
# STATISTICAL ANALYSIS PLAN

| | |
|---|---|
| **Study Protocol Number:** | E2006-G000-303 |
| **Study Protocol Title:** | A Long-Term Multicenter, Randomized, Double-Blind, Controlled, Parallel-Group Study of the Safety and Efficacy of Lemborexant in Subjects With Insomnia Disorder |
| **Date:** | 21 August 2018 |
| **Version:** | Final Version 2.0 |

# TABLE OF CONTENTS

## LIST OF IN-TEXT TABLES



## LIST OF IN-TEXT FIGURES



## LIST OF APPENDICES

**Table 1 Summary of Changes (Version 2)**

| Change | Rationale |
|---|---|
| Throughout the document: Editorial comments are made to correct typos | For clarification purposes. |
| Section 3 Added wording on multiple database locks and reporting | Due to changes in planned reporting. |
| Section 3.1.2 "Secondary Endpoints" and Section 3.1.3 "Exploratory Endpoints" | Wording adjusted to look at endpoints for duration of exposure, not timepoints during Period 2 |
| Section 3.2.1 "Definitions of Analysis Sets": The definition of 6-Months Completer Analysis Set, On-Treatment Safety Analysis Set and On-Treatment Full Analysis Set are added. | Detail added. |
| Section 3.4 "Efficacy Analyses" | Description of 12-month LEM exposure summaries detailed |
| Section 3.4.1.4 "Supplementary Analyses" and Section 3.4.2.3 " Supplementary Analyses for Key Secondary Efficacy Analyses": | The supplementary analysis is added to incorporate feedback from regulatory authorities. |
| Section 3.4.1.4 "Other Secondary Efficacy Measures" | Clarity added on responder and persistence of loss of effect. Description of 12-month LEM exposure summaries detailed. FSS responder analysis added. |
| Section 3.6.1 "Treatment Definitions" | Description of 12-month LEM exposure summaries detailed. Up to 12 month and up to 6 month treatment definitions removed. |
| Section 3.6.2 "Extent of Exposure" | Updated to include 9 month exposure summary. |
| Section 3.6.3 "Adverse Events" | Up to 12 month and up to 6 month treatment definitions removed. |
| Section 3.6.3.1 "Selected Adverse Events": The summary of potential cataplexy and its definition (Appendix 2) are added. | Detail of definition of potential cataplexy added. Up to 12 month and up to 6 month treatment definitions removed. |
| Section 3.6.3.2 "Subgroup Analysis of TEAE" and Section 3.6.4.1 "Subgroup Analysis of Laboratory values" are added. | Further subgroups for safety analyses added. Up to 12 month and up to 6 month treatment definitions removed. |
| 3.6.4 "Laboratory Values", 3.6.4.1 "Subgroup Analysis of Laboratory Values", | Description of 12-month LEM exposure summaries detailed. |

**Table 1 Summary of Changes (Version 2)**

| Change | Rationale |
|---|---|
| 3.6.5 "Vital Signs", 3.6.6 "Electrocardiograms", 3.7.1 "EQ-5D-3L", 3.7.2 "Patient Global Impression (PGI) – Insomnia", 3.7.2.1 "WPAI-GH" | |
| Section 7 "Programming Specifications": The description of a planned tipping point analysis has been included | Further detail of tipping point analysis required. |

# LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

| Abbreviation | Term |
|---|---|
| AE | adverse event |
| ANCOVA | analysis of covariance |
| ATC | anatomical therapeutic class |
| AUC | area under the concentration-time curve |
| BAI | Beck Anxiety Inventory |
| BDI-II | Beck Depression Inventory - II |
| BMI | body mass index |
| CI | confidence interval |
| CL | total clearance |
| $C_{max}$ | maximum observed concentration |
| CMH | Cochran-Mantel-Haenszel |
| CSR | clinical study report |
| CV | coefficient of variation |
| ECG | electrocardiogram |
| ESS | Epworth Sleepiness Scale |
| eCRF | electronic case report form |
| eC-SSRS | electronic Columbia-Suicide Severity Rating Scale |
| EOS | end of study |
| ESS | Epworth Sleepiness Scale |
| FAS | full analysis set |
| FDA | Food and Drug Administration |
| FSS | Fatigue Severity Scale |
| IRLS | International Restless Legs Scale |
| ISI | Insomnia Severity Index |
| LDA | longitudinal data analysis |
| LEM5 | lemborexant, 5-mg dose |
| LEM10 | lemborexant, 10-mg dose |
| LNH | low-normal-high |
| LS | least squares |

| Abbreviation | Term |
|---|---|
| MAR | missing at random |
| MCMC | Monte Carlo Markov Chain |
| MI | Multiple Imputation |
| MNAR | missing not at random |
| MedDRA | Medical Dictionary for Regulatory Activities |
| NREM | non-REM sleep |
| PAB | performance assessment battery |
| PBO | placebo |
| PD | pharmacodynamic |
| PGI | Patient Global Impression |
| PK | pharmacokinetic |
| PK/PD | pharmacokinetic/pharmacodynamic |
| PT | preferred term |
| QTcB | QT interval calculated using Bazett's formula |
| QTcF | QT interval corrected for heart rate by Fridericia's formula |
| SAE | serious adverse event |
| SAP | statistical analysis plan |
| SD | Standard deviation |
| SE | sleep efficiency |
| SOC | system organ class |
| SDSB | Sleep Disorders Screening Battery |
| sSE | subjective sleep efficiency |
| sSOL | subjective sleep onset latency |
| sTIB | subjective time in bed |
| sTST | subjective total sleep time |
| sWASO | subjective wake after sleep onset |
| T-BWSQ | Tyrer Benzodiazepine Withdrawal Symptom Questionnaire |
| TEAE | treatment-emergent adverse event |
| TIB | time in bed |
| TST | total sleep time |

| Abbreviation | Term |
|---|---|
| US | United States |
| VAS | visual analog scale |
| WASO | wake after sleep onset |
| WBC | white blood cells |
| WPAI-GH | Work Productivity and Activity Impairment Questionnaire – General Health |
| TLG | tables, listings, and graphs |

# 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-G000-303.

This document is prepared on the basis of the final study protocol version 6.0 (dated 13 Aug 2018). The reader is referred to the study protocol, the electronic case report form (eCRF), general eCRF completion guidelines, and various data collection instruments employed in the study for details of study design, conduct and data collection.

This SAP is to be reviewed and approved prior to study database lock. If any updates are made upon blinded review of study data, or for any other reasons in the course of the study, such modifications and rationale are likewise to be documented and approved prior to unblinding of the study database.

## 1.1 Study Objectives

### 1.1.1 Primary Objective

The primary objective of the study is to:

Determine the efficacy of lemborexant 5 mg (LEM5) and 10 mg (LEM10) compared to placebo (PBO) on subjective sleep onset latency (sSOL) after 6 months of treatment in subjects with insomnia disorder

### 1.1.2 Secondary Objectives

**Key Secondary Objectives**

- Determine the efficacy of LEM5 and LEM10 compared to PBO on subjective sleep efficiency (sSE) after 6 months of treatment in subjects with insomnia disorder
- Determine the efficacy of LEM5 and LEM10 compared to PBO on subjective wake after sleep onset (sWASO) after 6 months of treatment in subjects with insomnia disorder

**Additional Secondary Objectives**

- Determine the efficacy of LEM5 and LEM10 compared to PBO on sSOL, sSE, sWASO, and subjective total sleep time (sTST):
  - over the 1st 7 nights of treatment
  - after 1 month of treatment
  - after 3 months of treatment
- Determine the efficacy of LEM5 and LEM10 compared to PBO on sTST at 6 months

- Evaluate the proportions of sleep onset and sleep maintenance responders to LEM5 and LEM10 compared to PBO as defined by response on sSOL or sWASO at 6 months and 12 months
- Evaluate the safety and tolerability of LEM5 and LEM10
- Evaluate the efficacy of LEM5 and LEM10 compared to PBO as measured by responses on the Insomnia Severity Index (ISI) and the Fatigue Severity Scale (FSS) after 6 months
- Evaluate rebound insomnia following discontinuation of treatment
- Evaluate morning sleepiness during and following completion of treatment
- Evaluate persistence of efficacy of LEM5 and LEM10 over 12 months

### 1.1.3 Exploratory Objectives

The following will be explored for both LEM5 and LEM10 compared to PBO over Period 1 and over Period 2 with analyses dependent on whether subjects received active treatment or PBO during Period 1:

- Efficacy on quality of sleep
- Health outcomes on the EuroQOL version 5D-3L (EQ-5D-3L), Work Productivity and Activity Impairment Questionnaire – General Health (WPAI-GH), and Patient Global Impression – Insomnia (PGI-Insomnia)
- Efficacy on sSOL, sSE, sWASO, sTST, ISI, and FSS
- Withdrawal symptoms after completion of treatment (Period 2 only)

The following will be explored for LEM5 and LEM10:

- Plasma concentrations of lemborexant and its metabolites M4, M9, and M10
- Population pharmacokinetic (PK) modeling for lemborexant
- PK/pharmacodynamic (PD) relationships between lemborexant concentrations and efficacy and safety variables

## 1.2 Overall Study Design and Plan

This study was a 12-month, multicenter, randomized, controlled, double-blind, parallel-group study of 2 doses of lemborexant in approximately 900 male and female subjects with insomnia disorder (subjects who complain of difficulties with sleep onset and/or sleep maintenance). Approximately 40% of the population was planned to be 65 years of age or older.

The study had 2 phases, the Prerandomization Phase and the Randomization Phase. The Prerandomization Phase comprised 3 periods and lasted up to 35 days: a Screening Period, a Run-In Period, and a Baseline Period. The Randomization Phase comprised a 6-month, placebo-controlled treatment period (Period 1) followed by a further 6 months in which

subjects received only active treatment (Period 2). Subjects were informed that they were to receive PBO at some point during the study and that all would receive active drug for at least 6 months. They were not informed of either the timing of these periods or the timing of the 2nd randomization. A 2-week Follow-Up Period was then to take place, followed by an End of Study Visit. The study design is shown in Figure 1.

