis of the dose finding study E2006-G000-201 (Study 201), across various lemborexant doses (1 to 25 mg) at Days 14 and 15, the standard deviation (SD) of change from baseline for log-transformed LPS is assumed to be 0.9. The LS mean treatment difference at Days 14/15 from Study 201 for log-transformed LPS of LEM5 and LEM10 compared with PBO was -0.75 and -1.15, respectively. Therefore, a sample size of 250 subjects for LEM5, 250 subjects for LEM10, and 200 subjects for PBO has at least 95% power for each treatment comparison, LEM10 with PBO, and LEM5 with PBO, based on 2-sided 2-sample t-test at 5% significance level (Table 1).

Power is also estimated for the key secondary objectives, the comparison of LEM5 and LEM10 to PBO on change from baseline of SE and WASO, and LEM5 and LEM10 to ZOL on change from baseline of WASO2H (Table 1). A sample size of 250 subjects each for LEM5, LEM10, and ZOL, and 200 subjects for PBO has at least 95% power for detecting a statistically significant difference between LEM and PBO for change from baseline in SE, at least 80% power for detecting a statistically significant difference between LEM10 and ZOL/PBO for change from baseline in WASO/WASO2H based on 2-sided 2-sample t-test at 5% significance level.

Table 1 Power and Sample Size Calculation for Change from Baseline of LPS, SE, WASO2H, and WASO

Endpoint (Test)	Estimated Treatment Difference	Estimated SD	Power
Log(LPS) (LEM5 vs PBO)	-0.75	0.9	>95%
Log(LPS) (LEM10 vs PBO)	-1.15	0.9	>95%
SE (LEM5 vs PBO)	5%	14%	>95%
SE (LEM10 vs PBO)	7%	14%	>95%
WASO (LEM5 vs PBO)	-10 min	55 min	48%
WASO (LEM10 vs PBO)	-15 min	55 min	81%
WASO2H (LEM5 vs ZOL)	-8 min	38 min	65%
WASO2H (LEM10 vs ZOL)	-11 min	38 min	89%

Estimated treatment difference and SD are based on Study 201.

3 STATISTICAL METHODS

All final statistical analyses will be performed by the sponsor or designee after the study is completed and the database is locked and released for unblinding.

All descriptive statistics for continuous variables will be reported using number of observations (n), mean (arithmetic unless otherwise specified), standard deviation (SD), median, minimum and maximum. Categorical variables will be summarized as number and percentage of subjects. In summaries for safety the denominator for all percentages will be the number of subjects in a given treatment.

All statistical tests will be based on the 5% level of significance (two-sided).

3.1 Study Endpoints

3.1.1 Primary Endpoint(s)

The primary endpoint is:

- Change from baseline of mean LPS on Days 29 and 30 of LEM10 and LEM5 compared to PBO

3.1.2 Secondary Endpoint(s)

Key Secondary Endpoints - US Only

- Change from baseline of mean SE on Days 29 and 30 of LEM10 and LEM5 compared to PBO
- Change from baseline of mean WASO on Days 29 and 30 of LEM10 and LEM5 compared to PBO
- Change from baseline of mean WASO2H on Days 29 and 30 of LEM10 and LEM5 compared to ZOL

Key Secondary Endpoints - Non-US Only

- Change from baseline of mean SE on Days 29 and 30 of LEM10 and LEM5 compared to PBO
- Change from baseline of mean WASO on Days 29 and 30 of LEM10 and LEM5 compared to PBO

Additional Secondary Endpoints - US and Non-US

- Change from baseline on the postural stability test of mean units of body sway on Days 2 and 3 of LEM5 and LEM10 compared to ZOL

- Change from baseline of mean LPS, WASO, and total sleep time (TST) on Days 1 and 2 and Days 29 and 30 of LEM5 and LEM10 compared to ZOL
- Change from baseline of mean subjective Sleep Diary variables including sSOL, sWASO, subject sleep efficiency (sSE) and subjective total sleep time (sTST) over the first 7 and last 7 nights of the Treatment Period of LEM5 and LEM10 compared to ZOL
- Change from baseline of mean LPS, SE, WASO, WASO2H, and TST on Days 1 and 2 of LEM5 and LEM10 compared to PBO
- Change from baseline of mean WASO2H and TST on Days 29 and 30 of LEM5 and LEM10 compared to PBO
- Change from baseline mean of subjective Sleep Diary variables including sSOL, sWASO, sSE and sTST over the first 7 and last 7 nights of the Treatment Period of LEM5 and LEM10 compared to PBO
- Proportion of responders on Days 1 and 2 and Days 29 and 30 (PSG), and over the first 7 nights and last 7 nights of treatment (Sleep Diary), to LEM5 and LEM10 compared to ZOL and PBO, such that
    - Objective sleep onset response is defined as LPS ≤20 minutes (provided mean baseline LPS was >30 minutes)
    - Subjective sleep onset response is defined as sSOL ≤20 minutes (provided mean baseline sSOL was >30 minutes)
    - Objective sleep maintenance response is defined as WASO ≤60 minutes (provided mean baseline WASO was >60 minutes and is reduced by >10 minutes compared to baseline)
    - Subjective sleep maintenance response is defined as sWASO ≤60 minutes (provided mean WASO was >60 minutes and is reduced by >10 minutes compared to baseline)
- Change from baseline of the score from items 4-7 on the ISI at Day 31 of LEM5 and LEM10 compared to ZOL and PBO
- Change from baseline on the FSS score at Day 31 of LEM5 and LEM10 compared to ZOL and PBO
- Change from baseline of mean power of attention, mean continuity of attention, mean quality of memory, and mean speed of memory retrieval on Days 2 and 3

3.1.3 Exploratory Endpoint(s) - US and Non-US

The change from baseline of WASO2H for LEM10 and LEM5 compared to ZOL is considered exploratory for non-US. The following endpoints will also be explored for LEM5 and LEM10. Except for PK endpoints, comparisons to both ZOL and PBO will be made.

- Change from baseline of the mean rating on the Quality of Sleep question from the Sleep Diary of the first 7 days and last 7 days of the Treatment Period

- Change from baseline of mean power of attention, mean continuity of attention, mean quality of memory, and mean speed of memory retrieval on Days 30 and 31
- From the postural stability test, change from baseline of mean units of body sway after the first 2 nights of the Treatment Period compared to PBO and the last 2 nights of the Treatment Period compared to ZOL and PBO
- Rebound insomnia endpoints as assessed from the Sleep Diary during the Follow-up Period
    - Change from baseline of sSOL on each of the first 3 nights, mean sSOL of the first 3 nights, mean sSOL of the first 7 nights, and mean sSOL of the second 7 nights of the Follow-up Period
    - Change from baseline of sWASO on each of the first 3 nights, mean sWASO of the first 3 nights, mean sWASO of the first 7 and mean sWASO of the second 7 nights of the Follow-up Period
    - Proportion of subjects whose sSOL is longer than at Screening at the following time points during Follow-up Period: each of the first 3 nights, mean of the first 3 nights, mean of the first 7 nights, mean of the second 7 nights
    - Proportion of subjects whose sWASO is higher than at Screening at the following time points during Follow-up Period: each of the first 3 nights, mean of the first 3 nights, mean of the first 7 nights, mean of the second 7 nights
- Mean rating on the morning sleepiness item of the Sleep Diary on the first 7 mornings and last 7 mornings of the Treatment Period
- Mean rating on the morning sleepiness item of the Sleep Diary on the first 7 mornings and second 7 mornings of the Follow-up Period
- Change from baseline of mean morning sleepiness ratings assessed at 1.5 hours after wake time when subjects are in clinic on Days 1 and 2, and Days 29 and 30
- Change from baseline of mean minutes and mean percentage (a) per time in bed (TIB) and (b) per total sleep time (TST) of sleep stage N1, N2, N3 (separately and combined) and REM on Days 1 and 2 and Days 29 and 30
- Change from baseline of mean REM latency, mean number of awakenings, and mean number of long awakenings at Days 1 and 2 and Days 29 and 30
- Number and percentage of subjects with a rating of a positive medication effect on each PGI-Insomnia item at Day 31
- Change from baseline on the EQ-5D-3L at Day 31
- Mean score on the T-BWSQ of LEM5 and LEM10 compared to ZOL and PBO at end of study
- Proportion of subjects who score ≥3 on the T-BWSQ of LEM5 and LEM10 compared to ZOL and PBO at end of study
- PK of lemborexant and its metabolites M4, M9, and M10

- Relationships between lemborexant PK, efficacy, and/or safety variables using PK/PD modeling

3.1.4 Other Endpoints

The following PSG endpoints will be explored on an exploratory basis:

- Wake after sleep onset in the first half of the night (WASO1H)
- Duration of awakenings after persistent sleep
- Duration of long awakenings after persistent sleep
- Minutes and percentage of sleep stages per TIB: wake, non-REM (N1, N2, N3 separately and combined), REM
- Minutes and percentage of sleep stages per TST: non-REM (N1, N2, N3 separately and combined), REM
- WASO by quarter of the night

3.2 Study Subjects

3.2.1 Definitions of Analysis Sets

Safety Analysis Set: The Safety Analysis Set is the group of randomized subjects who received at least 1 dose of randomized study drug and had at least 1 postdose safety assessment.

Full Analysis Set (FAS): The FAS is the group of randomized subjects who received at least 1 dose of randomized study drug and had at least 1 postdose primary efficacy measurement.

PK Analysis Set: The PK analysis set is the group of subjects who have at least 1 quantifiable plasma concentration of lemborexant or its metabolites, or zolpidem, with adequately documented dosing history.

Per Protocol Analysis Set (PP): The PP is the group of all randomized subjects who received protocol-assigned study drug and do not meet any of the following criteria:

- Violated inclusion/exclusion criteria
- Duplicate randomization
- Missing primary efficacy assessment
- Primary efficacy assessment out of window
- Prohibited concomitant medication
- Study drug not administered

- Incorrect study drug kit dispensed

Subjects who met any of the criteria listed above will be excluded from the PP due to possible introduction of bias.

The number and percentage of subjects in each analysis set will be summarized by treatment groups using descriptive statistics. The summaries for FAS and PP will be based on subjects "as randomized". The summary for Safety Analysis Set will be based on subjects "as treated".

