STATISTICAL ANALYSIS PLAN

Study Protocol Number:	E2006-G000-202
Study Protocol Title:	A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study with Open-Label Extension Phase of the Efficacy and Safety of Lemborexantin Subjects with Irregular Sleep-Wake Rhythm Disorder and Mild to Moderate Alzheimer's Disease Dementia
Date:	20 JUN 2018
Version:	Final Version 1.0

TABLE OF CONTENTS

1 LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
ADAS-Cog	Alzheimer's Disease Assessment Scale – Cognitive Subscale
AE	adverse event
AMP	Amplitude of the rest-activity rhythm
ANCOVA	analysis of covariance
aSE	Actigraphy sleep efficiency
ATC	anatomical therapeutic class
aWE	Actigaphy wake efficiency
BMI	body mass index
CGIC- ISWRD	Clinician's Global Impression of Change - Irregular Sleep-Wake Rhythm Disorder version
CI	confidence interval
CRF	case report form
CSR	clinical study report
E2006	lemborexant
EQ-5D-5L	EuroQOL version 5D-5L
FAS	full analysis set
IS	Interdaily Stability
ISWRD	Irregular Sleep-Wake Rhythm Disorder
IV	Intradaily Variability
LEM	lemborexant
LEM2.5	lemborexant 2.5 mg
LEM5	lemborexant 5 mg
LEM10	lemborexant 10 mg
LEM15	lemborexant 15 mg
LS	least squares
MCP-MOD	Multiple Comparisons and Modelling
MedDRA	Medical Dictionary for Regulatory Activities
MMSE	Mini Mental State Examination
NPI-10	Neuropsychiatric Inventory

Abbreviation	Term
PBO	Placebo
PD	pharmacodynamic
PK	pharmacokinetic
PSG	Polysomnography
PSQI	Pittsburgh Sleep Quality Index
RA	Relative amplitude of the rest-activity rhythm
SAE	serious adverse event
SAP	statistical analysis plan
SD	Standard deviation
SDI	Sleep Disorders Inventory
SE	Sleep efficiency (PSG)
SFI	Sleep Fragmentation Index
SI	Système International
TEAE	treatment-emergent adverse event
TIB	Time in bed
TLG	tables, listings, and graphs
WFI	Wake Fragmentation Index
WHO	World Health Organization
ZBI	Zarit Burden Interview – short form

2 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-G000-202 (core study). The Open-Label Extension of this study will be covered by a separate SAP. Additional exploratory or post-hoc analyses not identified in this SAP may be performed to facilitate interpretation of study results and documented in the clinical study report.

This document is prepared on the basis of the final study protocol version 10.0 amendment 6 (dated 20 Jun 2018). The reader is referred to the study protocol, the case report form (CRF), general CRF completion guidelines, and various data collection instruments employed in the study for details of study design, conduct, and data collection.

This SAP was 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 study database.

2.1 STUDY OBJECTIVES

2.1.1 Sleep-Related Objectives (Revised per Amendments 05 and 06)

- To determine the dose response of lemborexant 2.5 mg (LEM2.5), 5 mg (LEM5), 10 mg (LEM10) and 15 mg (LEM15) compared to placebo (PBO) on the change from baseline in actigraphy-derived Sleep Efficiency (aSE) during the last week of treatment in subjects with Alzheimer's disease dementia (AD-D) who have ISWRD.
- To determine the efficacy of LEM2.5, LEM5, LEM10 and LEM15 compared to PBO on the change from baseline aSE during each week of treatment.
- To determine the efficacy of lemborexant LEM2.5, LEM5, LEM10 and LEM15 compared to PBO on the change from baseline on the Sleep Fragmentation Index (SFI) during each week of treatment.
- To determine the change from baseline of the mean duration of wake bouts (aMeanDurWB) over each week of treatment.

2.1.2 Wake-Related Objectives (revised per Amendments 05 and 06)

- To determine the dose response of lemborexant 2.5 mg (LEM2.5), 5 mg (LEM5), 10 mg (LEM10) and 15 mg (LEM15) compared to placebo (PBO) on the change from baseline in actigraphy-derived Wake Efficiency (aWE) during the last week of treatment in subjects with Alzheimer's disease dementia (AD-D) who have ISWRD.
- To determine the efficacy of lemborexant LEM2.5, (LEM5, LEM10, and LEM15 compared to PBO on the change from baseline of actigraphy-derived Wake Efficiency (aWE) during each week of treatment.

- To determine the efficacy of lemborexant LEM2.5, LEM5, LEM10, and LEM15 compared to PBO on the change from baseline of the Wake Fragmentation Index (WFI) during each week of treatment.
- To determine the change from baseline of the mean duration of sleep bouts (aMeanDurSB) over each week of treatment.

2.1.3 Circadian Rhythm-Related Objectives

- To evaluate onset and treatment course effect as measured by change from baseline of intradaily variability (IV), interdaily stability (IS), amplitude of the rest-activity rhythm (AMP), relative amplitude of the rest-activity rhythm (RA), and other actigraphy variables during each week of treatment.

2.1.4 Additional Objectives (revised per Amendment 05 and 06)

- To evaluate the safety and tolerability of lemborexant.
- To explore the effects of LEM2.5, LEM5, LEM10, LEM15, and PBO at the end of 4 weeks of treatment (unless otherwise specified)` on the following:
- Change from baseline of sum of activity counts and change from baseline of the number of bouts >10 minutes of sleep in the first 3 hours after morning waketime on each of the first 3 days and last 3 days of treatment as an indicator of next-morning residual effects.
- Potential rebound ISWRD in the 2 weeks following 4 weeks of treatment.
- Onset and course of treatment effect as measured by change from baseline of Clinician's Global Impression of Change-ISWRD (CGIC-ISWRD) Scale on symptoms of ISWRD total score and domains.
- Change from baseline of Alzheimer's Disease Assessment Scale – Cognitive Subscale (ADAS-cog).
- Change from baseline in Mini Mental State Examination (MMSE).
- Change from baseline in sleep quality in caregivers as measured by the Pittsburgh Sleep Quality Index (PSQI).
- Change from baseline of caregiver burden on the Zarit Burden Interview – short form (ZBI).
- Change from baseline of Health outcomes of the subject and/or caregiver on the EuroQOL version 5D-5L (EQ-5D-5L) (subject Self Version, caregiver Self Version, caregiver Proxy 1 Version).
- Change from baseline of mood and behavior on the Neuropsychiatric Inventory (NPI- 10; by caregiver as proxy for the subject).
- To characterize the pharmacokinetics (PK) of lemborexant using the population approach.
- To explore the PK/pharmacodynamic (PD) relationship between exposure to lemborexant and selected efficacy variables and most frequently occurring treatment- emergent adverse events (TEAEs).
- To assess the plasma concentrations of cognitive enhancers (cholinesterase inhibitors and/or memantine) and lemborexant in subjects taking such drugs.