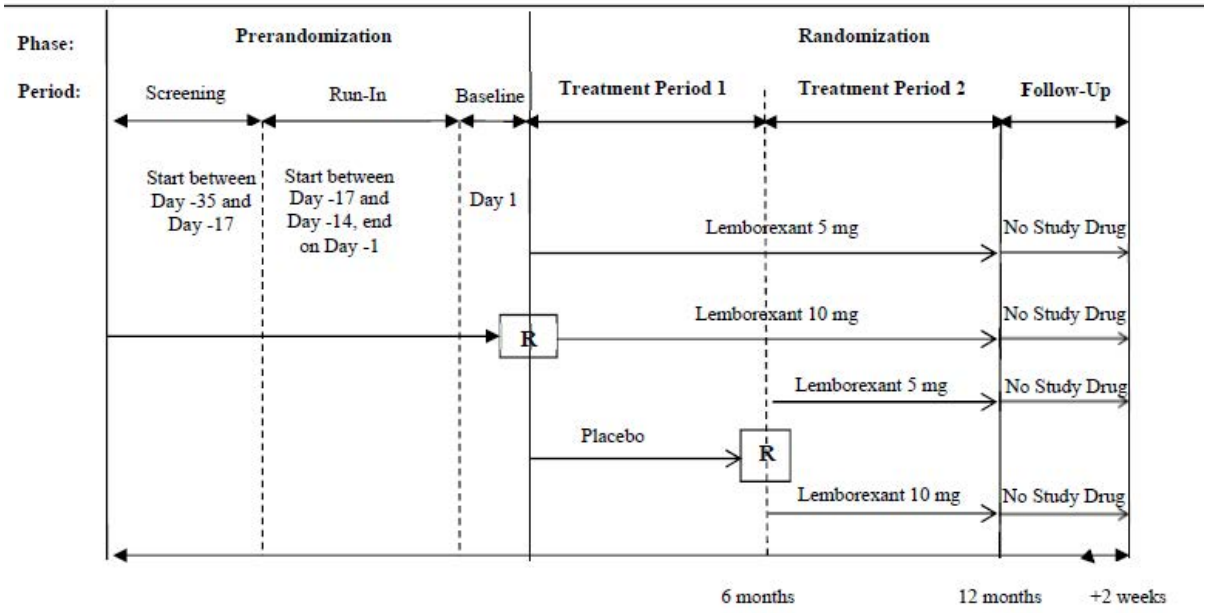



**Figure 1 Schematic Diagram of Study Design**

R = randomization

# 2 DETERMINATION OF SAMPLE SIZE

The sample size was estimated for the comparison of LEM10 and LEM5 with PBO, with respect to the mean change from Study Baseline at the end of Month 6 of the mean sSOL, the mean sSE and mean sWASO. This estimate was based on a sequential gate-keeping procedure at the 0.05 α-level in Section 3.3.3. There was sufficient power for both the primary endpoint (sSOL) and key secondary endpoints (sSE and sWASO).

On the basis of the dose-finding study E2006-G000-201 (Study 201), across various lemborexant doses (1 to 25 mg) at Days 8 to 15, the standard deviation of change from Study Baseline for sSOL was assumed to be 33 minutes; for sSE was assumed to be 12%; and for sWASO was assumed to be 43 minutes. The LS mean treatment differences at Days 8 to 15 from Study 201 were as follows: for sSE 9.5% for LEM10 and 5.5%, for LEM5; for sWASO -26.6 minutes for LEM10, and -11.3 minutes for LEM5. As a result of the non-normal distribution of sSOL, the LS mean treatment difference is not available (geometric mean ratios were used). Therefore, estimating treatment difference using medians at Days 8 to 15 from Study 201, leads to a median treatment difference of approximately -6.8 minutes for LEM10. For LEM5, a median treatment difference was approximately -13.2 minutes.

To detect a treatment difference in sSOL of at least -8.7 minutes, a sample size of 300 per treatment group at 5% (2-sided) level of significance had >90% power for comparing a dose of lemborexant with PBO.

To detect a treatment difference in sSE of at least 5.5%, a sample size of 300 per treatment group at 5% (2-sided) level of significance had >99% power for comparing a dose of lemborexant with PBO.

For sWASO, a total of 900 subjects (300 per treatment group) gave 90% power to detect a difference of -11.4 minutes for LEM5 and LEM10 compared to PBO (Table 2). The study was adequately powered to show the statistically significant difference from PBO for LEM10 as well as for LEM5.

**Table 2 Sample Size Calculations for Study E2006-G000-303**

| Alpha=0.05 | Difference between Active and Placebo CfB (min) | Power = 80% | | Power = 90% | |
|---|---|---|---|---|---|
| | | Sample size per group | Overall 3 group sample size | Sample size per group | Overall 3 group sample size |
| **sSOL (SD=33)** | -6.8 | 371 | 1113 | 496 | 1488 |
| | -7.6 | 300 | 900 | 398 | 1194 |
| | -8 | 269 | 807 | 359 | 1077 |
| | **-8.7** | 229 | 687 | **300** | **900** |
| | -9 | 213 | 639 | 284 | 852 |
| | -10 | 172 | 516 | 230 | 690 |
| | -11 | 143 | 429 | 191 | 573 |
| **sWASO (SD=43)** | -9 | 360 | 1080 | 481 | 1443 |
| | -9.9 | 300 | 900 | 398 | 1194 |
| | -10 | 292 | 876 | 390 | 1170 |
| | -11 | 241 | 723 | 323 | 969 |
| | **-11.4** | 225 | 675 | **300** | **900** |
| | -12 | 203 | 609 | 271 | 813 |

CfB = change from baseline, sSOL = subjective sleep onset latency, sWASO = subjective wake after sleep onset.

The study also had adequate power for the secondary analysis of responders. A sleep onset responder was defined as follows: sSOL at Study Baseline was ≥30 minutes and mean sSOL at 6 months was ≤20 minutes, and a sleep maintenance responder was defined as follows: sWASO at Study Baseline was ≥60 minutes and mean sWASO at 6 months was ≤60 minutes and showed a reduction of >10 minutes compared to Study Baseline. A total of 900 subjects gave >99% power to detect a treatment difference in sleep onset responder rates of 16% and sleep maintenance responder rates of 24.4%, both compared with placebo.

The above sample size also met regulatory safety requirements. Even with early discontinuation rates as high as 50% at Month 6 (ie, 150 subjects remaining) and 60% by Month 12 (ie, 120 subjects remaining in each treatment group), the requirement for 100 subjects in each of the elderly and non-elderly age-groups to complete 12-months of treatment on 5 or 10 mg lemborexant would be met. In addition, a total of 420 subjects would complete 6 months of treatment with 5 or 10 mg lemborexant.

# 3 STATISTICAL METHODS

All statistical analyses will be performed by the sponsor or designee. When all subjects have completed Treatment Period 1 (Period 1), all data will be unblinded to the sponsor. There will be 3 database locks. The first will be of Period 1 data only, where data up to and including the visit at Month 6 will be analyzed. The second will include all data (Period 1 and ongoing Period 2 data) after all continuing subjects have completed the Month 9 (Period 2) visit, all available data will be analyzed to support regulatory submissions. The third will include all data where the study has continued to completion.

For the Period 1 analysis, sponsor staff and certain designees (detailed in the Study Integrity Charter) will be unblinded to the treatment allocation; however, subjects, investigators, and site personnel will remain blinded. These are not considered formal interim analyses since the primary and key secondary efficacy endpoints are all at Month 6 (Period 1), and there is no placebo control group in Period 2. Therefore, no Type I error adjustment is needed for efficacy analysis.

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 group. sSOL will be log transformed in analyses unless otherwise stated.

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

## 3.1 Study Endpoints

Definitions of baseline, primary, secondary and exploratory endpoints are defined in this section. Where Sleep Diary endpoints are described, the time points refer to the mean of the final 7 nights before the visit unless otherwise stated.

### 3.1.1 Primary Endpoint

The primary endpoint is the mean change from Study Baseline in sSOL at Month 6.

### 3.1.2 Secondary Endpoints

The key secondary endpoints are:

- Mean change from Study Baseline in sSE at Month 6

- Mean change from Study Baseline of sWASO at Month 6

Additional secondary endpoints are:

- Mean change from Study Baseline of sSOL, of sSE, of sWASO and of sTST, at the beginning of treatment (mean of the 7 nights after the 1st dose in Period 1), at Month 1 and at Month 3
- Mean change from Study Baseline of sTST at Month 6
- Proportion of responders at Month 6 and Month 12, where sleep onset responder is defined as follows: sSOL at Study Baseline is ≥30 minutes and mean sSOL at 6 months is ≤20 minutes, and sleep maintenance responder is defined as follows: sWASO at Study Baseline is ≥60 minutes and mean sWASO at 6 months is ≤60 minutes and shows a reduction of >10 minutes compared to Study Baseline.
- Change from Study Baseline in daytime functioning, assessed as the total score from the 4 items on daytime functioning, on the ISI, at Months 1, 3, and 6
- Change from Study Baseline on the FSS at Months 1, 3, and 6
- Ratings on the morning sleepiness item of the Sleep Diary, for:
  - The mean change from Study Baseline of the 1st 7 mornings after the 1st dose in Period 1 and Period 2
  - The mean change from Study Baseline at: Month 1, Month 3, and Month 6
  - The mean change from Study Baseline and from Treatment Period 2 Baseline (as appropriate): for subjects with 1, 3, 6, 9 and 12 months exposure
  - The mean change from Screening for the 1st 7 mornings and 2nd 7 mornings of the Follow-up Period
- Rebound insomnia endpoints as assessed from the Sleep Diary during the Follow-up Period
  - Change from Screening of sSOL on each of the 1st 3 nights, mean sSOL of the 1st 7 nights, and mean sSOL of the 2nd 7 nights of the Follow-up Period
  - Change from Screening of sWASO on each of the 1st 3 nights, mean sWASO of the 1st 7 nights and mean sWASO of the 2nd 7 nights of the Follow-up Period
  - Proportion of subjects whose sSOL is longer than at Screening for each of the 1st 3 nights, or whose mean sSOL is longer than at Screening for 1st 7 nights or 2nd 7 nights of the Follow-up Period
  - Proportion of subjects whose sWASO is higher than at Screening for each of the 1st 3 nights, or whose mean sWASO is higher than at Study Baseline for the 1st 7 nights or 2nd 7 nights of the Follow-up Period
- Persistence of Effect
  - Mean change from Study Baseline of sSOL, of sSE, of sWASO and of sTST at Months 3, 6, 9, and 12 compared to Month 1
  - Mean change from Treatment Period 2 Baseline (Month 6) of sSOL, of sSE, of sWASO and of sTST at Months 9 and 12 compared to Month 7 (the 1st month of treatment in Period 2)

- Mean change from Study Baseline and Treatment Period 2 Baseline (as appropriate) of sSOL, of sSE, of sWASO and of sTST at 3 and 6 months exposure compared to 1 month of exposure

Safety and Tolerability of Lemborexant

- During Period 1, compared to PBO
- For subjects exposed to lemborexant for 3, 6, 9, and 12

### 3.1.3 Exploratory Endpoints

The following endpoints will be explored for LEM5 and LEM10. Except for PK endpoints, comparisons to PBO will be made.

- Change from Study Baseline in the mean value of the item on quality of sleep from the Sleep Diary for:
  - The 1st 7 mornings after the 1st dose in Period 1
  - Months 1, 3, and 6
- Change from Study Baseline and Treatment Period 2 Baseline (as appropriate) in the mean value of the item on quality of sleep from the Sleep Diary for:
  - subjects with 1, 3, 6, 9 and 12 months exposure
- Change from Study Baseline in:
  - EQ-5D-3L at Months 1, 3, and 6
  - WPAI-GH at Months 3 and 6
- Change from Study Baseline and Treatment Period 2 Baseline (as appropriate) in:
  - EQ-5D-3L for subjects with 3, 6, 9 and 12 months exposure to lemborexant
  - WPAI-GH for subjects with 3, 6, 9 and 12 months exposure to lemborexant
- Number and percentage of subjects with a rating of a positive medication effect on each PGI-Insomnia item (1) at Months 1, 3, and 6 (placebo-controlled Treatment Period 1) and (2) with 3, 6, 9 and 12 months exposure
- Change from Study Baseline and Treatment Period 2 Baseline (as appropriate) of sSOL, sSE, sWASO and sTST with 1, 3, 6, 9 and 12 months exposure, ISI, and FSS with 3, 6, 9 and 12 months exposure
- Mean score on the T-BWSQ of LEM5, and LEM10 compared to PBO at End of Study
- Plasma concentrations of lemborexant and its metabolites M4, M9, and M10
- PK of lemborexant using population modeling
- Relationships between lemborexant PK, efficacy, and/or safety variables using PK/PD modeling

### 3.1.4 Other Endpoints

Additional endpoints may be further explored on an exploratory or post-hoc basis to facilitate interpretation of study results.