3.2.2 Subject Disposition

Subject disposition will be summarized by treatment group for all randomized subjects. The number and percentage of subjects who completed or discontinued prematurely from the study and their reason for discontinuation will be summarized by treatment group.

In addition, the number of subjects screened the number and percentage of screen failures and their primary reason for screen failures will be summarized. The number and percentage of randomized subjects will be summarized by region, country and sites by treatment group for all randomized subjects. The number and percentage of subjects in each of the analysis sets will also be summarized.

3.2.3 Protocol Deviations

Protocol deviations will be identified, reviewed and documented by the clinical team prior to database lock/treatment unblinding. All protocol deviations will be categorized according to major/minor and standard classifications including but not limited to the following:

- Violations of inclusion/exclusion criteria
- Noncompliance with or incorrect implementation of protocol procedures
- Noncompliance of study drug/dosage intervention
- Use of prohibited concomitant medication

Major protocol deviations will be summarized by category and treatment group.

3.2.4 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics for FAS and Safety Analysis Set will be summarized for each treatment group using descriptive statistics. Continuous demographic and baseline variables include age, height, weight, and BMI; categorical variables include sex, age group (55 to <65, 65 to <75, ≥75 years), BMI group (<18.5, 18.5 to <25, 25 to 30, >30), race and ethnicity.

The selected baseline assessments of Sleep Diary variables including sSOL, sWASO, sSE and sTST; PSG variables including LPS, WASO, SE, WASO2H and TST; ISI score and its

individual question score, and FSS will be summarized by treatment group. The BDI-II and BAI scores will also be summarized at study baseline.

3.2.4.1 Medical History

All medical histories as documented by the Medical History and Current Medical Conditions CRF will be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

The number and percent of subjects with medical history will be summarized by System Organ Class (SOC), preferred term for each treatment group based on Safety Analysis Set.

3.2.5 Prior and Concomitant Therapy

All investigator terms for medications recorded in the CRF will be coded to an 11-digit code using the World Health Organization Drug Dictionary (WHO DD; Mar 2017 or latest version).

Prior medications are defined as medications that stopped before the first dose of study drug, including placebo during the Run-In Period. Concomitant medications are defined as medications that (1) started before the first dose of study drug (including the Run-In Period) and are continuing at the time of the first dose of study drug, or (2) started on or after the date of the first dose of study drug (including the Run-In Period) to the last dose day plus 14 days.

The number and percentage of subjects who take prior and concomitant medications will be summarized using the Safety Analysis Set by treatment group, Anatomical Therapeutic Chemical class (ATC), and WHO DD preferred term (PT). If a subject takes the same medications for the same class level or drug name, the subject will be counted only once for that class level or drug name. Separate summary will be provided for subjects who take concomitant medication during Run-in Period and Treatment Period.

3.2.6 Treatment Compliance

Treatment compliance (in %) is defined as follows:

$$\frac{100 \times (\text{total number of tablets dispensed} - \text{total number of tablets returned or lost})}{\text{number of tablets expected to be taken}}$$

Treatment compliance during the Run-in and Treatment Period will be summarized separately using descriptive statistics based on Safety analysis set. Treatment compliance will also be summarized by treatment group using the categories <80%, ≥80% to ≤100%, >100% to ≤120%, and >120%. In addition to overall treatment compliance, separate summaries will also be provided for tablets that are matched to LEM and tablets that are matched to ZOL.

3.3 Data Analysis General Considerations

The FAS will be used as the primary population for all efficacy analyses. The Per Protocol analysis set will be used for sensitivity analyses to corroborate the primary efficacy endpoints.

3.3.1 Pooling of Centers

This study was a multicenter, international study with an estimated 105 centers participating in the study. Due to small expected number of subjects in each center, sites will be pooled within specific regions for primary and secondary efficacy analyses. Other analyses will be performed with all centers pooled across the study unless stated otherwise. Consistency of results across regions (North America and Europe) will be examined as specified in the respective sections in this document.

3.3.2 Adjustments for Covariates

Baseline assessment and age groups (55 to 64, and ≥ 65 years old) are used as covariates in the primary and secondary analyses.

3.3.3 Multiple Comparisons/Multiplicity

A sequential gate-keeping procedure will be used for the primary and the key secondary endpoint comparisons to control for the overall type I error at the 0.05 significance level (Figure 2). The first endpoint comparison will be tested at the 0.05 significance level. If the testing is found to be statistical significant, then proceed to the next endpoint testing at significance level of 0.05, otherwise stop testing.

The primary endpoints will be tested in the following order:

- Change from baseline of the mean LPS of Days 29 and 30 of LEM10 compared to PBO
- Change from baseline of the mean LPS of Days 29 and 30 of LEM5 compared to PBO

The key secondary endpoints will only be tested if both primary analyses are statistically significant at the 0.05 level. The key secondary endpoints will be tested in the following order:

US Only

- Change from baseline of the mean SE of Days 29 and 30 of LEM10 compared to PBO
- Change from baseline of the mean SE of Days 29 and 30 of LEM5 compared to PBO
- Change from baseline of the mean WASO of Days 29 and 30 of LEM10 compared to PBO

- Change from baseline of the mean WASO2H of Days 29 and 30 of LEM10 compared to ZOL
- Change from baseline of the mean WASO on Days 29 and 30 of LEM5 compared to PBO
- Change from baseline of the mean WASO2H on Days 29 and 30 of LEM5 compared to ZOL

Non-US Only

- Change from baseline of the mean SE of Days 29 and 30 of LEM10 compared to PBO
- Change from baseline of the mean SE of Days 29 and 30 of LEM5 compared to PBO
- Change from baseline of the mean WASO of Days 29 and 30 of LEM10 compared to PBO
- Change from baseline of the mean WASO on Days 29 and 30 of LEM5 compared to PBO

No multiplicity adjustment will be done on other efficacy analyses.

The gate-keeping testing procedure of the primary and secondary endpoints is illustrated in Figure 2:

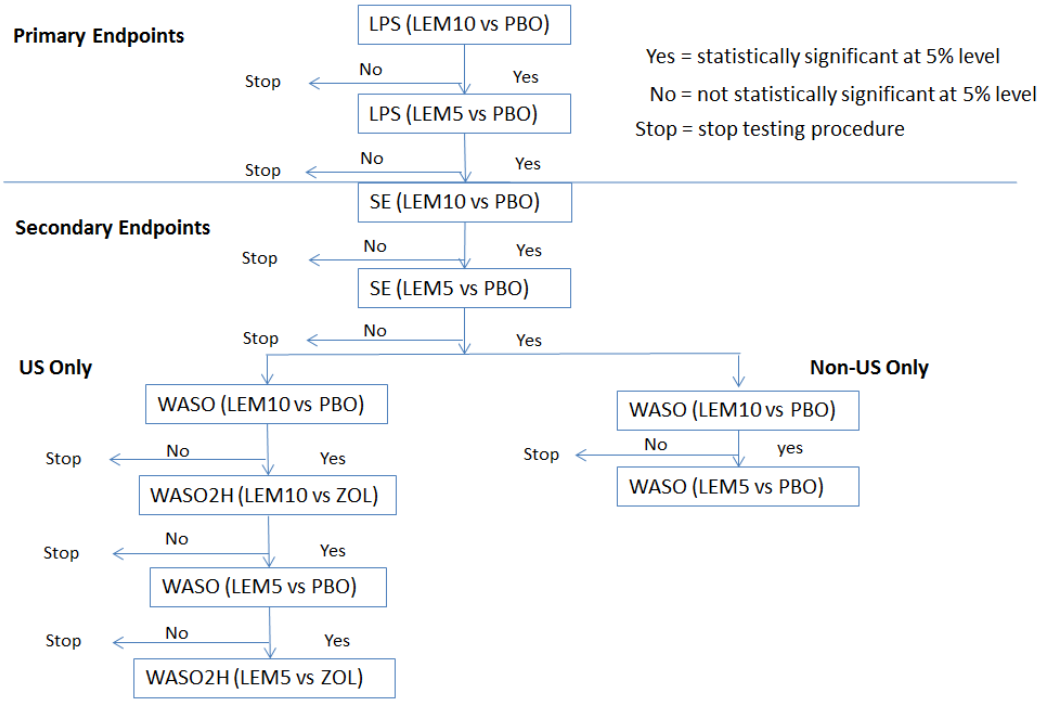

Figure 2 Flow Chart of Gate-Keeping Testing Procedure

LEM5 = lemborexant 5 mg, LEM10 = lemborexant 10 mg, LPS = latency to persistent sleep, PBO = placebo, SE = sleep efficiency, WASO = wake after sleep onset, WASO2H = wake after sleep onset in the second half of the night, ZOL = zolpidem tartrate extended release 6.25 mg.

3.3.4 Examination of Subgroups

Subgroup analysis of primary and key-secondary efficacy endpoints will be performed using age group (55 to <65, 65 to <75, ≥75 years old), alternative age group (55 to <65, ≥65 years old), sex (male and female), race (white, black, Asian, and other), region (North America and Europe), and BMI group (<18.5, 18.5 to <25, 25 to 30, >30) as detailed in Section 3.4.

3.3.5 Handling of Missing Data, Dropouts, and Outliers

Based on data on file and published clinical trials of similar mechanism (suvorexant, orexin receptor antagonist), the percentage of missing values related to efficacy is expected to be minimal and unlikely to affect the result of the primary and secondary efficacy analyses. Based primarily on data from the 1-month Phase 2 study of lemborexant (Study 201), the percentage of discontinued subjects from the lemborexant treatment group is expected to be approximately 5%. In suvorexant's Phase 3 program, the reported discontinuation rate due to any reason within 3 month of treatment was 8%, including less than 2% who discontinued due to lack of efficacy

The primary and key secondary efficacy endpoints will be analyzed using mixed effect model repeated measurement analysis (MMRM), the missing values will be imputed using

pattern-mixture multiple imputation (MI) assuming the missing data is missing not at random (MNAR) utilizing the complete case missing value pattern (CCMV). Additional sensitivity analyses will also be performed on primary and key secondary efficacy endpoints as follows:

MI Methods	Details	Analysis Type
Complete Case Missing Value (CCMV)	Subjects with missing data at any day are assumed to have a similar distribution as the ***completers within the respective treatment group*** (Pattern 1, below), where completers are defined as having no missing assessments for any post-baseline visits	Primary
Complete Case-4 CCMV(k=4)	This MI method will use all available monotone missing patterns to impute missing data assuming MNAR. This will relax the assumption of using only the complete cases as in the primary analysis. Study days where results are available: 1, 2, 29, 30 Pattern 1: x x x x Pattern 2: x x x . Pattern 3: x x . . Pattern 4: x . . .	Sensitivity
Tipping Point	Imputation towards the null hypotheses: A range of shifts will be used in the multiple imputation of missing data assuming MNAR to identify the specific shift and treatment effect that will tip the results from statistically significant to non-significant.	Sensitivity

Unless stated otherwise, missing values will be considered as non-responders in responder analyses and the continuous variables will be analyzed using MMRM assuming MAR. Details can be found in Section 3.4.