- To evaluate the long-term safety and tolerability of flexible doses of LEM5, LEM10, and LEM15 per day over a period of 30 months in subjects with ISWRD who have completed the Core Study.

2.2 Overall Study Design and Plan

E2006-G000-202 is a multicenter, randomized, double-blind, PBO-controlled, parallel-group study of 4 doses of lemborexant or PBO taken daily for 4 weeks in approximately 60 male or female subjects, ages 60 to 90 years, with mild or moderate AD-D who complain of disrupted sleep or multiple awakenings at night along with frequent periods of falling asleep during the day that impacts the quality of life of the subject. For each subject, an individual who knows the subject well and will provide the information about themselves will also be enrolled in the study (see Caregivers and Informants, below). Additional informants may also be associated with the study but will not be required to complete a consent form.

The study will have 3 phases: the Prerandomization Phase, the Randomization Phase, and the Extension Phase (Figure 1). The Prerandomization Phase will comprise 2 periods that will last up to a maximum of 42 days: a Screening Period and a Baseline Period. The Randomization Phase will comprise a Treatment Period during which subjects will be treated for 4 weeks, and a minimum 14-day Follow-Up Period before an End of Study visit. The Extension Phase comprises a 30-month Maintenance Period and a 14-day Follow-Up Period. Subjects who complete the Core Study End of Study (EOS) Visit within 30 days prior to enrollment in the Extension Phase will be eligible for participation. For subjects continuing directly from the Core Study into the Extension Phase, the EOS Visit of the Core Study will be the start of the Extension Phase. Subjects who complete the Core Study, but who do not elect to immediately continue into the Extension Phase have up to 30 days after the EOS Visit to participate. These subjects will be required to return to the site within 30 days of completion of the EOS Visit to repeat selected assessments before being dispensed drug for the Extension Phase.