## 3.2 Study Subjects

### 3.2.1 Definitions of Analysis Sets

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

On-Treatment Safety Analysis Set: On-Treatment Safety Analysis Set is the group of subjects who received at least 1 dose of lemborexant and had at least 1 post dose 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.

On-Treatment Full Analysis Set (FAS): On-Treatment FAS is the group of subjects who received at least 1 dose of lemborexant and had at least 1 post dose primary efficacy measurement.

6-Month Completer Analysis Set: The 6-Month Completer Analysis Set is the group of subjects in the FAS who had all efficacy assessments up to and including Month 6 (ie, Week 1 and Months 1 to 6 visits) without missing primary or key secondary efficacy assessments at any of these visits.

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

PK/PD Analysis Set: The PK/PD Analysis Set is the group of subjects receiving either lemborexant or PBO who have efficacy or safety data with documented dosing history. In addition, subjects receiving lemborexant should have at least 1 quantifiable lemborexant concentration data point as per the PK Analysis Set.

Per Protocol Analysis Set (PP): The PP is the group of subjects who sufficiently complied with the protocol. This analysis set will include all subjects in the Full Analysis Set who sufficiently complied with the protocol during Period 1. The criteria for exclusion in this analysis set are listed below:

- Violated inclusion/exclusion criteria
- Duplicate randomization
- Missing primary efficacy assessment
- Prohibited concomitant medication

- Study drug not administered
- Incorrect study drug kit dispensed
- Non compliance of study medication

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

The number of subjects screened, the number and percentage of screen failures, and their primary reason for screen failure will be summarized.

Subject disposition will be summarized by treatment groups for all randomized subjects. The number and percentage of subjects who completed or prematurely terminated from the treatment phase and from the study will be summarized separately by their primary reason for termination for each treatment group.

In addition, the number and percentage of randomized subjects who continued into Period 1 and Period 2 will be summarized by treatment group. The number and percentage of randomized subjects will be summarized by region, country, and sites for each treatment group. The number and percentage of randomized subjects in each of the study populations will also be presented.

### 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 randomized study drug and dosage
- 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, Safety Analysis Set and On-Treatment 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 (< 65, 65 to < 75, ≥75), BMI group (<18.5, 18.5 to <25, 25 to 30, >30), region (North America, Europe and New Zealand, Asia), race, and ethnicity.

The selected baseline assessments of Sleep Diary variables including sSOL, sWASO, sSE, sTST, sTIB, number of awakenings, quality of sleep, morning sleepiness; ISI total score and its individual question score; FSS will be summarized by treatment group. The BDI-II and BAI scores measured at screening will also be summarized.

#### 3.2.4.1 Medical History

All medical histories as documented by the Medical History and Current Medical Conditions eCRF 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. A subject data listing of medical history and current medical conditions will be provided.

### 3.2.5 Prior and Concomitant Therapy

All investigator terms for medications recorded in the eCRF will be coded to an 11-digit code using the World Health Organization Drug Dictionary (WHO DD; Mar 2017 or latest version). If the Safety Analysis Set and FAS differ, then the prior and concomitant medication summaries will be repeated on the FAS.

Prior medications are defined as medications that stopped before the first dose of study drug, where study drug includes placebo during the Run-In Period.

Concomitant medications are defined as medications that (1) started before the first dose of study drug (including placebo 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 placebo Run-In Period) to the last dose day plus 14 days.

For the Run-In Period, a concomitant medication is defined as a medication that started on or before the 1st dose of study medication of the Run-In Period and continued during the Run-In Period, or started during the Run-In Period. For Period 1, a concomitant medication is defined as a medication that started on or before the 1st dose of study medication of Period 1 and continued during Period 1, or started during Period 1. For Period 2, a concomitant medication is defined as a medication that started on or before the 1st dose of study medication in Period 2 and continued during Period 2, or started during Period 2.

The number (percentage) of subjects who take prior and concomitant medications will be summarized on the Safety Analysis Set by treatment group, Anatomical Therapeutic Chemical class (ATC) and World Health Organization Drug Dictionary-preferred term (PT). If the Safety Analysis Set and FAS differ, then the prior and concomitant medication summaries will be repeated on the FAS.

If the subject has taken the same concomitant medication (as coded to preferred WHO-drug term) more than once, the subject will be counted only once in the tabulation.

Separate summaries of concomitant medications will be provided for Run-in Period (using Safety Analysis Set), Period 1 (using Safety Analysis Set), Period 2 (using On-Treatment Safety Analysis Set) and Period 1 plus Period 2 (using On-Treatment Safety Analysis Set).

### 3.2.6 Treatment Compliance

Treatment compliance (in %) for each study drug will be calculated on the basis of number of tablets dispensed, lost and returned, separately for each type of tablet, for all randomized subjects as follows:

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

Treatment compliance will be summarized using descriptive summary statistics and number (percentage) of subjects using the categories <80%, ≥ 80% to ≤100%, >100% to ≤ 120%, and > 120% for each treatment group. Separate summaries will be provided for Run-in Period (using Safety Analysis Set), Treatment Period 1 (using Safety Analysis Set), and Treatment Period 2 (using On-Treatment Safety Analysis Set).

## 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 120 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, Europe and New Zealand, Asia) will be examined as specified in the respective sections in this document.

### 3.3.2 Adjustments for Covariates

Baseline assessment, region, and age groups 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 primary endpoint and secondary endpoint comparisons to control for the overall type 1 error at the 0.05 significance level. 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 Study Baseline at Month 6 in sSOL, LEM10 compared to PBO
- Change from Study Baseline at Month 6 in sSOL, 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:

- Change from Study Baseline at Month 6 in sSE, LEM10 compared to PBO
- Change from Study Baseline at Month 6 in sSE, LEM5 compared to PBO
- Change from Study Baseline at Month 6 in sWASO, LEM10 compared to PBO
- Change from Study Baseline at Month 6 in sWASO, LEM5 compared to PBO

No multiplicity testing will be performed for other secondary or exploratory endpoints.

### 3.3.4 Examination of Subgroups

Subgroup analysis of primary and key-secondary efficacy endpoints will be performed using age group (< 65, 65 to < 75, ≥75), sex (male and female), race (white, black, Asian, and other), region, and BMI group (< 18.5, ≥18.5 to <25, ≥25 to 30, and >30) as detailed in Section 3.4 and Section 3.6.

### 3.3.5 Handling of Missing Data, Dropouts, and Outliers

This study collected sleep diary data directly from subjects via a vendor-provided hand-held device; it is possible that some values entered by subjects may have resulted from input errors. For example, it is not possible for a total waketime to be longer than the time the subject spends in bed for the night. In such cases, erroneous data will be considered as missing values in data analyses. Efficacy data handling will be mentioned in the Section 6.1.

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:

## Table 3 Multiple Imputation Methods

<table>
<tr><th>MI Methods</th><th>Assumption</th><th>Analysis Type</th></tr>
<tr><td>Complete Case Missing Value (CCMV)</td><td>Subjects with missing data at any day have a similar distribution as the <u>completers within the respective treatment group</u> (Pattern 1, below), where completers are defined as having no missing assessments for any post-baseline visits</td><td>Primary</td></tr>
<tr><td>Complete Case-7<br>CCMV(k=7)</td><td>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.
<table>
<tr><th>Timepoints where results are available</th><th>Wk1</th><th>M1</th><th>M2</th><th>M3</th><th>M4</th><th>M5</th><th>M6</th></tr>
<tr><td>Pattern 1</td><td>x</td><td>x</td><td>x</td><td>x</td><td>x</td><td>x</td><td>x</td></tr>
<tr><td>Pattern 2</td><td>x</td><td>x</td><td>x</td><td>x</td><td>x</td><td>x</td><td>.</td></tr>
<tr><td>Pattern 3</td><td>x</td><td>x</td><td>x</td><td>x</td><td>x</td><td>.</td><td>.</td></tr>
<tr><td>Pattern 4</td><td>x</td><td>x</td><td>x</td><td>x</td><td>.</td><td>.</td><td>.</td></tr>
<tr><td>Pattern 5</td><td>x</td><td>x</td><td>x</td><td>.</td><td>.</td><td>.</td><td>.</td></tr>
<tr><td>Pattern 6</td><td>x</td><td>x</td><td>.</td><td>.</td><td>.</td><td>.</td><td>.</td></tr>
<tr><td>Pattern 7</td><td>x</td><td>.</td><td>.</td><td>.</td><td>.</td><td>.</td><td>.</td></tr>
</table>
Wk1=week1; M1=month1; M2=month2; M3=month3; M4=month4; M5=month5; M6=month6; x=observed data; .=missing data</td><td>Sensitivity</td></tr>
<tr><td>Tipping Point</td><td>Imputation towards the null hypotheses: A range of shifts starting from the primary analysis (CCMV) 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.</td><td>Sensitivity</td></tr>
<tr><td>Tipping Point</td><td>Imputation towards the null hypotheses: A range of shifts will be used in the multiple imputation of missing data assuming MAR to identify the specific shift and treatment effect that will tip the results from statistically significant to non-significant.</td><td>Sensitivity</td></tr>
</table>

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 efficacy endpoint (change from baseline in mean sSOL on Month 6) 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.

**Table 4 Estimands**

| Estimand | Description | Population | Intervention Effect of Interest | Analysis Type |
|---|---|---|---|---|
| Difference in outcome improvement for all randomized subjects | - all randomized subjects regardless of the treatment subjects actually received<br>- 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 |
| Difference in outcome improvement for all randomized subjects | - all randomized subjects regardless of what treatment subjects actually received<br>- include data after dropout | FAS | missing values imputed using MI assuming MNAR utilizing CCMV-7 missing value pattern (all available up to 7 monotone missing patterns)<br><br>(Assumes the probability of missing observations for any subject depends on the unobserved events. For the missing pattern, complete cases up to 7 monotone missing patterns will be used in the imputation – see Section 3.4.1.3for details. ***Thus this method relaxes the assumption of the primary analysis of using only completers to impute the missing data.*** | Sensitivity (CCMV-7) |
| Difference in outcome improvement for all randomized subjects | - all randomized subjects regardless of what treatment subjects actually received<br>- include data after dropout | FAS | a range of shifts starting from the primary analysis (CCMV) 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 the treatment subjects actually received<br>- include data after dropout | FAS | missing values imputed using MMRM model assuming MAR<br>(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.))*** | Supplementary (MMRM analysis assuming MAR) |
| Difference in outcome improvement for all randomized subjects | - all randomized subjects regardless of the treatment subjects actually received | 6-Months Completer Analysis Set | subjects who completed all primary and secondary efficacy assessments without missing visits | Supplementary (completer |

| Estimand | Description | Population | Intervention Effect of Interest | Analysis Type |
|---|---|---|---|---|
| | - subjects who complete the study without missing efficacy assessments | | | analysis) |
| Difference in outcome improvement for those who adhere to treatment | - subjects without major protocol violations that would impact efficacy assessments<br>- 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.)*** | Supplement ary (PP analysis) |

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;

## 3.4 Efficacy Analyses

Efficacy analyses for Treatment Period 1 will be performed on the FAS unless otherwise specified. Different baseline values will be used for endpoints defined during Treatment Period 1 (ie, Day 1 to Month 6) and during Treatment Period 2 (ie Month 7 to Month 12). Baseline values for each efficacy parameter are specified in Section 6.1.1. The On-Treatment FAS will be used for the on treatment summaries produced by time on treatment.