All safety analyses will be performed based on the observed data only.

3.3.6 Other Considerations

The following estimands are evaluated for the primary and key secondary efficacy endpoints in this study (Mallinckrodt, et al., 2012, and ICH E9(R1) Final Concept Paper, 2014). The details of the analysis method are discussed in Section 3.4.

Estimand	Description	Population	Intervention Effect of Interest	Analysis Type
Difference in outcome improvement for all randomized subjects	- all randomized subjects regardless of what treatment subjects actually received - include data after dropout	FAS	missing values imputed using MI assuming MNAR utilizing CCMV missing value pattern (complete cases) (Assumes the probability of missing observations for any subject depends on the unobserved events. For the missing pattern, complete cases will be used in the imputation. ***Thus this method assumes dropouts or subjects with missing values have similar treatment effect as the completers within the respective treatment group.***)	primary

Estimand	Description	Population	Intervention Effect of Interest	Analysis Type
Difference in outcome improvement for all randomized subjects	- all randomized subjects regardless of what treatment subjects actually received - include data after dropout	FAS	missing values will not be imputed; MMRM model is used on all available data assuming MAR (Assumes subjects with missing values behave the same as the observed data within that treatment group, i.e., the missingness is independent of unobserved data after accounting for the observed data in the model. ***Thus the dropouts or subjects with missing values may continue to benefit from the treatment as if they were still on treatment (just like completers.)***	Sensitivity (MMRM analysis assuming MAR)
Difference in outcome improvement for all randomized subjects	- all randomized subjects regardless of what treatment subjects actually received - include data after dropout	FAS	missing values imputed using MI assuming MNAR utilizing CCMV-4 missing value pattern (all available up to 4 monotone missing patterns) (Assumes the probability of missing observations for any subject depends on the unobserved events. For the missing pattern, complete cases up to 4 monotone missing patterns will be used in the imputation – see Section 3.4.1.3 for details. ***Thus this method relaxes the assumption of the primary analysis of using only completers to impute the missing data.***	Sensitivity (CCMV-4)
Difference in outcome improvement for all randomized subjects	- all randomized subjects regardless of what treatment subjects actually received - include data after dropout	FAS	a range of shifts will be used in the multiple imputation of missing data assuming MNAR to identify the specific shift and treatment effect that will tip the results from statistically significant to non-significant	Sensitivity (tipping point)
Difference in outcome improvement for all randomized subjects	- all randomized subjects regardless of what treatment subjects actually received - subjects who complete the study without missing efficacy assessments	FAS	subjects who completed all primary and secondary efficacy assessments without missing visits	Sensitivity (completer analysis)
Difference in outcome improvement for those who adhere to treatment	- subjects without major protocol violations that would impact efficacy assessments - include data after dropout	PP	missing values imputed using MI assuming MNAR utilizing CCMV missing value pattern (complete cases) (Assumes the probability of missing observations for any subject depends on the unobserved events. For the missing pattern, complete cases will be used in the imputation. ***Thus this method assumes dropouts or subjects with missing values have similar treatment effect as the completers within the respective treatment group.)***	sensitivity (PP analysis)
Difference in outcome improvement for those who adhere to treatment	- all randomized subjects; subject will be analyzed based on the actual treatment received - include data after dropout	FAS	missing values imputed using MI assuming MNAR utilizing CCMV missing value pattern (complete cases) (Assumes the probability of missing observations for any subject depends on the unobserved events. For the missing pattern, complete cases will be used in the imputation. ***Thus this method assumes dropouts or subjects with missing values have similar treatment effect as the completers within the respective treatment group.)***	sensitivity (as-treated analysis)

CCMV = complete case missing value; FAS = full analysis set; MI = multiple imputation; MAR = missing at random; MMRM = mixed effect model with repeated measurement; MNAR = missing not at random; PP = per-protocol analysis set; WASO2H = wake after sleep onset in the second half of the night

3.4 Efficacy/Pharmacodynamic Analyses

Unless specified otherwise, all efficacy endpoints will be summarized and analyzed using FAS. Baseline values for each efficacy parameter are defined in Section 6.2.

Unless specified otherwise, all efficacy/pharmacodynamic endpoints will be derived by calculating the averages of pairs of values [eg, average of LPS on Day 1 and Day 2 (denoted as Days 1/2 hereafter), average of LPS on Day 29 and Day 30 (denoted as Days 29/30 hereafter), …, etc.]

The primary and key secondary endpoints comparisons are tested following the gate-keeping testing procedure described in Section 3.3.3, Multiple Comparison/Multiplicity, to control for the overall type I error at the 0.05 significance level. The first primary efficacy endpoint comparison will be performed at the 0.05 significance level. The subsequent testing will only proceed if the previous test is statistically significant at the 0.05 level.

3.4.1 Primary Analyses

3.4.1.1 Primary Analysis

The primary efficacy endpoint is the change from baseline of LPS on Days 29/30 of LEM10 and LEM5 compared to PBO.

The null hypothesis of primary objective is that no difference exists in the mean change from baseline of LPS of Days 29/30 for treatment with LEM10 (or LEM5) as compared with PBO, and the corresponding alternative hypothesis is that a difference exists in the mean change from baseline of LPS of Days 29/30 for LEM10 (or LEM5) compared to PBO. The change from baseline of LPS on Days 1/2 and Days 29/30, will be analyzed using the mixed effect model repeated measurement analysis (MMRM) with factors of age group (55 to 64, and ≥ 65 years old), region (North America and Europe), treatment, visit (Days 1/2 and Days 29/30), and treatment-by-visit interaction as fixed effect, and baseline LPS as a covariate based on FAS. Since LPS is known to be non-normally distributed, a log-transformation will be used in the analysis. The unstructured covariance matrix (UN) will be used in the analysis. In the case of non-convergence of UN, the autoregressive [AR(1)] covariance matrix will be used in the model. Before the implementation of the MMRM model, the missing values will be imputed using pattern-mixture model multiple imputation (MI) assuming the missing values are missing not at random (MNAR) utilizing the complete case missing value pattern (CCMV - subjects who completed primary efficacy assessments without missing values). The missing values for a given visit will be imputed using all available values including the retrieved measurement from the post-discontinuation data.

The treatment comparison will be performed using contrasts. The p-value, least square (LS) means and the 95% confidence interval (CI) for the treatment difference will also be provided.

MULTIPLE IMPUTATION

Step 1 (imputing missing data): Thirty multiple imputed complete datasets were to be constructed using the imputation regression model of age, sex, race (white, black, and other), and region (North America, and Europe), baseline BMI, baseline log(LPS), baseline ISI, baseline sSOL, and individual log(LPS) assessments on Days 1, 2 , 29, and 30, with a predefined arbitrary seed number (seed=2359). SAS PROC MI will be used to implement the imputation procedure using all available values. The dataset will be converted into monotone missing pattern by imputing arbitrary missing data as the first step. The monotone data will then be imputed with monotone regression method and MNAR. The sample SAS statement can be found in Section 7.

Step 2 (performing MMRM using each imputed dataset): The MMRM model with factors of age group (55 to 64, and ≥ 65 years old), region (North America, and Europe), treatment, visit (Days 1/2, and Days 29/30), and treatment-by-visit interaction as fixed effect, and the baseline log(LPS) as a covariate will be applied to each imputed dataset. SAS PROC MIXED will be used for the MMRM analysis. The sample SAS statement can be found in Section 7.

Step 3 (combine results): Resulting treatment effect parameter estimators and standard errors from each of 30 multiple imputed datasets from Step 2 will be combined using SAS PROC MIANALYZE to obtain the pooled treatment effect and variance parameter estimators according to Rubin's rules (Rubin DB, 1987). The sample SAS statement can be found in Section 7.

3.4.1.2 Subgroup Analyses

The primary endpoint described in Section 3.4.1 will be summarized using descriptive statistics by each subgroup listed below. The MMRM model assuming MAR will be applied to provide the LS means and 95% CI for the treatment difference. No hypothesis testing (p-value) will be performed in the subgroup analyses.

In addition, forest plot, and median and median change over time of LPS will also be provided for each subgroup listed below.

- Age group (55 to 64, 65 to 74, ≥75 years old)
- Sex (male and female)
- Race (white, black, Asian and other)
- Region (North America and Europe)
- BMI group (<18.5, 18.5 to <25, 25 to 30, >30)

3.4.1.3 Sensitivity Analyses

The following analyses will be considered as sensitivity analyses:

- PP analysis: The same primary efficacy analyses described in Section 3.4.1 (MMRM analysis with MI for missing value imputation) will be repeated based on PP analysis set.
- Completer analysis: The same primary efficacy analyses described in Section 3.4.1 (MMRM analysis without missing value imputation) will be repeated on subjects who completed all primary efficacy assessments and have no missing visits.
- As-treated analysis: The same primary efficacy analyses described in Section 3.4.1 (MMRM analysis with MI for missing value imputation) will be repeated based on the actual treatment the subject received regardless of randomization.
- MMRM analysis assuming MAR: The same primary endpoint analysis described above will be analyzed using MMRM assuming the missing values are missing at random (MAR; MMRM analysis without missing value imputation).
- MI Imputation assuming MNAR utilizing CCMV-4: The same MMRM method used in the primary analysis will be applied utilizing CCMV-4 (ie, up to 4 monotone missing patterns will be used for missing value imputation as follows):

Study days where results are available	**1**	**2**	**29**	**30**
Pattern 1	x	x	x	x
Pattern 2	x	x	x	.
Pattern 3	x	x	.	.
Pattern 4	x	.	.	.
x = result present; . = result missing				

- Tipping point analysis: A range of shifts will be used in the multiple imputation of missing data assuming MNAR to identify the specific shift and treatment effect that will tip the results from statistically significant to non-significant.