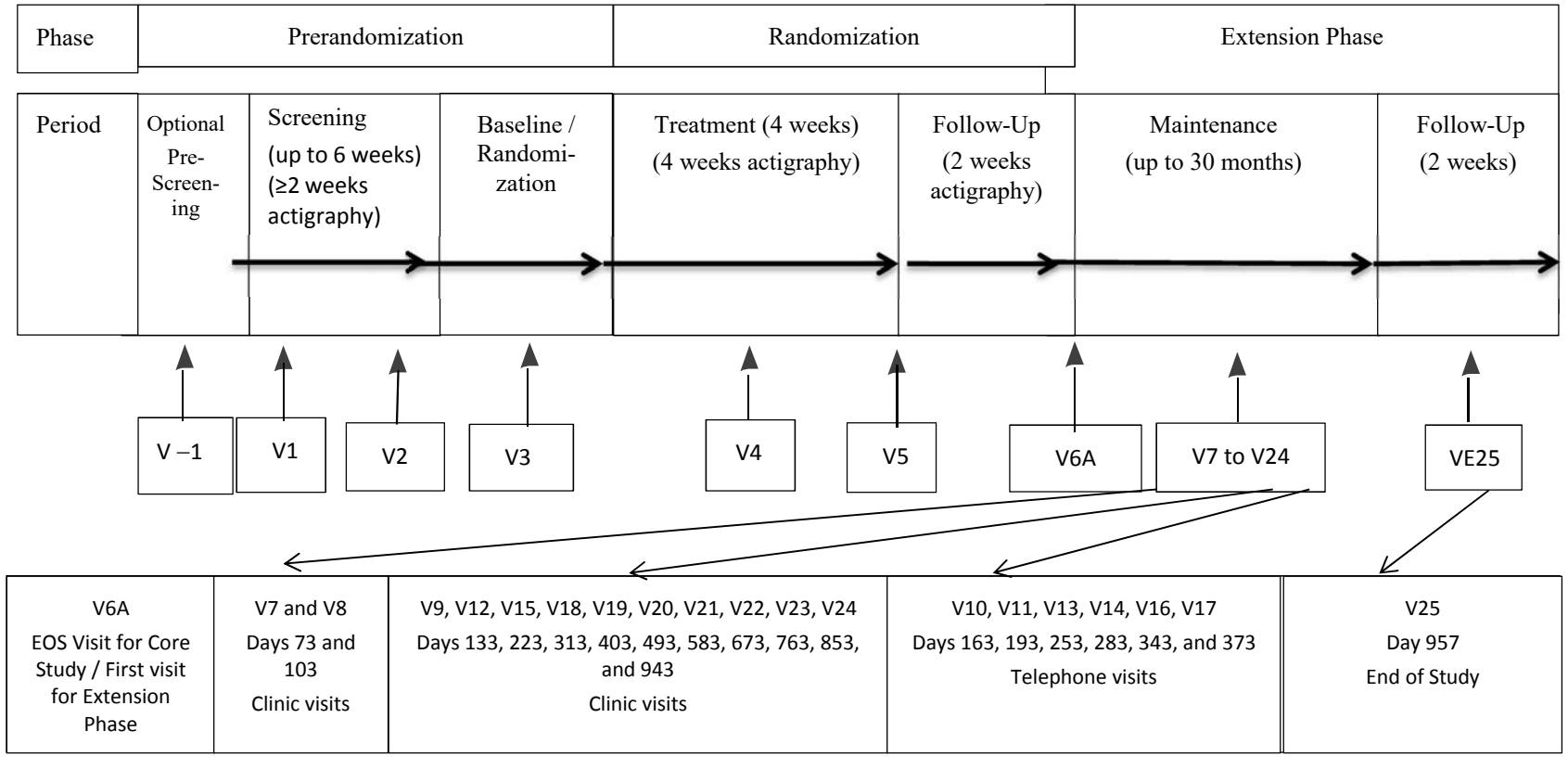

Figure 1 Study Design

Note: Visit –1 is an optional Prescreening Period for subjects and caregivers who are unsure whether the subject has Irregular Sleep-Wake Rhythm Disorder (ISWRD). Subjects will wear an activity tracker for approximately 4 days, then return to the site to have the data from the activity tracker downloaded and analyzed to determine whether they are candidates for the study. (revised per Amendment 05)
Note: Visit 2 is a caregiver visit for downloading actigraphy data to determine eligibility; Visit 3 is the baseline visit for both subject and caregiver; Visit 4 is a visit for both subject and caregiver to download actigraphy data and perform safety assessments; Visit 5 is the end-of-treatment assessments visit; and Visit 6A is for end-of-study assessments for the Core Study. For subjects continuing directly from the Core Study into the Extension Phase, Visit 6A will be the start of the Extension Phase. Subjects who complete the Core Study, but who do not elect to immediately continue into the Extension Phase have up to 30 days after Visit 6A to participate. These subjects will be required to repeat selected assessments (Visit 6B) before being dispensed drug for the Extension Phase. Some subjects will have fewer visits based on the availability of lemborexant commercially.
ISWRD = Irregular Sleep-Wake Rhythm Disorder, V = Visit.

3 DETERMINATION OF SAMPLE SIZE

Approximately 60 subjects will enroll in this proof-of- concept study. (revised per Amendment 06)

Some supporting exploratory power estimation is based on testing the dose response using the MCP-Mod package in R; several dose response curves were explored at one-sided 0.05 level of significance with different values of n (12, 13, and 14 subjects) per treatment group (Table 1, Table 2, and Table 3) and are presented as conservative scenarios based on a 10-percentage point difference between the active dose and placebo.

Table 1 Power Calculations for Actigraph Sleep Efficiency Under Different Dose Response Curves (n=12 and sigma=12)

Model (R variables)	Optimal Contrasts (PBO, 2.5, 5, 10, 15)	Power (%)
BetaMod 1 (0.339,0.05)	–0.82, –0.074, 0.116, 0.336, 0.442	0.6674254
BetaMod 2 (0.5,0.3)	–0.834, –0.073, 0.187, 0.414, 0.305	0.6588761
BetaMod 3 (0.19,0.15)	–0.887, 0.124, 0.236, 0.304, 0.223	0.6807728
E-max 1 (0.66)	–0.883, 0.095, 0.209, 0.277, 0.301	0.6953901
E-max 2 (1.1)	–0.869, 0.033, 0.195, 0.301, 0.340	0.6879214
E-max 3 (3)	–0.804, –0.123, 0.133, 0.349, 0.445	0.6782399
Exponential 1 (6.2)	–0.378, –0.319, –0.231, 0.097, 0.832	0.6843346
Exponential 2 (3)	–0.285, –0.274, –0.251, -0.071, 0.881	0.6890958
Linear[a]	–0.540, –0.332, –0.125, 0.291, 0.706	0.6830898
Linear-log (3.3)	–0.732, –0.210, 0.048, 0.351, 0.543	0.670295
Logistic 1 (0.13,0.732)	–0.892, 0.165, 0.241, 0.243, 0.243	0.714469
Logistic 2 (5,1)	–0.531, –0.459, –0.016, 0.500, 0.507	0.8027349
Logistic 3 (6,2)	–0.517, –0.399, –0.130, 0.460, 0.587	0.7642835
Quadratic (-0.029)	–0.643, –0.307, –0.023, 0.386, 0.586	0.6840342
SigEMax 1 (20.46,0.5)	–0.809, –0.093, 0.106, 0.329, 0.467	0.662415
SigEMax 2 (0.89,2.4)	–0.892, 0.162, 0.233, 0.247, 0.249	0.7117789

Table 2 Power Calculations for Actigraph Sleep Efficiency Under Different Dose-Response Curves (n=13 and sigma=12)

Model (R variables)	Optimal Contrasts (PBO, 2.5, 5, 10, 15)	Power (%)
BetaMod 1 (0.339,0.05)	–0.82, –0.074, 0.116, 0.336, 0.442	0.6990587
BetaMod 2 (0.5,0.3)	–0.834, –0.073, 0.187, 0.414, 0.305	0.6908521
BetaMod 3 (0.19,0.15)	–0.887, 0.124, 0.236, 0.304, 0.223	0.7140683
E-max 1 (0.66)	–0.883, 0.095, 0.209, 0.277, 0.301	0.7286372
E-max 2 (1.1)	–0.869, 0.033, 0.195, 0.301, 0.340	0.7206645
E-max 3 (3)	–0.804, –0.123, 0.133, 0.349, 0.445	0.7098843
Exponential 1 (6.2)	–0.378, –0.319, –0.231, 0.097, 0.832	0.7170187
Exponential 2 (3)	–0.285, –0.274, –0.251, -0.071, 0.881	0.7225359
Linear[a]	–0.540, –0.332, –0.125, 0.291, 0.706	0.7148674
Linear-log (3.3)	–0.732, –0.210, 0.048, 0.351, 0.543	0.7011122
Logistic 1 (0.13,0.732)	–0.892, 0.165, 0.241, 0.243, 0.243	0.7477969
Logistic 2 (5,1)	–0.531, –0.459, –0.016, 0.500, 0.507	0.8321559
Logistic 3 (6,2)	–0.517, –0.399, –0.130, 0.460, 0.587	0.7948821
Quadratic (-0.029)	–0.643, –0.307, –0.023, 0.386, 0.586	0.7160553
SigEMax 1 (20.46,0.5)	–0.809, –0.093, 0.106, 0.329, 0.467	0.6949806
SigEMax 2 (0.89,2.4)	–0.892, 0.162, 0.233, 0.247, 0.249	0.7447213

Table 3 Power Calculations for Actigraph Sleep Efficiency Under Different Dose Response Curves (n=14 and sigma=12)