Unless specified otherwise, all efficacy endpoints will be derived by calculating the average of weekly (7 days) diary parameter values (Section 6.1).

12-month LEM exposure data will be summarized by treatment (LEM5, LEM10) and duration of exposure (1 month [30 days], 3 months [90 days], 6 months [180 days], 9 months [270 days] and 12 months [365 days]). Treatment groups are LEM5 and LEM10, and include both (a) LEM Period 1 subjects on the On-Treatment Full Analysis Set using the change from Study Baseline and (b) LEM Period 2 subjects previously receiving PBO on the On-Treatment Full Analysis Set using the change from Period 2 Baseline. No hypothesis test will be performed for efficacy endpoints beyond the Month 6 visit.

The primary and key secondary endpoints comparisons are tested following the gate-keeping testing procedure described in Section 3.3.3 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 testings will only proceed if the previous test is statistically significant at the 0.05 level.

### 3.4.1 Primary Efficacy Analyses

The primary efficacy endpoint is the change from Study Baseline of sSOL at Month 6 of LEM10 and LEM5 compared to PBO.

Null Hypothesis: No difference exists in the mean change from Study Baseline to Month 6 of treatment with LEM10 (or LEM5) as compared with PBO.

Alternative Hypothesis: For sSOL, a difference exists in the mean change from Study Baseline to Month 6 for LEM10 (or LEM5) as compared with PBO.

The sSOL change from Study Baseline to Month 6 will be analyzed using the mixed effect model repeated measurement (MMRM) analysis on the FAS. The model will be adjusted for the corresponding Study Baseline value, region (North America, Europe and New Zealand, Asia), age group (<65 years old, ≥65 years old), treatment, time (1st 7 nights, Month 1, Month 2, Month 3, Month 4, Month 5 and Month 6) and the interaction of treatment by time. Since the sSOL 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 toeplitz covariance matrix (TOEP) will be used. In the case of non-convergence with TOEP, the autoregressive covariance matrix [AR (1)] will be used in the model.

Before the implementation of the MMRM model, the missing values will be imputed using a pattern mixture model utilizing multiple imputation (MI) assuming the missing values are missing not at random (MNAR) utilizing the complete case missing value pattern (CCMV - subjects who completed all 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.

#### 3.4.1.1 Multiple Imputation

Step 1 (imputing missing data): Thirty multiple imputed complete datasets were to be constructed using imputation regression model of baseline sSOL, baseline ISI, age, sex, race (white, black, and other), and region (North America, Europe and New Zealand, Asia), baseline BMI. baseline log(sSOL), baseline ISI with a predefined arbitrary seed number (seed=2006). 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 (<65 years old, ≥65 years old), region (North America, Europe and New Zealand, Asia), treatment, visit (1st 7 nights, Month 1, Month 2, Month 3, Month 4, month 5 and Month 6), and treatment-by-visit interaction as fixed effect, and the baseline log(sSOL) 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 endpoints described in Section 3.4.1 will be summarized using descriptive statistics by each subgroup listed below. No hypothesis testing will be performed in the subgroup analyses.

- Age group 1 (< 65, 65 to < 75, ≥75 years old)
- Age group 2 (<65, ≥65 years old)
- Sex (male and female)
- Race (white, black, Asian and other)
- Region (North America, Europe and New Zealand, Asia)
- BMI group (kg/m2) (<18.5, 18.5 to <25, ≥25 to 30, >30 kg/m2)

#### 3.4.1.3 Sensitivity Analyses

The following analyses will be considered as sensitivity analyses:

- MI Imputation assuming MNAR utilizing CCMV-7: The same MMRM method used in the primary analysis will be applied utilizing CCMV-7 (ie, up to 7 monotone missing patterns will be used for missing value imputation as follows):

**Table 5 Monotone Missing Patterns for Primary Efficacy Analysis**

| Timepoints where results are available | Wk1 | M1 | M2 | M3 | M4 | M5 | M6 |
|---|---|---|---|---|---|---|---|
| Pattern 1 | x | x | x | x | x | x | x |
| Pattern 2 | x | x | x | x | x | x | . |
| Pattern 3 | x | x | x | x | x | . | . |
| Pattern 4 | x | x | x | x | . | . | . |
| Pattern 5 | x | x | x | . | . | . | . |
| Pattern 6 | x | x | . | . | . | . | . |
| Pattern 7 | x | . | . | . | . | . | . |

Wk1=week1; M1=month1; M2=month2; M3=month3; M4=month4; M5=month5; M6=month6; x=observed data; .=missing data

- Tipping point analysis: A range of shifts starting from the primary analysis (CCMV) 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.
- The same primary efficacy analyses described in Section 3.4.1 (MMRM analysis with MI for missing value imputation) will be repeated based on non-Log transformed data.
- A MI Imputation assuming MNAR utilizing CCMV-7: The same MMRM method used in the primary analysis will be applied utilizing CCMV-7 based on non-Log transformed data.
- Tipping point analysis based on non-Log transformed data: A range of shifts starting from the primary analysis (CCMV) 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.

#### 3.4.1.4 Supplementary Analyses

The following analyses will be considered as supplementary 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 6-Months Completer Analysis Set.

- 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).
- 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) based on non-Log transformed data.

### 3.4.2 Secondary Efficacy Analyses

#### 3.4.2.1 Key Secondary Efficacy Analyses

The change from Study Baseline of key secondary endpoints sSE and sWASO at Month 6 comparing LEM10/LEM5 to PBO will be analyzed as follows.

**CHANGE FROM STUDY BASELINE OF SSE AT MONTH 6**

The change from Study Baseline of sSE at Month 6 will be analyzed using the mixed effect model repeated measurement (MMRM) analysis on the FAS. The model will be adjusted for the corresponding Study Baseline value, region (North America, Europe and New Zealand, Asia), age group (<65 years old, ≥65 years old), treatment, time (1st 7 nights, Month 1, Month 2, Month 3, Month 4, Month 5 and Month 6) and the interaction of treatment by time. The unstructured covariance matrix (UN) will be used in the analysis. In the case of non-convergence of UN, the toeplitz covariance matrix (TOEP) will be used. In the case of non-convergence with TOEP, the autoregressive covariance matrix [AR(1)] will be used in the model.

Before the implementation of the MMRM model, the missing values will be imputed using a pattern mixture model utilizing multiple imputation (MI) assuming the missing values are missing not at random (MNAR). 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**

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 (North America, Europe and New Zealand, Asia), baseline BMI, baseline mean sSE, baseline ISI, and mean sSE at first 7 nights, mean sSE at Month 1, mean sSE at Month 2, mean sSE at Month 3, mean sSE at Month 4, mean sSE at Month 5 and mean sSE at Month 6. 50 multiple imputed datasets will be used for sSE, where data are truncated to 0 (where imputations <0) and 100 (where imputations >100).

#### CHANGE FROM STUDY BASELINE OF sWASO ON MONTH 6

The change from Study Baseline of sWASO on Month 6 will be analyzed using the mixed effect model repeated measurement (MMRM) analysis on the FAS. The model will be adjusted for the corresponding Study Baseline value, region (North America, Europe and New Zealand, Asia), age group (<65 years old, ≥65 years old), treatment, time (1st 7 nights, Month 1, Month 2, Month 3, Month 4, Month 5 and Month 6) and the interaction of treatment by time. The unstructured covariance matrix (UN) will be used in the analysis. In the case of non-convergence of UN, the toeplitz covariance matrix (TOEP) will be used. In the case of non-convergence with TOEP, the autoregressive covariance matrix [AR(1)] will be used in the model.

Before the implementation of the MMRM model, the missing values will be imputed using a pattern mixture model utilizing multiple imputation (MI) assuming the missing values are missing not at random (MNAR). 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

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 (North America, Europe and New Zealand, Asia), baseline BMI, baseline mean sWASO, baseline ISI, and mean sWASO at first 7 nights, mean sWASO at Month 1, mean sWASO at Month 2, mean sWASO at Month 3, mean sWASO at Month 4, mean sWASO at Month 5 and mean sWASO at Month 6.

#### 3.4.2.2 Sensitivity Analyses for Key Secondary Efficacy Analyses

The following analyses will be considered as sensitivity analyses:

- MI Imputation utilizing CCMV-7: The same MMRM method used in the primary analysis will be applied utilizing CCMV-7 (ie, up to 7 monotone missing patterns will be used for missing value imputation as follows):

**Table 6 Monotone Missing Patterns for Secondary Efficacy Analysis**

| Timepoints where results are available | Wk1 | M1 | M2 | M3 | M4 | M5 | M6 |
|---|---|---|---|---|---|---|---|
| Pattern 1 | x | x | x | x | x | x | x |
| Pattern 2 | x | x | x | x | x | x | . |
| Pattern 3 | x | x | x | x | x | . | . |
| Pattern 4 | x | x | x | x | . | . | . |
| Pattern 5 | x | x | x | . | . | . | . |
| Pattern 6 | x | x | . | . | . | . | . |
| Pattern 7 | x | . | . | . | . | . | . |

Wk1=week1; M1=month1; M2=month2; M3=month3; M4=month4; M5=month5; M6=month6; x=observed data; .=missing data

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

#### 3.4.2.3 Supplementary Analyses for Key Secondary Efficacy Analyses

The following analyses will be considered as supplementary 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 6-Months Completer Analysis Set.
- 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).

#### 3.4.2.4 Other Secondary Efficacy Analyses

For all other secondary endpoints described below, the continuous variable will be summarized by using descriptive statistics by time point and the categorical variable will be summarized as the number and percentage of subjects by time point.

Unless is covered from the same model from the primary and secondary efficacy endpoints, or specified otherwise, 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 region and age group. Missing values will be considered as non-responders in all responder analyses. No multiplicity adjustment will be made for all analyses.

The following endpoints will be analyzed.

**EFFICACY MEASURES DERIVED FROM THE SLEEP DIARY**

- Change from Study Baseline of mean sSOL, mean sSE, mean sWASO and mean sTST for the 1st 7 nights, Months 1, 2, 3, 4 and 5; and mean sTST at Month 6; for LEM5 and LEM10 compared to PBO
- Change from Study Baseline/ Treatment Period 2 Baseline of mean sSOL, mean sSE, mean sWASO and mean sTST for 1, 3, 6, 9 and 12 months exposure to LEM5 and LEM10. . Data from subjects randomized to LEM5 or LEM10 at the start of the study will use Study Baseline, data from subjects who were re-randomized to LEM5 or LEM10 in Period 2 will use Treatment Period 2 Baseline.
- The proportion of responders, separately for sSOL and sWASO, will be analyzed, for the first 7 nights, Months 1, 2, 3, 4, 5 and 6 for LEM5 and LEM10 compared to PBO:
  - The sleep onset responder is defined as: sSOL at Study Baseline is >30 minutes and mean sSOL at each month (first 7 nights, Months 1, 2, 3, 4, 5, 6) is ≤20 minutes,
  - The sleep maintenance responder is defined as: sWASO at Study Baseline is>60 minutes and mean sWASO at each month (first 7 nights, Months 1, 2, 3, 4, 5, 6) is ≤60 minutes and shows a reduction of >10 minutes compared to Study Baseline.