3.4.2 Secondary Analyses

3.4.2.1 Key Secondary Analyses

CHANGE FROM BASELINE OF SE ON DAYS 29/30

The change from baseline of SE on Days 1/2 and on Days 29/30 will be analyzed using the same MMRM model as the primary efficacy endpoint with factors of age group (55 to 64, and ≥65 year old), region (North America, and Europe), treatment, visit (Days 1/2, and Days 29/30), and treatment-by-visit interaction as fixed effect, and baseline SE as covariates based on FAS. The unstructured covariance matrix will be used in the analysis. In case of non-convergence, the AR(1) will be used in the model. The missing values will be imputed using a pattern mixture model utilizing MI assuming MNAR. Before the implementation of the MMRM model, the missing values for a given visit will be imputed using all available values including the retrieved measurement from the post-discontinuation data.

The treatment comparison will be performed using contrasts. The p-value, least square (LS) means and the 95% confidence interval (CI) of the treatment differences will also be provided.

Multiple Imputation

The same 3 steps (imputing missing data, performing MMRM using each imputed dataset, and combine results) will be implemented as described in Section 3.4.1. The complete data sets will be constructed using regression model of age, sex, race (white, black, and other), region, baseline BMI, baseline SE, baseline ISI, baseline sSE, and individual SE assessments on Days 1, 2 , 29, and 30.

CHANGE FROM BASELINE OF WASO2H ON DAYS 29/30

The change from baseline of WASO2H on Days 1/2 and on Days 29/30 will be analyzed using the same MMRM model as the primary efficacy endpoint with factors of age group (55 to 64 years, and ≥ 65 year old), region (North America, and Europe), treatment, visit (Days 1/2, and Days 29/30), and treatment-by-visit interaction as fixed effect, and the baseline WASO2H as covariates based on FAS. The unstructured covariance matrix will be used in the analysis. In case of non-convergence, the AR(1) covariance matrix will be used in the model. The missing values will be imputed using a pattern mixture model utilizing MI assuming MNAR. Before the implementation of the MMRM model, the missing values for a given visit will be imputed using all available values including the retrieved measurement from the post-discontinuation data.

The treatment comparison will be performed using contrasts. The p-value, least square (LS) means and the 95% confidence interval (CI) of the treatment differences will also be provided.

Multiple Imputation

The same 3 steps (imputing missing data, performing MMRM using each imputed dataset, and combine results) will be implemented as described in Section 3.4.1. The complete data sets will be constructed using regression model of age, sex, race (white, black, and other), region, baseline BMI, baseline WASO2H, baseline ISI, baseline sWASO, and individual WASO2H assessments on Days 1, 2 , 29, and 30.

CHANGE FROM BASELINE OF WASO ON DAYS 29/30

The change from baseline of WASO on Days 1/2 and on Days 29/30 will be analyzed using the same MMRM model as the primary efficacy endpoint with factors of age group (55 to 64, and ≥65 year old), region (North America, and Europe), treatment, visit (Days 1/2, and Days 29/30), and treatment-by-visit interaction as fixed effect, and the baseline WASO as covariates based on FAS. The unstructured covariance matrix will be used in the analysis. In case of non-convergence, the AR(1) covariance matrix will be used in the model. The missing values will be imputed using a pattern mixture model utilizing MI assuming MNAR. Before the implementation of the MMRM model, the missing values for a given visit will be imputed using all available values including the retrieved measurement from the post-discontinuation data.

The treatment comparison will be performed using contrasts. The p-value, least square (LS) means and the 95% confidence interval (CI) of the treatment differences will also be provided.

Multiple Imputation

The same 3 steps (imputing missing data, performing MMRM using each imputed dataset, and combine results) will be implemented as described in Section 3.4.1, Primary Analyses. The complete data sets will be constructed using regression model of age, sex, race (white, black, and other), region, baseline BMI, baseline WASO, baseline ISI, baseline WASO, and individual WASO assessments on Days 1, 2 , 29, and 30.

The subgroup analyses including plots (forest plot, and mean and mean change from baseline over time) described in Section 3.4.1.2 and the sensitivity analyses described in Section 3.4.1.3 will be repeated for all key secondary endpoints.

3.4.2.2 Other Secondary Analyses

Unless it is covered from the same model from the primary and secondary efficacy endpoints, or specified otherwise, for all other secondary endpoints, the change from baseline assessments will be analyzed using MMRM assuming MAR (no missing value imputation) and the portion of responders will be analyzed using the Cochran Mantel Haenszel (CMH) test adjusted for age group. Missing values will be considered as non-responders in all responder analyses. No multiplicity adjustment will be made for all analyses.

POLYSOMNOGRAPHY

The following endpoints will be analyzed from PSG:

- Change from baseline of LPS, SE, WASO on Days 1/2 of LEM5 and LEM10 compared to PBO
- Change from baseline of LPS, SE, WASO on Days 1/2 and Days 29/30 of LEM5 and LEM10 compared to ZOL
- Change from baseline of WASO2H on Days 1/2 of LEM5 and LEM10 compared to ZOL
- Change from baseline of WASO2H on Days 1/2 and Days 29/30 of LEM5 and LEM10 compared to PBO
- Change from baseline of WASO on Days 1/2 of LEM5 and LEM10 compared to PBO
- Change from baseline of WASO on Days 1/2 and Days 29/30 of LEM5 and LEM10 compared to ZOL
- Change from baseline of TST on Days 1/2 and Days 29/30 of LEM5 and LEM10 compared to ZOL and PBO
- Proportion of responders on Days 1/2 and Days 29/30 of LEM5 and LEM10 compared to ZOL and PBO in which the responder is defined as follows:

- Objective sleep onset responder: defined as LPS ≤20 minutes provided baseline LPS >30 minutes
- Objective sleep maintenance responder: defined as WASO ≤60 minutes, a reduction from baseline by >10 minutes provided baseline WASO >60 minutes

ELECTRONIC SLEEP DIARY

The following endpoints will be analyzed from the Sleep Diary:

- Change from baseline of mean sSOL, sWASO, sTST, and sSE over the first 7 and last 7 nights of the treatment period of LEM 5 and LEM 10 compared to ZOL and PBO. The derivation of sSOL, sWASO, sTST and sSE is detailed in Appendix 2.
- Proportion of responders over the first 7 and last 7 nights of the treatment period of LEM5 and LEM10 compared to ZOL and PBO in which the responder is defined as follows:
  - Subjective sleep onset responder: defined as sSOL ≤20 minutes and baseline sSOL >30 minutes
  - Subjective sleep maintenance responder: defined as sWASO ≤60 minutes, reduction from baseline by > 10 minutes, and baseline sWASO >60 minutes

POSTURAL STABILITY USING THE CDR POSTURE ASSESSMENT

- Change from baseline of units of body sway on Days 2/3 of the Treatment Period compared to ZOL

INSOMNIA SEVERITY INDEX AND FATIGUE SEVERITY SCALE

The following endpoints will be analyzed from the ISI and FSS:

- Change from baseline of the total score from items 1-7 as well as items 4-7 on the ISI at Day 31 of LEM5 and LEM10 compared to ZOL and PBO
- Change from baseline on the FSS score at Day 31 of LEM5 and LEM10 compared to ZOL and PBO

COGNITIVE PERFORMANCE ASSESSMENT BATTERY

The following endpoints will be analyzed from computerized performance assessment battery (PAB)

- Change from baseline of the 4 composite domain factor scores of PAB (power of attention, mean continuity of attention, mean quality of memory, and mean speed of memory retrieval) on Days 2/3

3.4.3 Other Efficacy/Pharmacodynamic Analyses

The following endpoints are considered exploratory. Comparison of LEM10 and LEM5 will be made with ZOL and PBO.

Unless specified otherwise, for all other efficacy analyses endpoints, the change from baseline assessment will be analyzed using MMRM assuming MAR and the portion of responders will be analyzed using the Cochran Mantel Haenszel test adjusted for age group. Missing values will be considered as non-responders in all responder analyses. No multiplicity adjustment will be made for all analyses.

POLYSOMNOGRAPHY

- Change from baseline of total duration (in minutes) of sleep stage of non-REM (N1, N2, N3 separately and combined) and REM on Days 1/2 and Days 29/30
- Percentage of the change from baseline of total duration of sleep stage of non-REM (N1, N2, N3 separately and combined) and REM
    - per time in bed (TIB) on Days 1/2 and Days 29/30
    - per TST on Days 1/2 and Days 29/30
- Change from baseline of REM latency (defined as the first sleep epoch to first REM sleep epoch) on Days 1/2 and Days 29/30
- The change from baseline of mean REM latency will be analyzed separately for Days 1/2 and for Days 29/30 using Wilcoxon rank sum test. The treatment difference will be estimated using Hodges-Lehmann estimation, and the asymptotic (Moses) 95% CI for the difference will be provided.
- Change from baseline in number of awakenings on Days 1/2 and Days 29/30
- Change from baseline in number of long awakenings (defined as awakenings of 5 minutes or longer) on Days 1/2 and Days 29/30

ELECTRONIC SLEEP DIARY

- Change from baseline of the mean rating on the Quality of Sleep question from the Sleep Diary of the first 7 days and last 7 days of the Treatment Period
- Rebound insomnia endpoints during the Follow-up Period. Rebound insomnia is defined as worsened sleep (ie, higher value of sSOL or sWASO) relative to Screening after study drug treatment is completed.
    - Change from baseline of sSOL on each of the first 3 nights, mean of the first 3 nights, mean sSOL of the first 7 nights, and mean sSOL of the second 7 nights of the Follow-up Period
    - Change from baseline of sWASO on each of the first 3 nights, mean of the first 3 nights, mean sWASO of the first 7 and mean sWASO of the second 7 nights of the Follow-up Period

- Proportion of subjects whose sSOL is longer than at Screening at the following time points of the Follow-up Period by at least 5 minutes: each of the first 3 nights, mean of the first 3 night, mean of the first 7 nights, and mean of the second 7 nights
- Proportion of subjects whose sWASO is higher than at Screening at the following time points of the Follow-up Period by at least 5 minutes: each of the first 3 nights, mean of the first 3 nights, mean of the first 7 nights, and mean of the second 7

The actual value of sSOL and sWASO will be analyzed separately using analysis of covariance model (ANCOVA) with factors of age group (55 to <65, and ≥65 years older), region (North America, and Europe), and treatment for each time point (baseline, each of the first 3 night, mean of the first 3 nights, mean of the first 7 days, and mean of the last 7 days). The 95% CI of the treatment difference will be constructed for each time point. It will be considered as having strong evidence of rebound insomnia if the lower bound of the 95% CI of sSOL or sWASO for each of the 3 night, the mean of the first 3 nights, mean of the first 7 days, and mean of the second 7 nights of the Follow-up Period exceeds the upper bound of a 95% CI for the values during the Screening Period in the given treatment group. If the LS means for sSOL and sWASO for the Follow-up Period are all lower than for the Screening Period, then no rebound insomnia is suggested.