Model (R variables)	Optimal Contrasts (PBO, 2.5, 5, 10, 15)	Power (%)
BetaMod 1 (0.339,0.05)	–0.82, –0.074, 0.116, 0.336, 0.442	0.7293262
BetaMod 2 (0.5,0.3)	–0.834, –0.073, 0.187, 0.414, 0.305	0.7217665
BetaMod 3 (0.19,0.15)	–0.887, 0.124, 0.236, 0.304, 0.223	0.7453197
E-max 1 (0.66)[a]	–0.883, 0.095, 0.209, 0.277, 0.301	0.7587574
E-max 2 (1.1)	–0.869, 0.033, 0.195, 0.301, 0.340	0.7508701
E-max 3 (3)	–0.804, –0.123, 0.133, 0.349, 0.445	0.7396949
Exponential 1 (6.2)	–0.378, –0.319, –0.231, 0.097, 0.832	0.7473776
Exponential 2 (3)	–0.285, –0.274, –0.251, -0.071, 0.881	0.7537709
Linear[a]	–0.540, –0.332, –0.125, 0.291, 0.706	0.7447648
Linear-log (3.3)	–0.732, –0.210, 0.048, 0.351, 0.543	0.7311277
Logistic 1 (0.13,0.732)	–0.892, 0.165, 0.241, 0.243, 0.243	0.7783044
Logistic 2 (5,1)	–0.531, –0.459, –0.016, 0.500, 0.507	0.8576103
Logistic 3 (6,2)	–0.517, –0.399, –0.130, 0.460, 0.587	0.8224386
Quadratic (-0.029)	–0.643, –0.307, –0.023, 0.386, 0.586	0.7454875
SigEMax 1 (20.46,0.5)	–0.809, –0.093, 0.106, 0.329, 0.467	0.725321
SigEMax 2 (0.89,2.4)	–0.892, 0.162, 0.233, 0.247, 0.249	0.7753528

4 STATISTICAL METHODS

All descriptive statistics for continuous variables will be reported using mean, standard deviation (SD), median, minimum and maximum. Categorical variables will be summarized as number (percentage) of subjects.

All statistical tests will be based on the 5% level of significance (2-sided), unless otherwise stated. If statistical comparisons are not defined, all pairwise comparisons will be tested.

4.1 Study Endpoints

4.1.1 Sleep-Related Endpoints

The sleep-related endpoints are:

- The change from baseline of aSE during each week of treatment with LEM compared to PBO
- Change from baseline in mean SFI during each week of treatment
- Change from baseline of the aMeanDurWB during each week of treatment

4.1.2 Wake-Related Endpoints

- The change from baseline of mean aWE with LEM compared to PBO during each week of treatment
- Change from baseline of the aMeanDurSB during each week of treatment
- Change from baseline of mean WFI during each week of treatment

4.1.3 Circadian Rhythm-Related Endpoints

- Change from baseline of IV, IS, L5, M10, AMP and RA over each week of treatment

4.1.4 Additional Endpoints

- The following additional endpoints will be explored for LEM2.5, LEM5, LEM10 and LEM15 compared to PBO:
- Safety and tolerability of LEM, including AEs and SAEs
- Change from baseline in Clinician's Global Impression of Change-ISWRD (CGIC-ISWRD) Scale on symptoms of ISWRD total score and domains. (revised per Amendments 01 and 06)
- Change from baseline of the sum of activity counts and change from baseline in the number of bouts >10 minutes of sleep in the first 3 hours after morning waketime on the first 3 days and last 3 days of treatment.
- Number and percentage of subjects in each category of the CGIC-ISWRD Scale at Day 29 (revised per Amendment 01)
- Rebound sleep and wake fragmentation endpoints as assessed from actigraphy during the Follow-Up Period
- Change from baseline in mean aSE of the first 7 nights, and aSE of the second 7 nights of the Follow-Up Period
- Change from baseline in mean aWE of the first 7 days and mean aWE of the second 7

days of the Follow-Up Period

- Proportion of subjects whose mean aSE is higher than at baseline for the first 7 nights or the second 7 nights of the Follow-Up Period
- Proportion of subjects whose mean aWE is longer than at baseline for the first 7 days or the second 7 days of the Follow-Up Period (revised per Amendment 03)
- Change from baseline on ADAS-Cog at Day 29
- Change from baseline on MMSE at Day 29
- Change from baseline sleep quality in caregivers as measured by the PSQI at Day 29
- Change from baseline of caregiver burden on all scores of the ZBI-short form at Day 29. (revised per Amendment 06)
- Change from baseline on the EQ-5D-5L utility and Visual Analogue Scale (VAS) scores at Day 29 for both subject and caregiver
- Change from baseline of the total score of NPI-10 at Day 29
- Change from baseline on SDI at Day 29
- Characterize the PK of lemborexant using the population approach and descriptive statistics for the plasma concentrations of its metabolites M4, M9, and M10. (revised per Amendment 06)
- Relationships between exposure to lemborexant, efficacy, and/or safety variables using PK/PD modeling
- Assess the plasma concentrations of cognitive enhancers and lemborexant in subjects taking both drugs.
- Evaluate the long-term safety and tolerability of flexible doses of LEM5, LEM10, and LEM15 per day over a period of 30 months in subjects with ISWRD who have completed the Core Study. (revised per Amendments 05 and 06)

4.2 Study Subjects

4.2.1 Definitions of Analysis Sets

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

The Full Analysis Set (FAS) is the group of randomized subjects who received at least 1 dose of randomized study drug.

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

The PK/PD Analysis Set is the group of subjects receiving either lemborexant or placebo 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.

4.2.2 Definitions of Analysis Sets (Open Label Extension)

The Safety Analysis Set is the group of subjects who received at least 1 dose of study drug during the open label extension.

4.2.3 Subject Disposition

The number of subjects screened and the number failing screening (overall and by reason for failure) will be summarized. Screen failure data will be listed. The number of subjects randomized along with the number of subjects in each of the study populations will also be presented.

The number of subjects completing the core study will be presented. Subjects who prematurely terminated their participation in the study will be summarized by their primary reason for study termination. Other reasons for study drug and study terminations will also be summarized. These tabulations will be produced for all randomized subjects by treatment group.

The number of subjects enrolling in the open label extension will be presented.