  The proportion of responders will be analyzed based on the subjects with sSOL > 30 minutes at the baseline for sSOL and separately the subjects with sWASO > 60 minutes at the baseline for sWASO.
- The proportion of responders, separately for sSOL and sWASO, at Months 7, 8, 9, 10, 11 and 12 will also be summarized by using two types of baseline (ie., Study Baseline, Period 2 Baseline). The definition of the responder at Month 7, 8, 9, 10, 11 and 12 will be as follows:
  - The sleep onset responder is defined as: sSOL at Study Baseline (or Period 2 Baseline) is >30 minutes and mean sSOL at each month (Months 7, 8, 9, 10, 11 and 12) is ≤20 minutes,
  - The sleep maintenance responder is defined as: sWASO at Study Baseline (or Period 2 Baseline) is >60 minutes and mean sWASO at each month (Months 7, 8, 9, 10, 11

and 12) is ≤60 minutes and shows a reduction of >10 minutes compared to Study Baseline (or Period 2 Baseline).

The proportion of responders will be analyzed based on the subjects with sSOL > 30 minutes at the baseline for sSOL and separately the subjects with sWASO > 60 minutes at the baseline for sWASO.

- Morning sleepiness item of the Sleep Diary:
  - Change from Study Baseline of mean rating on morning sleepiness at the 1st 7 mornings after the 1st dose in Treatment Period 1, Month 1, Month 2, Month 3, Month 4, Month 5 and Month 6.
  - Change of mean rating on morning sleepiness from Study Baseline / Treatment Period 2 Baseline for 1st 7 mornings after the 1st LEM dose, 1, 3, 6, 9 and 12 months of exposure to LEM5 and LEM10. Data from subjects randomized to LEM5 or LEM10 at the start of the study will use Study Baseline, data from subjects who were re-randomized to LEM5 or LEM10 in Period 2 will use Treatment Period 2 Baseline. Change from Screening of mean rating on morning sleepiness of 1st 7 mornings and 2nd 7 mornings in Follow-Up Period.
- Rebound insomnia endpoints during the Follow-up Period as follows:
  - Change from Screening of sSOL on each of the 1st 3 nights, mean sSOL of the 1st 3 nights, mean sSOL of the 1st 7 nights, and mean sSOL of the 2nd 7 nights of the Follow-up Period
  - Change from Screening of sWASO on each of the 1st 3 nights, mean sWASO of the 1st 3 nights, mean sWASO of the 1st 7 nights and mean sWASO of the 2nd 7 nights of the Follow-up Period
  - Proportion of subjects whose sSOL is longer than at Screening by at least 5 minutes for each of the 1st 3 nights, mean sSOL of the 1st 3 nights, mean of the 1st 7 nights, and mean of the 2nd 7 nights of the Follow-up Period
  - Proportion of subjects whose sWASO is higher than at Screening by at least 5 minutes for each of the 1st 3 nights, mean sWASO of the 1st 3 nights, mean of the 1st 7 nights, and mean of the 2nd 7 nights of the Follow-up Period.

The actual value of sSOL and sWASO will be analyzed separately using analysis of covariance model (ANCOVA) with factors of age group (<65, and ≥65 years older), region (North America, Europe and New Zealand, Asia), and treatment for each time point (baseline, each of the first 3 nights, mean of the first 3 nights, mean of the first 7 nights, and mean of the last 7 nights of the Follow-up Period). The 95% CI of the treatment difference will be constructed for each time point. It will be considered as strong evidence of rebound insomnia if the lower bound of the 95% CI of sSOL or sWASO for each of the 3 nights, the mean of the first 3 nights, mean of the first 7 nights, 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. In this analysis, the treatment groups will be defined as the last randomized study treatment just before the Follow-up period.

- Analyses for persistence versus loss of effect will be conducted for sSOL, sSE, sWASO and sTST at Months 2 through Month 12 compared to Month 1.

  The 95% CI for the change from Study Baseline of sSOL, sSE, sWASO and sTST at each visit will be calculated using the MMRM model assuming MAR (no missing value imputation) with factors of age group (<65, and ≥65 years older), region (North America, Europe and New Zealand, Asia), treatment, visit and treatment-by-visit interaction as fixed effect, and baseline value as covariate based on FAS. At each month beyond Month 1, the change from Study Baseline will be compared to either the lower bound of the 95% CI (for sSE and sTST) or the upper bound of the 95% CI (for sSOL and sWASO) at Month 1. Persistence of effect will be defined as present if the mean change from Study Baseline at Month 6 is above the lower bound of the 95% CI at Month 1 for sSE or sTST and below the upper bound of the 95% CI at Month 1 for sSOL and sWASO.

  The same analyses for persistence versus loss of effect over Period 2 will compare sSOL, sSE, sWASO, and sTST at Months 7 through 12 to these measures at Month 7 for only the subjects who were randomized to PBO in Period 1 and who continued into Period 2. At each month beyond Month 7, the change from Treatment Period 2 Baseline will be compared to either the lower bound of the 95% CI (for sSE and sTST) or the upper bound of the 95% CI (for sSOL and sWASO) at Month 7. Persistence of effect will be defined as present if the mean change from Treatment Period 2 Baseline at Month 12 is above the lower bound of the 95% CI at Month 7 for sSE or sTST and below the upper bound of the 95% CI at Month 7 for sSOL and sWASO.

  Analyses for persistence versus loss of effect over duration of exposure will be conducted for On-Treatment Full Analysis Set subjects. These analyses will compare 1 month of duration of exposure of sSOL, sSE, sWASO and sTST at 3 and 6 months duration of exposure, for a) LEM Period 1 subjects using the change from Study Baseline and (b) LEM Period 2 subjects previously receiving PBO using the change from Period 2 Baseline. At 3 and 6 months duration of exposure, the change from Study Baseline/Treatment Period 2 Baseline will be compared to either the lower bound of the 95% CI (for sSE and sTST) or the upper bound of the 95% CI (for sSOL and sWASO) at 1 month duration of exposure. Persistence of effect will be defined as present if the mean change from Study Baseline/Treatment Period 2 Baseline at 6 months of exposure is above the lower bound of the 95% CI at 1 month of exposure for sSE or sTST and below the upper bound of the 95% CI at 1 month of exposure for sSOL and sWASO.

#### INSOMNIA SEVERITY INDEX AND FATIGUE SEVERITY SCALE

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

- Change from Study Baseline of the total score from items 1-7, and items 4-7, separately, on the ISI at Months 1, 3, 6 of LEM5 and LEM10 compared to PBO.

- Change from Study Baseline and Treatment Period 2 Baseline (as appropriate) of the total score from items 1-7, and items 4-7, separately, on the ISI for subjects with 3, 6, 9 and 12 months of exposure.
- Change from Study Baseline on the FSS score at Months 1, 3, 6 of LEM5 and LEM10 compared to PBO
- Change from Study Baseline and Treatment Period 2 Baseline (as appropriate) on the FSS score for subjects with 3, 6, 9 and 12 months of exposure.

The FSS will be further analyzed to evaluate whether lemborexant treatment affected fatigue severity in those who entered the study with clinically significant levels of fatigue. These subgroup analyses will include only those whose total FSS score was ≥18 at Study Baseline, which can be considered the lower threshold on the FSS for the presence of clinically significant fatigue. Analyses using the FSS analysis model will then be applied. In addition, S-plot figures will show the proportion of subjects with varying rates of response on the FSS as follows:

- Cumulative proportion of subjects with varying absolute decreases from Study Baseline in total FSS score: the x-axis will show decrease from baseline in 5-unit increments up to a ≥40-unit decrease from baseline at Months 1, 3 and 6. Responder rates of -10 and -30 units will be marked on the S-Plot.

- Cumulative proportion of subjects with varying relative decreases from Study Baseline in total FSS score: the x-axis will show decrease from baseline in 10% increments up to a 100% decrease from baseline at Months 1, 3 and 6. Responder rates of -10% and -30% will be marked on the S-Plot.

## 3.5 Pharmacokinetic, Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

### 3.5.1 Pharmacokinetic Analyses

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

A population PK approach will be used to characterize the PK of lemborexant. For this approach, PK analysis data from this study will be pooled with relevant data from Phase 1 and 2 studies, and other Phase 3 studies if available. The effect of covariates (eg, demographics, concomitant medications) on the PK of lemborexant will be evaluated. The PK model will be parameterized for oral clearance (CL/F) and volumes of distribution. Derived exposure parameters such as AUC and $C_{max}$ of lemborexant and any other relevant parameters will be calculated from the model using the individual estimates parameterized for oral clearance and dosing history. A separate analysis plan for the population PK analyses will be developed and finalized before the database lock.

### 3.5.2 Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

The PK/PD relationship between exposure to lemborexant and the efficacy variables including but not limited to sSOL, sSE, and sWASO, and the safety variables including but not limited to morning sleepiness and frequently occurring TEAEs, will be explored graphically. Any emergent PK/PD relationships will be evaluated by population PK/PD modeling. The population PK/PD analysis plan will be described and results will be reported in a separate document.

Population PK and PK/PD analyses will be performed using NONMEM Version 7.2 or later.

## 3.6 Safety Analyses

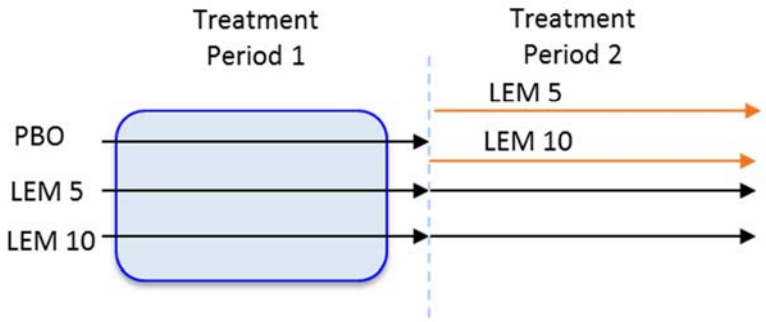
Evaluations of safety will be performed on the relevant Safety Analysis Set and On-Treatment Safety Analysis Set, as appropriate. All safety analyses will be performed based on observed data using tabulation or descriptive statistics only.

No hypothesis testing will be performed for safety analyses

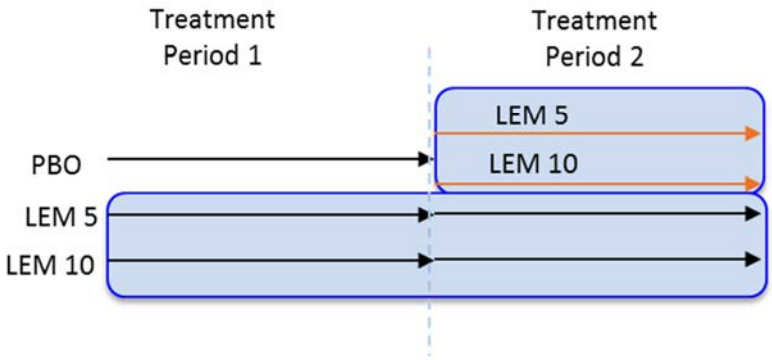
### 3.6.1 Treatment Definitions

For visit-based Safety data (ie, laboratory values, vital signs, ECGs, etc) data will be summarized by treatment (LEM5, LEM10) and duration of exposure to treatment (3, 6, 9 and 12 months). Treatment groups are LEM5 and LEM10, and include both (a) LEM Period 1 subjects on the On-Treatment Safety Analysis Set using the change from Study Baseline and (b) LEM Period 2 subjects previously receiving PBO on the On-Treatment Safety Analysis Set using the change from Period 2 Baseline. Summaries of exposure, AEs, TEMAV, clinically abnormal vital sign and ECG values will be summarized for:

- Period 1 data, where treatment groups are PBO, LEM5 and LEM10 in the Safety Analysis Set (ie, subjects who received at least 1 dose of study drug in Treatment Period 1)




- 12-month LEM exposure, where treatment groups are LEM5 and LEM10 in the On-Treatment Safety Analysis Set (ie., subjects who received at least 1 dose of lemborexant), where data will be summarized by exposure to treatment (3, 6, 9 and 12 months)




### 3.6.2 Extent of Exposure

The extent of exposure (mean daily dose, cumulative dose, duration of exposure) to study drug will be summarized descriptively. Period 1 data with LEM5 or LEM10 will be summarized for: the number (percent) of subjects who completed 1 month (30 or more days), 3 months (90 or more days), and 6 months (180 or more days). 12-month exposure with LEM5 or LEM10 will be summarized for: the number (percent) of subjects who completed 3 months (90 or more days), 6 months (180 or more days), 9 months (270 or more days) and 12 months (365 or more days) of dosing.