- Mean rating on morning sleepiness over the first 7 mornings and last 7 mornings of the Treatment Period and over the first 7 mornings and last 7 mornings of the Follow-up Period.

MORNING SLEEPINESS QUESTIONNAIRE

- Change from baseline of morning sleepiness ratings on Days2/3, and Days 30/31

POSTURAL STABILITY USING THE CDR POSTURE ASSESSMENT

- Change from baseline of units of body sway on Days 2/3 of the Treatment Period compared to PBO and on Days 30/31 of the Treatment Period compared to ZOL and PBO

COGNITIVE PERFORMANCE ASSESSMENT BATTERY

- Change from baseline of power of attention, continuity of attention, quality of memory, and speed of memory retrieval on Days 30/31

OTHER POLYSOMONGRAPHY ASSESSMENTS

The following endpoints from PSG will also be summarized using frequency count or descriptive statistics by treatment groups for exploratory purpose. No hypothesis testing will be performed on these endpoints.

Unless specified otherwise, the following endpoints will be summarized for Days 1/2 and Days 29/30:

- WASO1H

- Number of awakenings after persistent sleep
- Number and duration of long awakenings after persistent sleep
- Minutes of sleep stages: WASO, non-REM (N1,N2, N3 separately and combined), REM
- Percentage of minutes for each sleep stages per TIB: total wake time, non-REM (N1, N2, N3 separately and combined), REM
- Percentage of minutes for each sleep stages per TST: non-REM (N1, N2, N3 separately and combined), REM
- REM latency (defined as the first sleep epoch to first REM sleep epoch)
- Number of subjects with REM latency within 15 minutes of sleep onset
- WASO by quarter (every 2 hours) of the night

3.5 Pharmacokinetic, Pharmacogenomic, and Other Biomarker Analyses

3.5.1 Pharmacokinetic Analyses

The plasma concentrations of lemborexant and its metabolites M4, M9, and M10, as well as zolpidem (where quantified) will be summarized using descriptive statistics by dose, time and day based on Safety Analysis Set.

A separate analysis plan for the population PK analyses will be developed and finalized before the database lock.

3.5.2 Pharmacokinetic/Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

A separate analysis plan for the PK/PD analyses will be developed and finalized before the database lock.

3.6 Safety Analyses

All safety analyses will be performed based on observed data using the Safety Analysis Set. Safety data will be summarized on an “as treated” basis using descriptive statistics or frequency count only. No hypothesis testing will be performed for safety analyses.

3.6.1 Extent of Exposure

The extent of exposure (mean daily dose, cumulative dose, duration of exposure) to study drug will be summarized using descriptive statistics by treatment group. Duration of exposure of study drug will be defined as the number of days between the date the subject received the first dose of study drug during Treatment Period and the date the subject received the last dose of study drug during Treatment Period, inclusive.

3.6.2 Adverse Events

The adverse event (AE) verbatim descriptions (investigator terms from the CRF) will be classified into standardized medical terminology using the MedDRA. Adverse events will be coded to the MedDRA (Version 20.1 or higher) lower level term closest to the verbatim term. The linked MedDRA PT and primary system organ class (SOC) are also captured in the database.

A treatment-emergent AE (TEAE) is defined as an AE that emerges during treatment (including the Run-In Period up to 14 days after the last dose of study drug from the Treatment Period), having been absent at pretreatment (before the Run-In Period) or

- Reemerges during treatment (including the Run-In Period up to 14 days after the last dose of study drug from the Treatment Period), having been present at pretreatment (before the Run-In Period) but stopped before the last dose of study drug plus 14 days, or
- Worsens in severity during treatment (including the Run-In Period up to 14 days after the last dose of study drug from the Treatment Period) relative to the pretreatment state, when the AE is continuous.

For TEAEs occurred during the Run-in Period, the incidence of TEAEs will be summarized by SOC and PT.

An overview table of TEAE occurred during Treatment Period, including number of subjects with TEAEs, treatment-emergent serious adverse events (SAEs), deaths, severe TEAEs, study drug related TEAEs, TEAEs leading to study drug withdrawal during the Treatment Period will be provided. In addition, the following summaries will be produced for the TEAEs occurred during the Treatment Period:

- Incidence of TEAEs by PT in descending order
- Incidence of TEAEs by SOC and PT
- Incidence of treatment-related TEAEs by SOC and PT
- Incidence of TEAEs by SOC, PT, and severity
- Incidence of treatment-related TEAEs by SOC, PT, and severity
- Incidence of TEAEs by SOC, PT, and relationship to treatment
- Incidence of non-serious TEAEs (>5%) by SOC and PT

If a subject experiences more than one TEAE within a preferred term, the subject will be counted only once in the calculation of incidence of TEAE within that preferred term. Similarly, if a subject experiences more than one TEAE within a SOC, the subject will be counted only once in the calculation of incidence of TEAE within that SOC. If a subject experiences more than one TEAE within a preferred term (or SOC), the occurrence with the highest severity will be used in the calculation of the incidence of TEAE within that preferred term (SOC) by severity. If a subject experiences more than one TEAE within a preferred term (or SOC), the occurrence considered most closely related to study drug will be used in

the calculation of the incidence of TEAE with that preferred term (SOC) by relationship (given by investigator).

The following summaries will also be presented for the treatment-emergent SAEs occurred during the Treatment Period:

- Incidence of treatment-emergent SAEs by SOC and PT
- Incidence of treatment-emergent SAEs by SOC, PT, and relationship to treatment.

In addition, number and percentage of subjects with TEAEs and treatment-related TEAEs leading to discontinuation from study treatment during the Treatment Period will also be summarized by MedDRA SOC, PT for each treatment group.

3.6.2.1 Selected Adverse Events

The following significant AEs will be summarized by SOC and PT:

- Cataplexy
- Falls
- Seizures
- Abuse liability events

Cataplexy includes the TEAEs with MedDRA PT of cataplexy, and drop attack.

Falls includes the TEAEs with MedDRA PT of "fall" only.

Seizure includes TEAEs with MedDRA PTs belonging to MedDRA Standardized MedDRA Query (SMQ) of "Convulsions" (Narrow Terms).

Abuse liability events includes TEAEs with MedDRA PT listed in Appendix 3.

3.6.3 Laboratory Values

Laboratory results will be summarized using Système International (SI) units, as appropriate. With the exception of urinalysis, all quantitative parameters listed in protocol Section 9.5.1.5.5 Laboratory Measurements, the actual value and the change from baseline will be summarized at each visit using descriptive statistics by treatment group. For urinalysis, the actual and the change from baseline of pH and specific gravity will be summarized at each visit by treatment group. Analysis of changes from baseline will be based on the number of subjects with both nonmissing baseline and relevant postbaseline results.

Laboratory test results will be assigned a low- normal-high (LNH) classification according to whether the value was below (L), within (N), or above (H) the laboratory parameter's reference range. Shifts from baseline (LNH) to the Day 31, End of Treatment and the EOS visit will be provided by treatment groups for each laboratory parameter.

The Sponsor's Grading for Laboratory Values (Appendix 1) presents the criteria that will be used to identify subjects with treatment-emergent markedly abnormal laboratory values (TEMAV). Except for phosphate, a TEMAV is defined as a postbaseline value with an increase from baseline to a grade of 2 or higher. For phosphate, a TEMAV was defined as a postbaseline value with an increase from baseline to a grade of 3 or higher. When displaying the incidence of TEMAVs, each subject will be counted once in the laboratory parameter high and in the laboratory parameter low categories, as applicable.

3.6.4 Vital Signs

For each vital signs parameters (ie, diastolic and systolic BP, pulse, respiration rate, temperature) and weight, the actual value and changes from Study Baseline will be summarized by treatment group at each visit using descriptive statistics. Analysis of changes from baseline will be based on the number of subjects with both nonmissing baseline and relevant postbaseline results.

In addition, clinically notable vital sign values will be identified using the criteria in Table 2. The clinically notable vital sign values will be summarized using frequency count at each visit by treatment group.

Table 2 Vital Sign Criteria

	Criterion value[a]	Change relative to baseline[a]	Clinically notable range
Heart rate	>120 bpm	Increase of 15 bpm	H
	<50 bpm	Decrease of ≥15 bpm	L
Systolic BP	>180 mmHg	Increase of ≥20 mmHg	H
	<90 mmHg	Decrease of ≥20 mmHg	L
Diastolic BP	>105 mmHg	Increase of ≥15 mmHg	H
	<50 mmHg	Decrease of ≥15 mmHg	L
Weight	--	Increase of ≥7%	H
	--	Decrease of ≥7%	L
Respiratory Rate	>20 bpm	--	H
	< 10 bpm	--	L

BP = blood pressure, H = high, L = low.

a. Clinically notable means that a value must meet the criterion value and must attain the specified magnitude of change relative to baseline.