The number and percentage of subjects will be summarized with the number in the FAS as the denominator, together with the number of subjects randomized but never dosed:

- Randomized but never dosed (only showing number of subjects)
- Completed treatment
- Prematurely discontinued treatment and the reasons for discontinuations
- Completed study
- Prematurely discontinued the study and the reasons for discontinuations

Kaplan-Meier analysis for time to treatment discontinuation will be performed to produce graphical presentations of the treatment completion curves and to estimate treatment discontinuation rates by treatment group. Time to treatment discontinuation (in days) is calculated as: last dose date – first dose date +1. Subjects who complete the treatment period will be considered to have no event. Subjects who prematurely discontinued treatment will be considered to have an event at the time of treatment discontinuation.

A listing will be provided for subjects who discontinued treatment or who discontinued study with reasons for discontinuation. A randomization listing will be provided with subjects ordered by randomization date.

4.2.4 Protocol Deviations

Protocol deviations, as specified by the monitor, will be categorized into major and minor deviations prior to unblinding and will be listed. Major protocol deviations will be summarized by treatment group.

4.2.5 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics for the Safety Analysis Set will be summarized for each treatment group using descriptive statistics. Continuous demographic and baseline variables include age, height, and weight; body mass index (BMI); categorical variables include sex, age group (<60, 60 to 74 years, 75 to 84 years, and 85 and older), BMI group (<18.5, 18.5 to <25, 25 to ≤30, and >30 kg/m^2, or other suitable categorization), race, ethnicity, and country. Other baseline characteristics for actigraphy data and MMSE will also be summarized. Caregiver demography including relationship of the caregiver to the subject, caregiver's sleeping location, etc.) will also be summarized.

If the Safety Analysis Set and FAS differ substantially, the demographic summaries will be repeated on the FAS.

4.2.5.1 Medical History

The number (percentage) of subjects in the Safety Analysis Set reporting a history of any medical condition, as recorded on the CRF, will be summarized for each treatment group and overall. A subject data listing of medical and surgical history will be provided.

4.2.6 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 2016 or latest version). 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 WHO DD -preferred term (PT). 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.

Concomitant medications are defined as medications that (1) started before the first dose of study drug 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 to the last dose day plus 14 days. All medications will be presented in subject data listings.

4.2.7 Treatment Compliance

Compliance for each study drug will be calculated on the basis of the number of tablets dispensed, lost and returned, separately for each row of tablets and overall, for all randomized subjects.

Compliance = (Total number of tablets dispensed - total number of tablets returned or lost) x 100
Number of tablets expected to be taken by the subject for the treatment period

Summaries will provide descriptive summary statistics and number (percentage) of subjects using the categories <80%, >= 80% to <= 100%, >100% to <= 120%, and > 120% for each treatment group.

If deemed necessary, treatment compliance will be calculated for the Safety Analysis Set and FAS as well.

4.3 Data Analysis General Considerations

The FAS will be used as the primary population for all efficacy analyses.

4.3.1 Pooling of Centers

This study was a multicenter study across the United States, UK and Japan, with approximately 40-50 sites across the two countries. As country is used as randomization stratification factor, country will be used as a covariate in many planned analyses. Therefore, data from centers of the same country will be pooled together as strata. It was planned that each country would enroll a minimum of 20 subjects in the study. UK subjects will be pooled with US subjects if deemed necessary. In the case of low enrollment as the study approaches completion, decisions regarding eliminating country as a factor in the analysis models will be made and documented before treatment unblinding.

4.3.2 Adjustments for Covariates

In the statistical models, the covariate of country will be included in the model since country was used for the randomization stratification. Details of adjustment for covariates for other efficacy analyses are defined in the description of each individual efficacy analysis.

4.3.3 Multiple Comparisons/Multiplicity

No multiplicity adjustments will be made. The study is aimed to find the best dose response curve for each endpoint and will not be used for labelling claims.

However, this proof of concept (PoC) study uses the multiple comparison procedures-modeling (MCP-Mod) approach with set of candidate dose response models and test for a dose response relationship via model associated statistics. The multiplicity adjusted p-value and critical values are numerically computed based on the assumption that the multiple T test statistics, are jointly multivariate *t*-distributed under the null hypothesis of no dose response.

4.3.4 Examination of Subgroups

Due to the sample size reduction, there will bo no formal subgroup analysis. Basic summary statistics may be explored for various subgroups in order to assess whether the treatment effects are consistent across different subgroups. These subgroups include: age group (<60, 60 to 74

years, 75 to 84 years, and 85 and older), sex (male and female), race (White, Black or African American, Japanese, Chinese, Other Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other), country (USA, Japan), BMI group (< 18.5, >=18.5 to <25, >=25 to <30, and >=30), severity of baseline aSE (less than 65%, 65% to 75%), severity of baseline aWE (less than 75%, 75% to 87%), frequency of complaint of sleep and wake fragmentation at baseline (4, 5, 6 or 7 days per week, maybe subsumed depending on data), APOε4 genotype, and severity of AD-D based on MMSE (10 to 20, greater than 20 to 25). Age group is classified by the subject's age at the time of Informed Consent; sex and race are defined as entered on the study CRF, country is the same as that used for the randomization stratification; BMI group is defined by weight taken from the baseline vital sign visit; and severity of aSE, aWE, MMSE and frequency of complaint of sleep and wake fragmentation are from data collected during screening.

For some subgroup variables, in the event that the sample size in a subgroup is so small that interpretation of the results is challenging, some subgroups with small sample sizes may be regrouped. For example, BMI group can be re-grouped as <25 and >=25; race can be re-grouped as White, Black or African American, Japanese, and All Others. If re-grouping of subgroups is warranted, it will be documented before treatment unblinding.

Additional subgroup analyses may also be conducted, such as subgrouping by combinations of aforementioned variables or other criteria of clinical relevance, if deemed necessary to provide additional understanding of study data, interpretation of results, or to address questions of clinical interest, and will be documented in the updates of this SAP or clinical study report.

4.3.5 Handling of Missing Data, Dropouts, and Outliers

Specific methods for handling missing values in calculating the efficacy endpoints (aSE and aWE) are detailed in Section 8 Definitions and Conventions for Data Handling. In general, missing values due to subject discontinuation, missed or unusable assessments will not be imputed.

Since this is actigraphy study, and as long as subject wears his/her actigraph during the last week of treatment, primary endpoint will most likely not be missing.