The duration of exposure to study drug is defined as the date of last dose – the date of first dose date +1.

The exposure to Run-In Period placebo will be presented in the subject data listings only.

### 3.6.3 Adverse Events

The AE verbatim descriptions (investigator terms from the eCRF) will be classified into standardized medical terminology using the MedDRA. Adverse events will be coded to the MedDRA (Version 17.0 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.

Only those AEs that are treatment-emergent will be included in summary tables. All AEs, treatment-emergent or otherwise, will be presented in subject data listings. AEs will be classified as TEAEs up to 14 days after the last study treatment (see below).

A 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 first dose of study drug), having been absent before the randomized study medication is started or

- Reemerges during active treatment, having been present at pretreatment (before the randomized study medication is started) but stopped before treatment, or
- Worsens in severity during active treatment (after randomized study medication starts) relative to the pretreatment state, when the AE is continuous.

Separately, a TEAE during each Period (ie, Run-In, Treatment Period 1 or Treatment Period 2) is defined as:

1. TEAE during the Run-In: AE that emerges prior to the date of first dose in Treatment Period 1 or the date of last dose in Run-In + 14 days, whichever comes earlier.
2. TEAE during Period 1: AE that emerges after the date of first dose in Treatment Period 1 but prior to (a) the date of first dose in Period 2 or (b) the date of last dose in Treatment Period 1 + 14 days, whichever comes earlier.
3. TEAE during Period 2: AE that emerges after the date of first dose in Treatment Period 2 but prior to the date of last dose in Treatment Period 2 + 14 days, whichever comes earlier.

Adverse events will be summarized using the Safety Analysis Set and On-Treatment Safety Analysis Set by treatment groups.

The number (percentage) of subjects with TEAEs during the Run-In Period will be summarized by SOC and PT.

The following summary tables will be produced for the Period 1 (Safety Analysis Set)and On-Treatment (On-Treatment Safety Analysis Set) separately.

An overview table of TEAEs, 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 will be provided for each treatment group.

In addition, the following summaries will be produced for the TEAEs by treatment group:

- 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 maximum severity
- Incidence of treatment-related TEAEs by SOC, PT, and maximum severity
- Incidence of TEAEs by SOC, PT, and relationship to treatment

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:

- 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 will also be summarized by MedDRA SOC, PT for each treatment group.

#### 3.6.3.1 Selected Adverse Events

The following summary tables will be produced for the Period 1 (Safety Analysis Set)and 12-month LEM exposure (On-Treatment Safety Analysis Set) separately.

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

- Cataplexy
- Potential cataplexy
- Falls
- Seizures
- Abuse liability events

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

Cataplexy and potential cataplexy includes the TEAEs with MedDRA PT listed in Appendix 2 plus all events that have been flagged by investigator as potential event of cataplexy in the eCRF.

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

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

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

#### 3.6.3.2 Subgroup Analysis of TEAE

Subgroup summaries will be presented for the following summary tables. These summaries will be repeated for the Period 1 (Safety Analysis Set)and 12-month LEM exposure (On-Treatment Safety Analysis Set) separately.

- Overview table of TEAEs, including number of subjects with TEAEs, treatment-related TEAEs, severe TEAEs, treatment-emergent serious AEs (TESAEs), deaths, TEAEs leading to study drug withdrawal
- Incidence of TEAEs by SOC and PT
- Incidence of TESAEs by SOC and PT

The following demographic and baseline information will be used in the subgroup summary:

- Age group 1 (<65, 65 to <75, and ≥75 years old)
- Age group 2 (<65 and ≥65 years old)
- Sex (male and female)
- Race (white, black, Asian, and other)
- Region (North America, Europe and New Zealand, Asia)
- BMI (<25, 25 to 30, >30 kg/m2)

Percentages will be based on the total number of subjects in the relevant treatment group. A subject with 2 or more TEAEs with the same PT (or SOC) will be counted only once for that PT (or SOC).

### 3.6.4 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 9.5.1.5.3 Laboratory Measurements, the actual value and the change from Study Baseline will be summarized at each scheduled visit using descriptive statistics by treatment group. The change from Study Baseline/Treatment Period 2 Baseline (as appropriate) will also be repeated for the 12-month LEM exposure (On-Treatment Safety Analysis Set), summarized by treatment group and duration of exposure.

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 Study Baseline (LNH) to each scheduled visit will be provided by treatment group for each laboratory parameter based on Safety Analysis Set. The shift tables will also be repeated for the 12-month LEM exposure (On-Treatment Safety Analysis Set), summarized by treatment group and duration of exposure.

The Sponsor's Grading for Laboratory Values (Appendix 1) presents the Eisai criteria that will be used to identify subjects with treatment-emergent markedly abnormal laboratory values (TEMAV). Except for phosphate, a TEMAV is determined if a postbaseline grade increases from baseline and the postbaseline grade is greater than or equal to a grade of 2. For phosphate, a TEMAV was defined as a postbaseline value with an increase from baseline to a grade of 3 or higher.

The number (percentage) of subjects with TEMAVs will be summarized by treatment group using the Safety Analysis Set. The number (percentage) of subjects with TEMAVs will be produced separately for the Period 1, up to 12-months treatment, up to 6-months treatment and On-Treatment. 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.

#### 3.6.4.1 Subgroup Analysis of Laboratory Values

Subgroup summaries will be presented for the following summary tables. These summaries will be repeated for the Period 1 (Safety Analysis Set) and 12-month LEM exposure (On-Treatment Safety Analysis Set) separately.

- Mean and mean change from baseline by visit for selected parameters in Table 7
- TEMAVs

**Table 7 Selected Laboratory Parameters for Subgroup Analysis**

| Category | Parameters |
|---|---|
| Hematology | hematocrit, hemoglobin, platelets, red blood cell count, and white blood cell count |
| Chemistry – Liver Function Tests | alanine aminotransferase, alkaline phosphatase, aspartate aminotransferase, gamma glutamyl transpeptidase, direct bilirubin, total bilirubin |
| Chemistry – Renal Function Tests | blood urea/blood urea nitrogen, creatinine |

The following demographic and baseline information will be used in the subgroup summary:

- Age group 1 (<65, 65 to <75, and ≥75 years old)
- Age group 2 (<65 and ≥65 years old)
- Sex (male and female)
- Race (white, black, Asian, and other)
- Region (North America, Europe and New Zealand, Asia)
- BMI (<25, 25 to 30, >30 kg/$m^2$)

### 3.6.5 Vital Signs

Descriptive statistics for vital signs parameters (ie, diastolic and systolic BP, pulse, respiration rate, temperature) and weight, and changes from Study Baseline and Period 2 Baseline will be presented by visit and treatment group. Analysis of changes from baseline will be based on the number of subjects with both nonmissing baseline and relevant postbaseline results.

Clinically notable vital sign values will be identified on the listings as those above (H) or below (L) a clinically notable range, as per Table 8. The clinically notable vital sign values will be presented for change from Study Baseline (Safety Analysis Set) by treatment group and by time point. This will also be repeated for the 12-month LEM exposure (On-Treatment Safety Analysis Set), summarized by treatment group and duration of exposure.

**Table 8 Vital Sign Criteria**

| Variable | Criterion Value[a] | Change Relative to Study Baseline/Period 2 Baseline[a] | Clinically Notable Range |
|---|---|---|---|
| Pulse 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 | >30 bpm | Increase of >=10 bpm | H |
| | <8 bpm | Decrease of >=4 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 Study Baseline or Period 2 Baseline, as appropriate.

### 3.6.6 Electrocardiograms

Descriptive statistics for ECG parameters and changes from Study Baseline (Safety Analysis Set) will be presented by treatment group. This will also be repeated for the 12-month LEM exposure (On-Treatment Safety Analysis Set), summarized by treatment group and duration of exposure. Shift tables will present changes from Study Baseline in ECG interpretation (categorized as normal; abnormal, not clinically significant; and abnormal, clinically significant) by time point.

In addition, maximum postbaseline measurement will also be tabulated separately for the Period 1 and 12-month LEM exposure 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

### 3.6.7 Other Safety Analyses

#### 3.6.7.1 Columbia-Suicide Severity Rating Scale (C-SSRS)

Suicidality is 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 occurrence of suicidal ideation or suicidal behavior at each visit will be summarized by treatment group using frequency count. A subject will be counted once for each of these categories if at least one question is answered positive belonging to the category.

In addition, self-injurious non-suicidal behavior will also be summarized separately for occurrence in a similar fashion to suicidal ideation and behavior.

#### 3.6.7.2 Tyrer Benzodiazepine Withdrawal Symptom Questionnaire (T-BWSQ)

Withdrawal symptoms are assessed using the T-BWSQ at the End of Study (EOS) visit. Subjects are 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. Since this questionnaire is completed at the EOS visit only, this analysis will be performed at the EOS visit based on Safety Analysis set. The same analysis will be repeated for the subgroup based on the duration of exposure.

## 3.7 Other Analyses

### 3.7.1 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 Study Baseline and each postbaseline timepoint (Month 1, Month 3 and Month 6) using frequency count on observed data only with no imputation. The change from Study Baseline of EQ VAS will be summarized by treatment group for Month 1, Month 3, and Month 6. The change from Study Baseline of EQ VAS will be analyzed using MMRM assuming MAR for Month 1, Month 3 and Month 6 on observed data only based on FAS with no imputation for missing values.

The each dimension and EQ-VAS change from Study Baseline/Treatment Period 2 Baseline (as appropriate) will be summarized for the 12-month LEM exposure (On-Treatment Safety Analysis Set), summarized by treatment group and duration of exposure .

The time point window presented in Section 6.2.3 will be used to determine the analysis visit.

### 3.7.2 Patient Global Impression (PGI) - Insomnia

The PGI-Insomnia questionnaire captures the global impression of the study medication's effect at each postbaseline time point. 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 summarized separately as number and percentage of subjects for each time point (Month 1, Month 3 and Month 6). Furthermore, each item will be analyzed using chi-square test for Month 1, Month 3 and Month 6 on observed data only based on FAS with no imputation for missing values, and repeated for age subgroups. For this analysis, each item will be categorized as follows: "positive medication effect" versus others for the first 3 item; "just right" versus others for the last item). This will also be repeated for the 12-month LEM exposure (On-Treatment Safety Analysis Set) will be summarized by treatment group and duration of exposure.