3.6.5 Electrocardiograms

For each ECG parameters (including PR interval, RR interval, QRS interval, QT interval, QTcB interval, QTcF interval and heart rate) and actual value and changes from baseline will be summarized by treatment group at each visit using descriptive statistics. Shift tables from

baseline to the Day 31, End of Treatment and the EOS visits will be presented by treatment group for ECG interpretation (categorized as normal and abnormal).

In addition, maximum postbaseline measurement will also be tabulated by treatment group as follows:

- Number and percentage of subjects with QTcF of >450 msec, and >500 msec during the treatment
- Number and percentage of subjects with a QTcF increment of >30 msec, and >60 msec from the baseline visit.
- Number and percentage of subjects with PR of >220 msec
- Number and percentage of subjects with QRS of >120 msec

3.6.6 Other Safety Analyses

3.6.6.1 Columbia-Suicide Severity Rating Scale (C-SSRS)

Suicidality will be assessed using a self-rated electronic version of the C-SSRS (eC-SSRS). The eC-SSRS assesses an individual's degree of suicidality, including both suicidal ideation and suicidal behavior. The incidence of suicidal ideation, suicidal behavior, and self-injurious non-suicidal behavior at each visit will be summarized by treatment group using frequency count. A subject will be counted once in a category if at least one question is answered positive in the category.

3.6.6.2 Tyrer Benzodiazepine Withdrawal Symptom Questionnaire (T-BWSQ)

Withdrawal symptoms will be assessed using the T-BWSQ at the EOS visit. Subjects will be asked about the presence/absence and severity of the symptoms listed in the questionnaire. For each listed symptom, the subject is to respond "No" (Score = 0), "Yes – moderate" (Score = 1) or "Yes – severe" (Score = 2). The sum of responses will be the subject's total score. The total score will be summarized by treatment group using descriptive statistics. In addition, the number and percentage of subjects with a total score of ≥3 will be summarized using frequency count.

3.7 Other Analyses

3.7.1 Health Outcome Economics Analyses

3.7.1.1 EQ-5D-3L

The EQ-5D-3L instrument comprises questions on 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a visual analogue score (EQ VAS). Each dimension has 3 levels: no problem, some problems, extreme problems and the EQ VAS is ranged from 0 ("Worst imaginable health state") to 100 ("Best imaginable health

state"). Each dimension score will be summarized separately at Baseline and Day 31 using frequency count on observed data only with no imputation. The change from baseline of EQ VAS will be analyzed using ANCOVA with factors of age group (55 to 64, and ≥65 years old), region (North America, and Europe), treatment based on FAS and baseline VAS as a covariate.

3.7.1.2 Patient Global Impression (PGI) - Insomnia

The PGI-Insomnia questionnaire captures the global impression of the study medication's effect at the end of treatment and is collected on Day 31 visit only. The PGI-Insomnia has 3 items related to study medication effect (helped/worsened sleep, decreased/increased time to fall asleep, and increased/decreased TST) on a 3-point scale (1=positive medication effect, 2=neutral medication effect, and 3=negative medication effect) and 1 item related to perceived appropriateness of study mediation strength also on a 3-point scale (medication: 1=too strong, 2=just right, and 3=too weak). Each item will be analyzed summarized separately ("positive medication effect" versus others for the first 3 item; "just right" versus others for the last item) using chi-square test on observed data only based on FAS with no imputation for missing values, and repeated for age subgroups.

3.8 Exploratory Analyses

None

4 INTERIM ANALYSES

An interim analysis is planned to be conducted after approximately 50% of subjects (approximately n=475 subjects) have been randomized and either completed Day 31 assessments or discontinued from the study. This interim analysis will be conducted for administrative reasons as detailed in the separate Interim Analysis Charter. When the specified number of subjects has completed the Day 31 assessments, an independent statistician external to the Sponsor will be provided with the relevant PSG dataset and will be unblinded to the primary endpoint, ie, change from baseline in WASO2H for the mean of Days 29 and 30. A conditional power will be calculated to predict the probability that the trial will achieve a significant treatment effect for WASO2H in the LEM10 versus ZOL arms at the end of the study, given what is observed at the time of interim analysis. The interim analysis will be limited to the comparison of LEM10 versus ZOL on the change from baseline in WASO2H for the mean of Days 29 and 30. No other endpoints, dose groups, or timepoints will be analyzed at the interim analysis. The study will not be terminated for either futility or efficacy. Therefore no impact to the type I error rate is expected.

The method of calculating the conditional power will be detailed in the Interim Analysis Charter, along with operational procedures, unblinding procedures, procedures for communicating the results of the conditional power calculation and recipients of this information. To preclude potential influence on the conduct of the remainder of the study, disclosure of the interim results will be limited to a prespecified set of executive-level individuals at the sponsor and sponsor's co-development partner. No individuals involved

with the conduct of the study will have access to the interim data or the results of the interim analysis (i.e., the conditional power of LEM10 versus ZOL on the change from baseline in WASO2H for the mean of Days 29 and 30).

Enrollment of subjects will not be stopped during the interval during which the interim analysis is conducted. The interim analysis may be waived or otherwise not conducted, for reasons including but not limited to a higher than anticipated enrollment rate which would make the interim analysis unnecessary as the majority of subjects would have been enrolled by the time the interim analysis is concluded.

5 CHANGES IN THE PLANNED ANALYSES

The following changes were made in version 2.0 of the SAP from version 1.0:

- Section 3.3.5 "Handling of Missing Data, Dropouts, and Outlier", Section 3.3.6 "Other Considerations", Section 3.4 "Efficacy/Pharmacodynamic Analyses": Sensitivity analyses are added to evaluate different missing value patterns. Sensitivity and subgroup analyses are added to key secondary efficacy endpoints for completeness.
- Section 3.6.2.1 "Selected Adverse Events" and Appendix 13.3 "List of Abuse Liability Events": The definition of "Abuse Lability Events" is updated to utilize MedDRA SMQs. Appendix 13.3 from version 1.0 is removed from this section.
- Section 6.1 "Visit Window": Visit window description is added for diary efficacy endpoints.
- Throughout the document: Editorial comments are made to correct typos or for clarification purposes.

The following changes were made in version 3.0 of the SAP from version 2.0:

- Section 3.1.2"Secondary Endpoints", Section 3.3.3 "Multiple Comparisons/Multiplicity": The order of the objectives and endpoints is updated to incorporate feedback from regulatory authorities.
- Section 3.2.1 "Definitions of Analysis Sets": List of exclusion reasons from the PP population is updated.
- Section 3.6.2.1 "Selected Adverse Events" and Appendix 3 "List of Abuse Liability Events": The definition of "Abuse Lability Events" is updated to revert back to the approach from version 1.0.
- Section 7 "Programming Specifications": The description of a planned tipping point analysis has been included.
- Throughout the document: Editorial comments are made to correct typos or for clarification purposes.

6 DEFINITIONS AND CONVENTIONS FOR DATA HANDLING

6.1 Visit Window

Study Day 1 is defined as the date of the first dose of study drug during the Treatment Period. The nominal visit (ie, study visit captured on the CRF) will be used as the analysis visits in all by-visit summaries except for sleep diary efficacy endpoints. The Early Term visit will be considered as unscheduled visit and will not be included in the by-visit summary. Where applicable, the Early Term visit will be used along with the Day 31 visit for completers as the End of Treatment visit for the safety analyses.

For diary efficacy endpoints, the following visit window will be applied:

Timepoint	Visit Window (in study days)
First 7 days of Treatment	2-8
Last 7 days of Treatment	22-36[a]

a: Last seven days within this window while on treatment

6.2 Baseline Assessment

Unless otherwise specified, baseline measurement is the last observed measurement, including unscheduled assessments, prior to the first dose of study medication of treatment period for a given assessment. For the following endpoints, baseline measurement is defined as follows:

- PSG parameters : average of the two PSG recordings during the Run-in Period
- Sleep diary parameters:
  - For rebound insomnia: the mean of diary data entered on the last 7 mornings before the Screening PSG during the Screening Period
  - Other Sleep Diary-derived endpoints: the mean of diary data entered on the last 7 mornings before the first Baseline PSG during the Run-In Period
- Morning sleepiness questionnaire at 1.5 hours after wake time on mornings after PSG recordings: Average of the 2 morning sleepiness ratings during the Run-in Period
- ISI: Last available ISI measurement on or prior to Visit 3
- Postural Stability parameters: Average of non-missing measurements from Visit 3 Visit 4
- Cognitive PAB parameters: Average of non-missing measurements from Visit 3 and Visit 4
- FSS: Last available FSS measurement on or prior to Visit 5
- EQ-5D-3L: Last available EQ-5D-3L measurement on or prior to Visit 5
- C-SSRS: Visit 5 (“since last visit” form)

6.3 Missing Data Handling

Unless stated otherwise, missing values will be considered as non-responders in responder analyses and the continuous variables will be analyzed using MMRM to handle the missing values assuming MAR in all other efficacy analyses. Details can be found in Section 3.4

All safety analyses will be performed based on the observed data only.

6.3.1 Polysomnography, Cognitive Performance Assessment, Posture Stability, and Morning Sleepiness Questionnaire

Each PSG, PAB, posture stability, and morning sleepiness questionnaire parameters will be derived by calculating the averages of pairs of values, i.e., the average of the two PSG recordings during the Run-in Period, Day 1 and Day 2, and Day 29 and Day 30. If one of each pair of values is missing, the other available value will be taken as the average of the pair; if both values are missing, then the parameter will be missing for the corresponding pair.

6.3.2 Sleep Diary

Each Sleep Diary parameter will be derived by calculating the average of weekly (7 days) diary parameter values. For the follow-up period, if the first 7 nights overlaps with the last 7 nights (eg, the follow-up period is less than 14 days in total), the last non-overlaps nights will be used in calculating the average value for the last 7 nights.

For each Sleep Diary parameter at baseline, if no more than 2 of the 7 nights' values are missing, the available values will be used to calculate the mean. If more than 2 values are missing, the parameter will be considered missing for baseline. For each Sleep Diary parameter during treatment period and follow-up period, if at least 4 of the 7 nights' values are available, the available values will be used to calculate the mean. If less than 4 values are available, the parameter will be considered missing for the corresponding time point.

7 PROGRAMMING SPECIFICATIONS

The rules for programming derivations and dataset specifications are provided in separate documents.