Incomplete/Missing data: Will not be imputed, unless otherwise specified; i.e., all missing values will remain as missing in all statistical analyses and listings, unless otherwise specified.

Outliers: No formal statistical analyses will be performed to detect and/or remedy the presence of statistical outliers.

4.3.6 Other Considerations

Individual subject data in the database will be presented in data listings.

The following estimands (Table 4) are evaluated for the primary efficacy endpoint (aSE and aWE change from baseline during last week of treatent) 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 4.4.

Table 4

Estimate	Description	Popu-lation	Intervention Effect of Interest	Analysis Type
Dose Response: Testing the hypothesis that (the dose-response curve is not flat)	Establish a dose-response signal using multiple comparison procedure. Based on the observed data, the model that shows a statistically significant trend test will be selected (at one-sided 5% If more than one is statistically significant, then the most optimal model using Akaike Information Criteria will be selected. This method prospectively controls the type I error at 5%.	FAS	Subjects who completed all efficacy endpoint assessments (aSE or aWE) during the 4th week of treatment without missing the week 4 value.	(This is first important analysis since there is no primary or secoandary analysis in this proof of concept study).
Difference in outcome improvement for all randomized subjects	- all randomized subjects regardless of what treatment subjects actually received - include data after dropout	FAS	Missing values will not be imputed; MMRM model is used on all available data assuming MAR (Assumes subjects with missing values behave the same as the observed data within that treatment group, i.e., the missingness is independent of unobserved data after accounting for the observed data in the model. Thus the dropouts or subjects with missing values may continue to benefit from the treatment as if they were still on treatment (just like completers).	(This is second important analysis since there is no primary or secoandary analysis in this proof of concept study). (MMRM analysis assuming MAR)

FAS = full analysis set; MAR = missing at random, MMRM = mixed effect for repeated measures model

4.3.7 Visit Window

All efficacy and safety data will be presented using nominal visits as reported in the data. No additional derivation will be performed for analysis visit window. Actigraphy data will be presented by visit as derived by the actigraphy vendor.

4.4 Efficacy Analyses

Efficacy analyses will be performed on the FAS unless otherwise specified.

4.4.1 Efficacy Analysis for the Sleep-Related, Wake-Related, and Circadian Rhythm-Related Endpoints

The change from baseline of the mean aSE and mean aWE (separately) for the last 7 nights/days on treatment will be analyzed using MCP-MOD (Multiple Comparisons and Modelling) approach. Dose response models that will be evaluated are linear, linear log, quadratic, exponential, e_{max}, sigmoid e_{max}, beta and logistic.

For all actigraphy parameters, baseline will be defined as the average value of the last 7 days of screening. For IS, IV, RA, AMP, L5, M10 parameters, the weekly averages are calculated by actigraphy vendor. For these variables, the last record of screening period will be considered as the baseline (the average of the last 7 days) of screening period.

Analysis stage – MCP-step: Establish a dose-response signal (the dose-response curve is not flat) using multiple comparison procedure. Based on the observed data, the model that shows a statistically significant trend test will be selected (at one-sided 5% significance). If more than one is statistically significant, then the most optimal model using Akaike Information Criteria will be selected. This method prospectively controls the type I error at 5%.

Analysis Stage – Mod-step: Dose response and target dose estimation will be based on dose response modelling. MCP-MOD approach allows for interpolation between doses.

The p-value and Akaike Information Criteria for the trend test will be reported for all models explored for the MCP-step. The dose response and target dose estimation for the chosen model will be reported for the Mod-step.

The change from baseline of secondary endpoints mean SFI, mean WFI, mean aMeanDurSB, mean aMeanDurWB, IV, IS, L5, M10, AMP and RA of the last 7 days of treatment (ie, Week 4) will be analyzed using mixed methods for repeated measures (MMRM) model on the FAS for LEM2.5, LEM5, LEM10 and LEM15 compared to PBO, as appropriate. The model will include all data and will be adjusted for the corresponding baseline value, country, treatment, time (Week 1, Week 2, Week 3 and Week 4) and the interaction of treatment by time. Treatment by time interaction will be used to construct the treatment comparisons at a specific time. The MMRM model accounts for any missing data, and assumes that the missing data are missing at random. An unstructured covariance matrix will be used, and if the model fails to converge, then an autoregressive matrix will be used. Where data are normally distributed, least square (LS) means, difference in LS means of each lemborexant dose compared to PBO, 95% CIs and P values at the appropriate time point will be presented.

The overall score of CGIC-ISWRD Scale at Day 29 will be analyzed using the Cochran–Mantel–Haenszel (CMH) test, adjusted for country.

4.4.2 Subgroup Analyses

Subgroup analysis, as deemed necessary, will be presented using basic summary statistics. No hypothesis testing will be made on subgroup analysis.

4.4.3 Additional Analyses

The mean aSE and aWE change from baseline will be analyzed using the MMRM. The linear least squares will be performed on aSE and aWE change from baseline at Week 1, Week 2, Week 3 and Week 4 of treatment on the FAS for LEM2.5, LEM5, LEM10 and LEM15 compared to PBO, where the model will be explored with baseline value, country and treatment as covariates/factors, and with and without age and MMSE score as further exploratory covariates. Least square (LS) means, difference in LS means of each lemborexant dose compared to PBO, 95% CIs and P values at the appropriate time point will be presented.

4.5 Other Efficacy: Exploratory and Pharmacodynamic Analyses

To assess residual morning sleepiness levels, change from baseline in the sum of activity counts in the 3 hour interval after morning waketime will be compared for each treatment group relative to placebo for each of the 6 mornings comprising the first 3 days and last 3 days of treatment. The change from baseline will be analyzed using analysis of covariance (ANCOVA), with treatment and baseline as fixed effects. LS means, difference in LS means of LEM2.5, LEM5, LEM10 and LEM15 compared to PBO, 95% CIs, and p-values will be presented. In addition, the change from baseline of the number of bouts >10 min scored as sleep will be determined. The change from baseline will be analyzed using ANCOVA, as above.

Rebound sleep and wake fragmentation is defined as worsened aSE or aWE compared to baseline after study drug treatment is discontinued. Actigraphy data from the Follow-Up Period will be compared to actigraphy data from the baseline to assess whether subjects experience rebound sleep or wake fragmentation. Specifically, a lower value for aSE or aWE during the Follow-Up Period compared to the mean aSE or aWE value during baseline will be considered worsened sleep or wake fragmentation.