The time point window presented in 8.2.3 will be used to determine the analysis visit.

#### 3.7.2.1 WPAI-GH

The WPAI-GH consists of six questions: Q1 = currently employed; Q 2 = hours missed due to health problems; Q3 = hours missed other reasons; Q4 = hours actually worked; Q5 = degree health affected productivity while working; Q6 = degree health affected productivity in regular unpaid activities.

The four main outcomes derived by the following definition will be summarized by visit. The change from Study Baseline for the main outcomes will be summarized by visit. This will also be repeated for the 12-month LEM exposure (On-Treatment Safety Analysis Set), summarized by treatment group and duration of exposure.

The time point window presented in 8.2.4 will be used to determine the analysis visit.

1. Percent work time missed due to health: Q2 / (Q2+Q4) * 100
2. Percent impairment while working due to health: Q5 / 10 * 100
3. Percent overall work impairment due to health: [{Q2 / (Q2+Q4)} + {1- (Q2 / (Q2+Q4))} *(Q5/10)] *100
4. Percent activity impairment due to health: Q6 /10 *100

The above 1 to 3 will be obtained from the subjects who were employed (ie., Q1 ='Y').

### 3.8 Exploratory and Pharmacodynamic Analyses

The change from Study Baseline for the mean score of the quality of sleep item on the Sleep Diary will be analyzed using MMRM assuming MAR (no missing value imputation) for the mean of the 1st 7 days of Period 1 and at Months 1, through 6.

Furthermore, the mean score of the quality of sleep item change from Study Baseline/Treatment Period 2 Baseline (as appropriate) for the 12-month LEM exposure (On-Treatment Safety Analysis Set), summarized by treatment group and duration of exposure.

## 4 INTERIM ANALYSES

### 4.1 DSMB (Data Safety Monitoring Board)

The DSMB will serve as an independent safety monitoring committee and will perform the safety data reviews. The interim safety analyses will be conducted by an independent statistician, who is working on behalf of Eisai from a CRO that is independent of study conduct. To maintain the credibility and integrity of the trial, procedures will be implemented to ensure the DSMB and the independent statistician have sole access to the unblinded interim safety data. Full details of the DSMB operation procedures are documented in the DSMB charter. The sponsor will be kept blinded until the unblinding of Period 1. No decision to stop the study based on superior efficacy or futility of lemborexant to placebo is planned. Data will be cleaned as required prior to the scheduled DSMB data reviews.

### 4.2 Interim Analysis

No formal interim analysis is planned for this study.

## 5 CHANGES IN THE PLANNED ANALYSES

The following major changes to the statistical methods in this analysis plan were made after Amendment 03 of the Protocol was approved.

- Additional time points (ie., Months 2, 4, 5) have been included in MMRM as covariates for the primary and key secondary analysis.

- An additional safety analysis set (ie., On-Treatment Safety Analysis Set) has been defined and the safety analysis based on the On-Treatment Safety Analysis Set has been added.

The full changes to the statistical methods in this analysis plan were made in version 2.0 of the SAP from version 1.0 are detailed in Table 1 at the beginning of the SAP.

# 6 DEFINITIONS AND CONVENTIONS FOR DATA HANDLING

## 6.1 Efficacy Data Handling

### 6.1.1 Baseline Efficacy

For the analyses of PBO-controlled endpoints (ie, Day 1 to Month 6), "Study Baseline" is defined as the data captured during the Run-In Period (or during the Baseline Period). For the analyses of data that are not placebo-controlled (ie, Month 7 to Month 12), data from the Month 6 visit will be used as the "Period 2 Baseline" for some analyses.

Unless otherwise specified, the Study Baseline measurement is the last observed measurement among all available measurements, including from unscheduled assessments, prior to the first dose of randomized study medication for a given assessment. For the following endpoints, baseline measurement is defined as follows:

- Sleep diary parameters: the mean of diary data entered on the last 7 mornings before Visit 3a during the Run-In Period
- ISI: Last available ISI measurement on or before Visit 3a (excluding Visit 1)
- FSS: Last available FSS measurement on or before Visit 3a (excluding Visit 1)
- EQ-5D-3L: Last available EQ-5D-3L measurement on or before Visit 3a (excluding Visit 1)
- WPAI-GH: Last available WPAI-GH measurement on or before Visit 3a (excluding Visit 1)

For each Sleep Diary parameter at baseline, if no more than 2 of the final 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 the 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 fewer than 4 values are available, the parameter will be considered missing for the corresponding time point.

For the analysis using the MMRM assuming MAR (no missing value imputation), the above handling will be used to obtain the average of weekly (7 days) diary parameter values.

### 6.1.2 Derivation of Primary and Key Secondary Endpoints for Sleep Diary

Each Sleep Diary parameter will be derived by calculating the average of weekly (7 days) diary parameter values.

Baseline is the mean of diary data entered on the last 7 mornings during the Run-In Period before randomization; sleep diary parameter of the first 7 nights is the mean of diary data entered on the first 7 mornings following the start of treatment; sleep diary parameter for each visit is the mean of diary data entered on the last 7 mornings prior to that visit.

For the follow-up period, if the first 7 nights overlap with the last 7 nights (eg, the follow-up period is less than 14 days in total), the last non-overlap nights will be used in calculating the average value for the last 7 nights.

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)

### 6.1.3 Visit Windowing

For the Sleep Diary, the following visit windowing rule in Table 9 will be applied based on measurements from unscheduled visits and early termination visits to attain the measurement for analysis for the given scheduled visit.

**Table 9 Windowing Rules for Efficacy Endpoints from Sleep Diary**

| Visit | Target Visit Day (in study days) | Visit Window (in study days) |
|---|---|---|
| First 7 nights | 2 | 2 - 8 |
| Visit 4 (Month 1) | 30 | 24 - 30 |
| Visit 5 (Month 2) | 60 | 54 – 60 |
| Visit 6 (Month 3) | 90 | 84 - 90 |
| Visit 7 (Month 4) | 120 | 114 - 120 |
| Visit 8 (Month 5) | 150 | 144 - 150 |
| Visit 9 (Month 6) | 180 | 174 - 180 (or last 7 days prior to subject entering Period 2, whichever is earlier) * |

**Table 9 Windowing Rules for Efficacy Endpoints from Sleep Diary**

| Visit | Target Visit Day (in study days) | Visit Window (in study days) |
|---|---|---|
| Visit 10 (Month 7) | 210 | 204 - 210 |
| Visit 11 (Month 8) | 240 | 234 - 240 |
| Visit 12 (Month 9) | 270 | 264 - 270 |
| Visit 13 (Month 10) | 300 | 294 - 300 |
| Visit 14 (Month 11) | 330 | 324 - 330 |
| Visit 15 (Month 12) | 364 | Last 7 days from study days 358 - 371 (or last dose date + 1 for subjects who completed 12 months treatment, whichever is earlier) |

*: If subjects discontinued the study and did not enter Treatment Period 2, last 7 days from study days 168 -180 will be used.

## 6.2 SAFETY DATA HANDLING

### 6.2.1 Baseline Safety

The baseline value for all safety endpoints will be the last non-missing measurement occurring prior to the first dose of the study medication.

For visit-based safety summaries there may be two different baseline value as followed:

- Study Baseline: defined as the last non-missing measurement occurring prior to the first dose of the study medication in Period 1.
- Treatment Period 2 Baseline: defined as the last non-missing measurement occurring prior to the first dose of the study medication in Period 2.

### 6.2.2 Handling of Missing Safety data, Drop-outs, and Outliers

For the purpose of summarizing maximum severity, if the severity of an AE is missing for a subject, then, if this subject has another AE with the same preferred term that has “severe” severity, the maximum severity of the AE will be noted as “severe’; otherwise the maximum severity will be noted as missing. Similarly, for the purpose of summarizing TEAE treatment relationship, if the relationship of an AE to study drug is missing, the AE will be noted to be related if there is another related AE with the same preferred term, otherwise this relationship will be noted as missing.

For determining treatment-emergent markedly abnormal laboratory values, a missing baseline laboratory value will be assumed to be of grade 0.

No special handling of missing data is planned for the analysis of any of the other safety variables.

Data exceptions or outliers will be determined by inspection of the tables, listings, and graphs in consultation with the clinical study team, if applicable.

All the listings will display the original missing values.

### 6.2.3 Visit Windowing

For the laboratory values, ECG, vital signs, weight, eC-SSRS, T-BWSQ, PGI-Insomnia, WPAI-GH and EQ-5D-3L, the following visit windowing rule in Table 10 will be applied based on measurements from unscheduled visits and early termination visits to attain the measurement for analysis for the given scheduled visit.

**Table 10 Windowing Rules for Safety Endpoints**

| Visit | Target Visit Day (in study days) | Visit Window (in study days) |
|---|---|---|
| Visit 4 (Month 1) | 30 | 2 - 45 |
| Visit 5 (Month 2) | 60 | 46 - 75 |
| Visit 6 (Month 3) | 90 | 76 - 135 |
| Visit 9 (Month 6) | 180 | 136 – Prior to the date of first dose of Period 2<br><br>If subjects discontinue treatment, the visit window will be 136 – 187 |
| Visit 12 (Month 9) | 270 | The date of first dose in Period 2 - 315<br><br>If subjects discontinue treatment, the visit window will be 188 – 315 |
| Visit 15 (Month 12) | 364 | 316 - 371 |
| End of Study (EOS) | 378 | Nominal visit<br>(ie., the visit captured in eCRF will be used) |

In case of multiple assessments within the same visit window, the assessment closest to the target visit day will be used in data analyses; in the event of two assessments being equally close to the scheduled visit day, the latest assessment in time will be used.

# 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=2006 out=<dataset1>;
  VAR age BMI baseline… visit1-visit8;
  MCMC chain=multiple nbiter=500 niter=300 impute=monotone;
  BY treatment;
RUN;
```

**IMPUTE MISSING VALUES:**

```
PROC MI data=<dataset1> nimpute=1 seed=2006 out=<dataset2>;
  CLASS treatment sex race region;
  MONOTONE regression ( /details);
  MNAR model (visit2-visit8/ 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;
  LSMEANS visit*treatment/DIFF 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**

- sSOL: age, baseline BMI, baseline ISI, baseline log(mean sSOL), log(mean sSOL) at first 7 nights, log(mean sSOL) at Month 1, log(mean sSOL) at Month 2, log(mean sSOL) at Month 3, log(mean sSOL) at Month 4, log(mean sSOL) at Month 5, log(mean sSOL) at Month 6

- sSE: age, baseline BMI, baseline ISI, baseline mean sSE, mean sSE at first 7 nights, mean sSE at Month 1, mean sSE at Month 2, mean sSE at Month 3, mean sSE at Month 4, mean sSE at Month 5, mean sSE at Month 6
- sWASO: age, baseline BMI, baseline ISI, baseline mean sWASO, mean sWASO at first 7 nights, mean sWASO at Month 1, mean sWASO at Month 2, mean sWASO at Month 3, mean sWASO at Month 4, mean sWASO at Month 5, mean sWASO at Month 6
- sTST: age, baseline BMI, baseline ISI, baseline mean sTST, mean sTST at first 7 nights, mean sTST at Month 1, mean sTST at Month 2, mean sTST at Month 3, mean sTST at Month 4, mean sTST at Month 5, mean sTST at Month 6

**To Impute Missing Values**

- sSOL: treatment, sex, race, region + variables used to create monotone missing data pattern
- sSE: treatment, sex, race, region + variables used to create monotone missing data pattern
- sWASO: treatment, sex, race, region + variables used to create monotone missing data pattern
- sTST: treatment, sex, race, region + variables used to create monotone missing data pattern

**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 sSOL (Log sSOL), sSE and sWASO using the multiple imputation methodology as described in the section above but with the following modifications:

1. Convert the dataset into monotone missing data patterns (30 datasets), as above.
2. Generate the dataset so that there is a missing data indicator for data that has been imputed (ie, 1= observed, 0= missing).
3. Perform CCMV imputation. Ensure the missing data indicator is kept in this dataset.
4. Using the CCMV dataset, 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 MNAR/CCMV assumption when the shift is zero. These shifts are to be applied to Month 6 only.
5. For each shift apply the PROC MIXED model by _imputation_ and combine results using PROC MIANALYZE
6. Store results from PROC MIANALYZE for each shift
7. If the MNAR (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.