The following sample SAS statement provides the framework for the MI method:

CONVERT DATASET INTO MONOTONE MISSING DATA PATTERN (IMPUTING ARBITRARY MISSING DATA):

```
PROC MI data=<dataset> nimpute=30 seed=2359 out=<dataset1>;
  VAR age BMI baseline… visit1-visit4;
  MCMC chain=multiple nbiter=500 niter=300 impute=monotone;
  BY treatment;
RUN;
```

IMPUTE MISSING VALUES:

```
PROC MI data=<dataset1> nimpute=1 seed=2359 out=<dataset2>;
  CLASS treatment sex race region;
  MONOTONE regression ( /details);
  MNAR model (visit1-visit4/ modelobs=CCMV);
  VAR treatment age sex race region BMI baseline….;
  BY _imputation_;
RUN;
```

PERFORMING MMRM:

```
PROC MIXED data=<dataset2>;
  CLASS subject treatment agegrp visit;
  MODEL value=treatment agegrp region visit visit*treatment / ddfm=kr;
  REPEAT visit/sub=subject type=UN group=treatment;
  LSMEANS visit*treatment;
  ESTIMATE '5mg – ZOL Days 1_2'     treatment 0 -1 1 0  visit*treatment   0 0 -1 0 1 0 0 0/CL;
  ESTIMATE '10mg – ZOL Days 1_2'    treatment 0 -1 0 1  visit*treatment   0 0 -1 0 0 0 1 0/CL;
  ESTIMATE '5mg – ZOL Days 29_30'   treatment 0 -1 1 0  visit*treatment   0 0 0 -1 0 1 0 0/CL;
  ESTIMATE '10mg – ZOL Days 29_30'  treatment 0 -1 0 1  visit*treatment   0 0 0 -1 0 0 0 1/CL;
  BY _imputation_;
  ODS output estimates=<dataset3>;
RUN;
```

COMBINE RESULTS:

```
PROC MIANALYZE data=<dataset3>;
  MODELEFFECTS estimate;
  STDERR stderr;
RUN;
```

VARIABLE ORDER TO BE USED IN THE PROC MI PROCEDURES:

To Create Monotone Missing Data Pattern

- LPS: age, baseline BMI, baseline ISI, baseline log(sSOL), baseline log(LPS), log(LPS) at Day1, log(LPS) at Day2, log(LPS) at Day29, log(LPS) at Day30
- SE: age, baseline BMI, baseline ISI, baseline sSE, baseline SE, SE at Day1, SE at Day2, SE at Day29, SE at Day30
- WASO2H: age, baseline BMI, baseline ISI, baseline sWASO, baseline WASO2H, WASO2H at Day1, WASO2H at Day2, WASO2H at Day29, WASO2H at Day30
- WASO: age, baseline BMI, baseline ISI, baseline sWASO, baseline WASO, WASO at Day1, WASO at Day2, WASO at Day29, WASO at Day30

To Impute Missing Values

- LPS: treatment, age, sex, race, region, baseline BMI, baseline ISI, baseline log(sSOL), baseline log(LPS), log(LPS) at Day1, log(LPS) at Day2, log(LPS) at Day29, log(LPS) at Day30
- SE: treatment, age, sex, race, region, baseline BMI, baseline ISI, baseline sSE, baseline SE, SE at Day1, SE at Day2, SE at Day29, SE at Day30
- WASO2H: treatment, age, sex, race, region, baseline BMI, baseline ISI, baseline sWASO, baseline WASO2H, WASO2H at Day1, WASO2H at Day2, WASO2H at Day29, WASO2H at Day30
- WASO: treatment, age, sex, race, region, baseline BMI, baseline ISI, baseline sWASO, baseline WASO, WASO at Day1, WASO at Day2, WASO at Day29, WASO at Day30

TIPPING POINT SENSITIVITY ANALYSIS:

The following sample SAS statements and algorithm provide the framework for the Tipping Point Sensitivity Analysis:

A tipping point sensitivity analysis will be conducted on the endpoints LPS (Log LPS), SE, WASO and WASO2H using the multiple imputation methodology as described in the section above but with the following modifications:

1. The second MI procedure (monotone missing values) is to be modified to introduce an adjustable shift (i.e sensitivity parameter) to the imputed values for only the treatment groups LEM10 and LEM5, corresponding to a MAR assumption when the shift is zero. These shifts are to be applied to Day29 and Day30 only.

```
PROC MI data=<dataset1> nimpute=1 seed=2359 out=<dataset2>;
    CLASS treatment sex race region;
    MONOTONE regression ( /details);
    MNAR adjust (visit3/ shift=<shift> adjustobs=(treatment=LEM5));
    MNAR adjust (visit4/ shift=<shift> adjustobs=(treatment= LEM5));
    MNAR adjust (visit3/ shift=<shift> adjustobs=(treatment=LEM10));
    MNAR adjust (visit4/ shift=<shift> adjustobs=(treatment= LEM10));
    VAR treatment age sex race region BMI baseline… V5 V6 V7 V8;
    BY _imputation_;
RUN;
```

2. If the MAR (shift=0) model has a p-value that is significant (<0.05), then the <shift> value will be systematically incremented until the resulting p-value is >=0.05.
3. Step 2 may be repeated iteratively starting with the shift found just prior to p-value >0.05 ending with the p-value>=0.05 found in Step 2, using smaller increments, until a shift is found where the rounded p-value has a value of 0.05 to a reasonable accuracy.
4. The values for <shift> in Step 1 will be applied uniformly (same shift) to both LEM 5mg and LEM 10gm at both Day29 and Day 30.

5. The values for <shift> will correspond to worsening values for the endpoint according to the following:

- Increasing positive shift values (LPS, WASO, WASO2H)
- Increasing negative shift values (SE)

6. The following specifies which comparisons are of interest when evaluating p-values for this procedure:

- LEM 5mg and LEM10mg vs. Placebo (LPS, SE, WASO)
- LEM 5mg and LEM 10mg vs. Zolpidem (WASO2H)

8 STATISTICAL SOFTWARE

Statistical analyses will be performed using SAS version 9.4 (or later versions). In the event that certain features graphical analyses cannot be implemented by SAS, other statistical software such as Splus can be employed.

The conditional power calculated for the interim analysis will be performed using EAST® version 6 (or later versions).

9 MOCK TABLES, LISTINGS, AND GRAPHS

The study tables, listings and graphs shells will be provided in a separate document, which will show the content and format of all tables, listings, and graphs in detail.

10 REFERENCES

ICH Final Concept Paper E9(R1): Addendum to statistical principles for clinical trials on choosing appropriate estimands and defining sensitivity analyses in clinical trials dated 22October 2014.

Mallinckrodt CH, Lin Q, Lipkovich I, Molenberghs G. A structured approach to choosing estimands and estimators in longitudinal clinical trials. Pharmaceutical Statistics 2012,11:456-461, 10 September 2012.

Rubin, DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons; 1987.

Appendix 1 Sponsor's Grading for Determining Markedly Abnormal Laboratory Results

	Grade 1	Grade 2	Grade 3	Grade 4
BLOOD/BONE MARROW				
Hemoglobin	<LLN – 10.0 g/dL <LLN – 100 g/L <LLN – 6.2 mmol/L	<10.0 – 8.0 g/dL <100 – 80 g/L <6.2 – 4.9 mmol/L	<8.0 g/dL <80 g/L <4.9 mmol/L; transfusion indicated	life-threatening consequences; urgent intervention indicated
Leukocytes (total WBC)	<LLN – 3.0×10^9/L <LLN – 3000/mm^3	<$3.0 – 2.0\times10^9$/L <3000 – 2000/mm^3	<$2.0 – 1.0\times10^9$/L <2000 – 1000/mm^3	<1.0×10^9/L <1000/mm^3
Lymphocytes	<LLN – 800/mm^3 <LLN – 0.8×10^9/L	<800 – 500/mm^3 <$0.8 – 0.5\times10^9$/L	<500 – 200/mm^3 <$0.5 – 0.2\times10^9$/L	<200/mm^3 <0.2×10^9/L
Neutrophils	<LLN – 1.5×10^9/L <LLN – 1500/mm^3	<$1.5 – 1.0\times10^9$/L <1500 – 1000/mm^3	<$1.0 – 0.5\times10^9$/L <1000 – 500/mm^3	<0.5×10^9/L <500/mm^3
Platelets	<LLN – 75.0×10^9/L <LLN – 75,000/mm^3	<$75.0 – 50.0\times10^9$/L <75,000 – 50,000/mm^3	<$50.0 – 25.0\times10^9$/L <50,000 – 25,000/mm^3	<25.0×10^9/L <25,000/mm^3
METABOLIC/LABORATORY				
Albumin, serum- low (hypoalbuminemia)	<LLN – 3 g/dL <LLN – 30 g/L	<3 – 2 g/dL <30 – 20 g/L	<2 g/dL <20 g/L	life-threatening consequences; urgent intervention indicated
Alkaline phosphatase	>ULN – 3.0×ULN	>3.0 – 5.0×ULN	>5.0 – 20.0×ULN	>20.0×ULN
ALT	>ULN – 3.0×ULN	>3.0 – 5.0×ULN	>5.0 – 20.0×ULN	>20.0×ULN
AST	>ULN – 3.0×ULN	>3.0 – 5.0×ULN	>5.0 – 20.0×ULN	>20.0×ULN
Bilirubin (hyperbilirubinemia)	>ULN – 1.5×ULN	>1.5 – 3.0×ULN	>3.0 – 10.0×ULN	>10.0×ULN
Calcium, serum-low (hypocalcemia)	<LLN – 8.0 mg/dL <LLN – 2.0 mmol/L	<8.0 – 7.0 mg/dL <2.0 – 1.75 mmol/L	<7.0 – 6.0 mg/dL <1.75 – 1.5 mmol/L	<6.0 mg/dL <1.5 mmol/L
Calcium, serum-high (hypercalcemia)	>ULN – 11.5 mg/dL >ULN – 2.9 mmol/L	>11.5 – 12.5 mg/dL >2.9 – 3.1 mmol/L	>12.5 – 13.5 mg/dL >3.1 – 3.4 mmol/L	>13.5 mg/dL >3.4 mmol/L
Cholesterol, serum-high (hypercholesterolemia)	>ULN – 300 mg/dL >ULN – 7.75 mmol/L	>300 – 400 mg/dL >7.75 – 10.34 mmol/L	>400 – 500 mg/dL >10.34 – 12.92 mmol/L	>500 mg/dL >12.92 mmol/L
Creatinine	>ULN – 1.5×ULN	>1.5 – 3.0×ULN	>3.0 – 6.0×ULN	>6.0×ULN
GGT (γ-glutamyl transpeptidase)	>ULN – 3.0×ULN	>3.0 – 5.0×ULN	>5.0 – 20.0×ULN	>20.0×ULN
Glucose, serum-high (hyperglycemia)	Fasting glucose value: >ULN – 160 mg/dL >ULN – 8.9 mmol/L	Fasting glucose value: >160 – 250 mg/dL >8.9 – 13.9 mmol/L	>250 – 500 mg/dL; >13.9 – 27.8 mmol/L; hospitalization indicated	>500 mg/dL; >27.8 mmol/L; life-threatening consequences
Glucose, serum-low (hypoglycemia)	<LLN – 55 mg/dL <LLN – 3.0 mmol/L	<55 – 40 mg/dL <3.0 – 2.2 mmol/L	<40 – 30 mg/dL <2.2 – 1.7 mmol/L	<30 mg/dL <1.7 mmol/L life-threatening consequences; seizures