To assess rebound sleep and wake fragmentation, both categorical analysis at the subject level and continuous analysis at the group mean level will be performed. For each of the 2 weeks of the Follow-Up Period the proportion of subjects whose corresponding value for aSE or aWE is less than the corresponding baseline value by 5% for aSE and 1% for aWE (which is approximately minutes based on the a typical 8-hour sleep period and 16-hour wake period) will be summarized by treatment group and compared to placebo. The percentage of ‘rebounders’ between each treatment and placebo group will be analyzed using a CMH test, adjusted for country.

To assess statistical significance using the continuous data at the group mean level, the data will be analyzed using ANCOVA, adjusted for country. The LS mean of each week of the Follow-Up Period will be compared to the baseline between each treatment group and placebo. If the upper bound of the 95% CI of aSE or aWE for the mean of each week of the Follow-Up Period is less

than the lower bound of a 95% CI for the values during the baseline in the given treatment group, it will be considered strong evidence for rebound sleep or wake fragmentation. If the LS means for aSE and aWE for the Follow-Up Period are all higher than for the baseline, then no rebound sleep or wake fragmentation is suggested. Otherwise, the degree to which the parameters worsen, and the time point(s) at which they worsen will be considered to evaluate whether clinically meaningful rebound sleep and/or wake fragmentation is present.

Each domain of CGIC-ISWRD Scale and the total score at Day 29 will be analyzed using the CMH test, adjusted for country.

The change from baseline of the EQ-5D-5L utility and VAS scores, the global score of PSQI, all scores of the ZBI-short form and the SDI at Day 29 will be analyzed using ANCOVA, with treatment and baseline as fixed effects. Provided that the data are normally distributed, LS means, difference in LS means of LEM2.5, LEM5, LEM10 and LEM15 compared to PBO, 95% confidence intervals (CIs) and p-values will be presented.

For the PSQI, caregivers who do not attend to the subject in the night will not be included in the summaries and analysis. Caregiver data will not be included in analyses if they do not live in the same dwelling as the subject, and will be presented as a subgroup data for exploratory purposes.

4.6 Pharmacokinetic, Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

4.6.1 Pharmacokinetic Analyses

The Safety Analysis Set will be used for individual plasma concentration listings of lemborexant and its metabolites M4, M9, and M10. The PK Analysis Set will be used for summaries of plasma concentrations of lemborexant and its metabolites M4, M9, and M10, by dose and day.

Further details of the PK and population PK analyses will be provided in a separate analysis plan.

4.6.2 Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

DNA samples will be collected and stored, and may be used to examine the role of genetic variability in absorption, distribution, metabolism, and excretion, or development of AEs. Variations in lemborexant exposure or AEs may be explored by correlation of single nucleotide polymorphisms with PK, safety, or efficacy data. Efficacy may be explored in relation to the APOε4 genotype as follows:

The change from baselines of the mean aSE and mean aWE (separately) for the last 7 nights/days on treatment may be explored by APOε4 genotype comparing total lemborexant to placebo.

4.7 Safety Analyses

Safety analyses will be performed on the Safety Analysis Set unless otherwise specified. Safety data, presented by treatment group, will be summarized on an "as treated" basis using descriptive statistics (eg, n, mean, standard deviation, median, minimum, maximum for continuous variables; n [%] for categorical variables). Safety variables include treatment-emergent adverse events (TEAEs), clinical laboratory parameters, vital signs, 12-lead ECG results, electronic Columbia-Suicide Severity Rating Scale (eC-SSRS). Study Day 1 for all safety analyses is defined as the date of the first dose of study drug.

4.7.1 Extent of Exposure

The extent of exposure (mean daily dose, cumulative dose, duration of exposure) to study drug will be summarized descriptively for lemborexant.

Compliance for lemborexant will be calculated on the basis of the number of tablets dispensed, lost, and returned, separately for each type (dose) of tablet. Summaries will provide descriptive summary statistics and number (percentage) of subjects below 80%, between 80% and 120%, and greater than 120%.

4.7.2 Adverse Events

The AE verbatim descriptions (investigator terms from the eCRF) will be classified into standardized medical terminology using the Medical Dictionary for Regulatory Activities (MedDRA). Adverse events will be coded to the MedDRA (Version 18.0 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.

A treatment-emergent adverse event (TEAE) is defined as an AE that emerges during treatment, having been absent at pretreatment (Baseline) or

- Reemerged during treatment, having been present at pretreatment (Baseline) but stopped before treatment, or
- Worsened in severity during treatment relative to the pretreatment state, when the AE is continuous.

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.

An overview of the TEAEs will be summarized by treatment group, including the number and percentage of subjects who experience TEAEs, treatment-related TEAEs, severe TEAEs, TEAEs leading to death and discontinuation from study/study drug.

The incidence of TEAEs will be reported as the number (percentage) of subjects with TEAEs by SOC and PT. A subject will be counted only once within an SOC and PT, even if the subject experienced more than 1 TEAE within a specific SOC and PT. The number (percentage) of

subjects with TEAEs will also be summarized by maximum severity (mild, moderate, or severe). The number (percentage) of non-serious TEAEs with an incidence of greater than 5% will be summarized.

The number (percentage) of subjects with TEAEs will also be summarized by relationship to study drug (Yes [related] and No [not related]). Treatment-related TEAEs include those events considered by the investigator to be related to study treatment.

The number (percentage) of subjects with TEAEs leading to death will be summarized by MedDRA SOC and PT for each treatment group. A subject data listing of all AEs leading to death will be provided.

The number (percentage) of subjects with treatment-emergent SAEs will be summarized by MedDRA SOC and PT for each treatment group. A subject data listing of all SAEs will be provided.

The number (percentage) of subjects with TEAEs leading to discontinuation from study drug will be summarized by MedDRA SOC and PT for each treatment group. A subject data listing of all AEs leading to discontinuation from study drug will be provided.