8. 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.
9. The values for <shift> in Step 1 will be applied uniformly (same shift) to both LEM5 and LEM10 at Month 6.
10. The values for <shift> will correspond to worsening values for the endpoint according to the following:

- Increasing positive shift values (sSOL, sWASO)
- Increasing negative shift values (sSE)

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

    LEM5 and LEM10 vs. Placebo (sSOL, sSE, sWASO)

For the MAR sensitivity analysis, the following procedure should be applied:

1. Convert the dataset into monotone missing data patterns (30 datasets), as above.
2. 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 Month 6 only.

```
PROC MI data=<dataset1> nimpute=1 seed=2006 out=<dataset2>;
    CLASS treatment sex race region;
    MONOTONE regression ( /details);
    MNAR adjust (visit8/ shift=<shift> adjustobs=(treatment=LEM5));
    MNAR adjust (visit8/ shift=<shift> adjustobs=(treatment= LEM10));
    VAR treatment age sex race region BMI baseline…visit8;
    BY _imputation_;
RUN;
```

3. 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.
4. 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.
5. The values for <shift> in Step 1 will be applied uniformly (same shift) to both LEM5 and LEM10 at Month 6.
6. The values for <shift> will correspond to worsening values for the endpoint according to the following:

- Increasing positive shift values (sSOL, sWASO)
- Increasing negative shift values (sSE)

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

- LEM5 and LEM10 vs. Placebo (sSOL, sSE, sWASO)

# 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.

# 9 MOCK TABLES, LISTINGS, AND GRAPHS

The study TLG 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 Imputations 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<br><LLN – 100 g/L<br><LLN – 6.2 mmol/L | <10.0 – 8.0 g/dL<br><100 – 80 g/L<br><6.2 – 4.9 mmol/L | <8.0 g/dL<br><80 g/L<br><4.9 mmol/L;<br>transfusion indicated | life-threatening consequences; urgent intervention indicated |
| Leukocytes (total WBC) | <LLN – $3.0\times10^9$/L<br><LLN – 3000/$mm^3$ | <$3.0 – 2.0\times10^9$/L<br><3000 – 2000/$mm^3$ | <$2.0 – 1.0\times10^9$/L<br><2000 – 1000/$mm^3$ | <$1.0\times10^9$/L<br><1000/$mm^3$ |
| Lymphocytes | <LLN – 800/$mm^3$<br><LLN – $0.8\times10^9$/L | <800 – 500/$mm^3$<br><$0.8 – 0.5\times10^9$/L | <500 – 200/$mm^3$<br><$0.5 – 0.2\times10^9$/L | <200/$mm^3$<br><$0.2\times10^9$/L |
| Neutrophils | <LLN – $1.5\times10^9$/L<br><LLN – 1500/$mm^3$ | <$1.5 – 1.0\times10^9$/L<br><1500 – 1000/$mm^3$ | <$1.0 – 0.5\times10^9$/L<br><1000 – 500/$mm^3$ | <$0.5\times10^9$/L<br><500/$mm^3$ |
| Platelets | <LLN – $75.0\times10^9$/L<br><LLN – 75,000/$mm^3$ | <$75.0 – 50.0\times10^9$/L<br><75,000 – 50,000/$mm^3$ | <$50.0 – 25.0\times10^9$/L<br><50,000 – 25,000/$mm^3$ | <$25.0\times10^9$/L<br><25,000/$mm^3$ |
| **METABOLIC/LABORATORY** | | | | |
| Albumin, serum- low (hypoalbuminemia) | <LLN – 3 g/dL<br><LLN – 30 g/L | <3 – 2 g/dL<br><30 – 20 g/L | <2 g/dL<br><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<br><LLN – 2.0 mmol/L | <8.0 – 7.0 mg/dL<br><2.0 – 1.75 mmol/L | <7.0 – 6.0 mg/dL<br><1.75 – 1.5 mmol/L | <6.0 mg/dL<br><1.5 mmol/L |
| Calcium, serum-high (hypercalcemia) | >ULN – 11.5 mg/dL<br>>ULN – 2.9 mmol/L | >11.5 – 12.5 mg/dL<br>>2.9 – 3.1 mmol/L | >12.5 – 13.5 mg/dL<br>>3.1 – 3.4 mmol/L | >13.5 mg/dL<br>>3.4 mmol/L |
| Cholesterol, serum-high (hypercholesterolemia) | >ULN – 300 mg/dL<br>>ULN – 7.75 mmol/L | >300 – 400 mg/dL<br>>7.75 – 10.34 mmol/L | >400 – 500 mg/dL<br>>10.34 – 12.92 mmol/L | >500 mg/dL<br>>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:<br>>ULN – 160 mg/dL<br>>ULN – 8.9 mmol/L | Fasting glucose value:<br>>160 – 250 mg/dL<br>>8.9 – 13.9 mmol/L | >250 – 500 mg/dL;<br>>13.9 – 27.8 mmol/L;<br>hospitalization indicated | >500 mg/dL;<br>>27.8 mmol/L;<br>life-threatening consequences |
| Glucose, serum-low (hypoglycemia) | <LLN – 55 mg/dL<br><LLN – 3.0 mmol/L | <55 – 40 mg/dL<br><3.0 – 2.2 mmol/L | <40 – 30 mg/dL<br><2.2 – 1.7 mmol/L | <30 mg/dL<br><1.7 mmol/L<br>life-threatening consequences; seizures |
| Phosphate, serum-low | <LLN – 2.5 mg/dL | <2.5 – 2.0 mg/dL | <2.0 – 1.0 mg/dL | <1.0 mg/dL |

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

| Code | Preferred Term |
|---|---|
| 10007737 | Cataplexy |
| 10009346 | Clonus |
| 10050093 | Consciousness fluctuating |
| 10012373 | Depressed level of consciousness |
| 10013036 | Diplopia |
| 10076060 | Eyelid myoclonus |
| 10078509 | Heteronymous diplopia |
| 10078508 | Homonymous diplopia |
| 10021118 | Hypotonia |
| 10024855 | Loss of consciousness |
| 10049565 | Muscle fatigue |
| 10072889 | Muscle tone disorder |
| 10028372 | Muscular weakness |
| 10028622 | Myoclonus |
| 10053854 | Opsoclonus myoclonus |
| 10036653 | Presyncope |
| 10050496 | Reversible ischaemic neurological deficit |
| 10071299 | Slow speech |
| 10042772 | Syncope |
| 10044380 | Transient global amnesia |
| 10044390 | Transient ischaemic attack |
| 10013887 | Dysarthria |
| 10033799 | Paralysis |
| 10009192 | Circulatory collapse |
| 10015995 | Eyelid ptosis |

# Appendix 3 List of Abuse Liability Events

| Code | Preferred Term |
|---|---|
| 10063746 | Accidental death |
| 10000381 | Accidental overdose |
| 10000383 | Accidental poisoning |
| 10054196 | Affect lability |
| 10074080 | Autoscopy |
| 10078602 | Behavioral addiction |
| 10067494 | Confusional arousal |
| 10012218 | Delirium |
| 10012335 | Dependence |
| 10077805 | Depersonalisation/derealisation disorder |
| 10013142 | Disinhibition |
| 10013457 | Dissociation |
| 10013462 | Dissociative disorder |
| 10013654 | Drug abuse |
| 10061111 | Drug abuser |
| 10080753 | Product administered at inappropriate site |
| 10013663 | Drug dependence |
| 10052237 | Drug detoxification |
| 10066053 | Drug diversion |
| 10061822 | Drug intolerance |
| 10052804 | Drug tolerance |
| 10052805 | Drug tolerance decreased |
| 10052806 | Drug tolerance increased |
| 10079381 | Drug use disorder |
| 10013752 | Drug withdrawal convulsions |
| 10013753 | Drug withdrawal headache |
| 10013754 | Drug withdrawal syndrome |
| 10014551 | Emotional disorder |
| 10049119 | Emotional distress |
| 10015535 | Euphoric mood |
| 10016322 | Feeling abnormal |
| 10016330 | Feeling drunk |
| 10017062 | Formication |
| 10019063 | Hallucination |
| 10019070 | Hallucination, auditory |
| 10019071 | Hallucination, gustatory |
| 10019072 | Hallucination, olfactory |
| 10062824 | Hallucination, synaesthetic |

| | |
|---|---|
| 10019074 | Hallucination, tactile |
| 10019075 | Hallucination, visual |
| 10019079 | Hallucinations, mixed |
| 10019133 | Hangover |
| 10020927 | Hypnagogic hallucination |
| 10020928 | Hypnopompic hallucination |
| 10049564 | Impaired driving ability |
| 10049976 | Impatience |
| 10021567 | Impulsive behaviour |
| 10021588 | Inappropriate affect |
| 10022523 | Intentional overdose |
| 10074903 | Intentional product misuse |
| 10076308 | Intentional product use issue |
| 10022998 | Irritability |
| 10023236 | Judgement impaired |
| 10026749 | Mania |
| 10048294 | Mental status changes |
| 10027940 | Mood altered |
| 10027951 | Mood swings |
| 10028747 | Nasal necrosis |
| 10028765 | Nasal septum perforation |
| 10028766 | Nasal septum ulceration |
| 10028896 | Needle track marks |
| 10061862 | Neonatal complications of substance abuse |
| 10033295 | Overdose |
| 10079763 | Performance enhancing product use |
| 10061355 | Poisoning |
| 10067669 | Prescription form tampering |
| 10069330 | Product tampering |
| 10070592 | Product used for unknown indication |
| 10037249 | Psychotic behaviour |
| 10061920 | Psychotic disorder |
| 10038001 | Rebound effect |
| 10040026 | Sensory disturbance |
| 10062684 | Somatic hallucination |
| 10066169 | Substance abuse |
| 10067688 | Substance abuser |
| 10076595 | Substance dependence |
| 10070964 | Substance use |
| 10079384 | Substance use disorder |
| 10072387 | Substance-induced mood disorder |

| | |
|---|---|
| 10072388 | Substance-induced psychotic disorder |
| 10043431 | Thinking abnormal |
| 10070863 | Toxicity to various agents |
| 10049414 | Treatment noncompliance |
| 10048010 | Withdrawal syndrome |
| MedDRA v21.0 | |

# SIGNATURE PAGE

Authors:

[electronic signature in eDMS]

PPD

PPD

Neuroscience Business Group, Eisai Ltd

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

Neurology Business Group, Eisai Ltd.

Date