	Grade 1	Grade 2	Grade 3	Grade 4
Phosphate, serum-low (hypophosphatemia)	<LLN – 2.5 mg/dL <LLN – 0.8 mmol/L	<2.5 – 2.0 mg/dL <0.8 – 0.6 mmol/L	<2.0 – 1.0 mg/dL <0.6 – 0.3 mmol/L	<1.0 mg/dL <0.3 mmol/L life-threatening consequences
Potassium, serum-high (hyperkalemia)	>ULN – 5.5 mmol/L	>5.5 – 6.0 mmol/L	>6.0 – 7.0 mmol/L hospitalization indicated	>7.0 mmol/L life-threatening consequences
Potassium, serum-low (hypokalemia)	<LLN – 3.0 mmol/L	<LLN – 3.0 mmol/L; symptomatic; intervention indicated	<3.0 – 2.5 mmol/L hospitalization indicated	<2.5 mmol/L life-threatening consequences
Sodium, serum-high (hypernatremia)	>ULN – 150 mmol/L	>150 – 155 mmol/L	>155 – 160 mmol/L hospitalization indicated	>160 mmol/L life-threatening consequences
Sodium, serum-low (hyponatremia)	<LLN – 130 mmol/L	N/A	<130 – 120 mmol/L	<120 mmol/L life-threatening consequences
Triglyceride, serum-high (hypertriglyceridemia)	150 – 300 mg/dL 1.71 – 3.42 mmol/L	>300 – 500 mg/dL >3.42 – 5.7 mmol/L	>500 – 1000 mg/dL >5.7 – 11.4 mmol/L	>1000 mg/dL >11.4 mmol/L life-threatening consequences
Uric acid, serum-high (hyperuricemia)	>ULN – 10 mg/dL ≤0.59 mmol/L without physiologic consequences	N/A	>ULN – 10 mg/dL ≤0.59 mmol/L with physiologic consequences	>10 mg/dL >0.59 mmol/L life-threatening consequences

ALT = alanine aminotransferase (serum glutamic pyruvic transaminase), AST = aspartate aminotransferase (serum glutamic oxaloacetic transaminase), GGT = γ-glutamyl transpeptidase, N/A = not applicable, LLN = lower limit of normal, ULN = upper limit of normal, WBC = white blood cell.
Based on Common Terminology Criteria for Adverse events (CTCAE) Version 4.0. Published: May 28, 2009 (v4.03: June 14, 2010).

Appendix 2 Derivations of Efficacy Endpoints from Electronic Sleep Diary

The following 7 questions are captured in the electronic Sleep Diary:

Q1: What time did you try to go to sleep?

Q2: How long did it take you to fall asleep?

Q3: How many times did you wake up, not counting your final awakening?

Q4: In total, how long did these awakenings last?

Q5: What time was your final awakening?

Q6: After your last awakening, how much longer did you try to sleep?

Q7: What time did you get out of bed for the day?

The efficacy endpoints from electronic Sleep Diary are defined as follows:

- sSOL= Q2
- sWASO = Q4 + Q7 – Q5
- sTST = TIB – time spent awake [where TIB = Q7 – Q1; and time spent awake = Q2 + Q4 + Q7 – Q5]
- sSE = sTST/TIB (as defined above)

Appendix 3 List of Abuse Liability Events

Code	PT
10061422	Abnormal behaviour
10000125	Abnormal dreams
10063746	Accidental death
10000381	Accidental overdose
10000383	Accidental poisoning
10001022	Acute psychosis
10054196	Affect lability
10001443	Affective disorder
10001488	Aggression
10001497	Agitation
10001666	Alice in wonderland syndrome
10001854	Altered state of consciousness
10053549	Altered visual depth perception
10001949	Amnesia
10061423	Amnestic disorder
10002368	Anger
10002511	Anhedonia
10002711	Anterograde amnesia
10002820	Antisocial behaviour
10002855	Anxiety
10002942	Apathy
10003472	Asocial behaviour
10003739	Attention-seeking behaviour
10049848	Balance disorder
10004224	Belligerence
10005885	Blunted affect
10050012	Bradyphrenia
10057668	Cognitive disorder
10061046	Communication disorder
10010219	Compulsions
10010297	Confabulation
10067494	Confusional arousal
10010305	Confusional state

10050093	Consciousness fluctuating
10010947	Coordination abnormal
10012177	Deja vu
10012218	Delirium
10012239	Delusion
10012335	Dependence
10077805	Depersonalisation/derealisation disorder
10012374	Depressed mood
10012378	Depression
10012411	Derailment
10012422	Derealisation
10013142	Disinhibition
10013395	Disorientation
10013457	Dissociation
10013462	Dissociative disorder
10013468	Dissociative identity disorder
10013496	Disturbance in attention
10061108	Disturbance in social behaviour
10013573	Dizziness
10061111	Drug abuser
10013659	Drug administered at inappropriate site
10052237	Drug detoxification
10066053	Drug diversion
10052804	Drug tolerance
10052806	Drug tolerance increased
10079381	Drug use disorder
10013752	Drug withdrawal convulsions
10013753	Drug withdrawal headache
10013754	Drug withdrawal syndrome
10013887	Dysarthria
10054940	Dyslogia
10014551	Emotional disorder
10049119	Emotional distress
10048779	Energy increased
10015535	Euphoric mood
10070246	Executive dysfunction

10016256	Fatigue
10016275	Fear
10016322	Feeling abnormal
10016330	Feeling drunk
10016338	Feeling jittery
10016344	Feeling of despair
10016352	Feeling of relaxation
10016754	Flashback
10016759	Flat affect
10016777	Flight of ideas
10017062	Formication
10019063	Hallucination
10019070	Hallucination, auditory
10019072	Hallucination, olfactory
10062824	Hallucination, synaesthetic
10019074	Hallucination, tactile
10019075	Hallucination, visual
10019079	Hallucinations, mixed
10019133	Hangover
10020400	Hostility
10048533	Hypervigilance
10020937	Hypoaesthesia
10021212	Ideas of reference
10021402	Illogical thinking
10021403	Illusion
10049564	Impaired driving ability
10071176	Impaired reasoning
10049976	Impatience
10021567	Impulsive behaviour
10021588	Inappropriate affect
10021630	Incoherent
10021703	Indifference
10022523	Intentional overdose
10074903	Intentional product misuse
10023118	Jamais vu
10023236	Judgement impaired

10024264	Lethargy
10024825	Loose associations
10025429	Magical thinking
10026749	Mania
10027175	Memory impairment
10061284	Mental disorder
10027374	Mental impairment
10048294	Mental status changes
10027940	Mood altered
10027951	Mood swings
10028330	Muscle rigidity
10028747	Nasal necrosis
10028765	Nasal septum perforation
10028766	Nasal septum ulceration
10028896	Needle track marks
10061862	Neonatal complications of substance abuse
10029216	Nervousness
10029412	Nightmare
10033295	Overdose
10033664	Panic attack
10033670	Panic reaction
10033775	Paraesthesia
10033848	Paramnesia
10033864	Paranoia
10061910	Parasomnia
10063117	Paroxysmal perceptual alteration
10034719	Personality change
10061355	Poisoning
10067669	Prescription form tampering
10069330	Product tampering
10070592	Product used for unknown indication
10037211	Psychomotor hyperactivity
10037213	Psychomotor retardation
10049215	Psychomotor skills impaired
10061920	Psychotic disorder
10053632	Reactive psychosis

10038001	Rebound effect
10038743	Restlessness
10038965	Retrograde amnesia
10039897	Sedation
10040026	Sensory disturbance
10061567	Sensory level abnormal
10041052	Sluggishness
10041317	Somatic delusion
10062684	Somatic hallucination
10041349	Somnolence
10041953	Staring
10042264	Stupor
10067688	Substance abuser
10070964	Substance use
10079384	Substance use disorder
10072387	Substance-induced mood disorder
10072388	Substance-induced psychotic disorder
10042635	Suspiciousness
10043114	Tangentiality
10043431	Thinking abnormal
10043495	Thought blocking
10052214	Thought broadcasting
10043496	Thought insertion
10043497	Thought withdrawal
10070863	Toxicity to various agents
10044380	Transient global amnesia
10056326	Transient psychosis
10049414	Treatment noncompliance
10048010	Withdrawal syndrome
MedDRA 20.1	

SIGNATURE PAGE

Author(s):

[electronic signature in eDMS]

PPD

PPD

Neuroscience Business Group, Eisai Inc.

Date

Approval:

[electronic signature in eDMS]

PPD

PPD

Neuroscience Business Group, Eisai Inc.

Date

[electronic signature in eDMS]

PPD

PPD

Neuroscience Business Group, Eisai Inc.

Date

[electronic signature in eDMS]

PPD

PPD

Neuroscience Business Group, Eisai Inc.

Date