Adverse events will be summarized by the following subgroups: age (≤64 years, 65 to 74 years, 75 to 84 years, ≥85 years), sex (male, female), race (white, black, Asian, other), and country

The number (percentage) of subjects with TEAEs of cataplexy that are characterized according to the customized MedDRA query PT as potential cataplexy-related events (Section 9.5.1.5 of the Protocol), or as seizure-related events will be summarized as deemed necessary. Adjudicated events will also be presented separately.

4.7.3 Laboratory Values

Clinical laboratory values will be evaluated for each laboratory parameter by subject. Abnormal laboratory values will be identified as those outside (above or below) the normal range. Reference (normal) ranges for laboratory parameters will be included in the clinical study report for this study. Descriptive summary statistics (eg, mean, SD, median, minimum, maximum for continuous variables, and number and percentage for categorical variables) for the laboratory parameters and changes from baseline will be evaluated by treatment group and visit.

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. Within-treatment comparisons will be based on 3 by 3 tables (shift tables) that, for a particular laboratory test, compare the Study Baseline LNH classification to the LNH classification at end of study/early termination, by treatment group.

Clinical laboratory results post-baseline will be evaluated for markedly abnormal values. A laboratory test will be considered markedly abnormal if the result worsens to meet Eisai grading criteria for laboratory values limit of Grade 2 or higher. If the Grade 2 limit is missing, the Grade 1 limit will be considered. Appendix 12.1 presents the Eisai grading criteria for

laboratory values that were used to identify subjects with markedly abnormal laboratory values. For the incidence of markedly abnormal laboratory values, each subject may be counted once in the laboratory parameter value high and in the laboratory parameter low categories as applicable.

4.7.4 Vital Signs

Descriptive statistics for vital signs parameters (ie, diastolic and systolic BP, pulse, respiration rate, temperature) and weight, and changes from Baseline will be presented by visit and treatment group.

Vital sign values will be listed. Clinically notable vital sign values will be identified on the listings as those above (H) or below (L) a clinically notable range (Table). Categorical analyses of subjects (number and percent) who fall outside the below clinically notable vital sign ranges will also be presented by treatment group and visit.

Table 5 Vital Sign Criteria

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

BP = blood pressure, bpm = beats per minute, H = high, L = low.
Clinically notable means that a value must meet the criterion value and must attain the specified magnitude of change relative to baseline.

4.7.5 Electrocardiograms

Descriptive statistics for ECG parameters and changes from Baseline will be presented by treatment group. Shift tables will present changes from Baseline in ECG interpretation (categorized as normal; abnormal, not clinically significant; and abnormal, clinically significant) by time point.

For each subject, the maximum observed corrected QT interval calculated using Fridericia's formula (QTcF), and the maximum prolongation from baseline in QTcF will be compiled. Categorical analyses of subjects (number and percent) with maximum observed QTcF values >450 msec, >480 msec, and >500 msec and maximum prolongations (from Baseline) in QTcF >30 msec and >60 msec will be presented by treatment group and by time point. Categorical analyses of subjects (number and percent) with maximum observed PR values >220 msec, and QRS values > 120 msec will be presented by treatment group and by time point.

4.7.6 Other Safety Analyses

The results of eC-SSRS assessments will be listed for each subject. The incidence of treatment-emergent suicidal ideation or suicidal behavior will be summarized by treatment group using descriptive statistics as appropriate.

The change from baseline of total score from the NPI-10, ADAS-Cog and MMSE at Day 29 will be analyzed with ANCOVA, with treatment and baseline as fixed effects on the Safety Analysis Set. Provided that the data are normally distributed LS means, difference in LS means of LEM2.5, LEM5, LEM10 and LEM15 compared to PBO, 95% confidence intervals (CIs) and p-values will be presented.

5 INTERIM ANALYSES

No formal interim analyses will be performed for this study. However, due to the fact that the last subject has enrolled at a very late stage in this study, analysis for top-line results will be performed based on the first 61 randomized subjects without unblinding the study site and CRO regarding the last subject. All TLGs then will be produced based on all data when all subjects complete the study.

6 CHANGES IN THE PLANNED ANALYSES

It has been decided to stop the study early after enrolling about 60 subjects while original sample size was 125 subjects. Due to the reduction of sample size, no formal subgroup analysis will be performed.

7 DEFINITIONS AND CONVENTIONS FOR DATA HANDLING

Actigraph data will be excluded from the analysis where there are fewer than 5 complete days out of 7 days' data. A day will be considered complete as long as data from 90% of the 24-hour period are scorable. Actigraphy vendor will not derive actigraphy parameters for weeks where subject has fewer than 5 complete days out of 7 days' data, and therefore such data will not be available for analysis nor it will be imputed.

Actigraph data will be calculated based on habitual TIB and waketime values for the protocol described analyses, and also based on actual TIB and waketime values as an exploratory analysis.

8 PROGRAMMING SPECIFICATIONS

For all actigraphy parameters, baseline will be defined as the average value of the last available seven days' values of screening period . For all other safety and efficacy enpoints, the last measurement before dosing will be defined as the baseline. Unless it is specified otherwise.

All efficacy and safety data will be presented using nominal visit as reported in the data. No additional derivation will be performed for analysis visit window.

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

9 STATISTICAL SOFTWARE

All statistical analyses will be performed using SAS v 9.3 or later.

10 MOCK TABLES, LISTINGS, AND GRAPHS

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

11 REFERENCES

Bornkamp B, Pinheiro J, Bretz F. MCPMod: An R Package for the Design and Analysis of Dose-Finding Studies. J Statistical Software. 2009 (7):29.

Gehrman PR, Marler M, Martin JL, et al. The timing of activity rhythms in patients with dementia is related to survival. J Gerontol Biol Sci Med Sci. 2004;59A:1050–1055.

Marler MR, Gehrman P, Ancoli-Israel S, et al. The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with non-sinusoidal shapes. Stat Med. 2006;25:3893–3904.

Martin et al, Sleep and Sleep Disorders in Older Persons: A Multicomponent Nonpharmacological Intervention Improves Activity Rhythms Among Nursing Home Residents With Disrupted Sleep/Wake Patterns Journal of Gerontology: MEDICAL SCIENCES, 2007, Vol. 62A, No. 1, 67–72

Martin J, Marler MR, Shochat T, et al. Circadian rhythms of agitation in institutionalized patients with Alzheimer's disease. Chronobiol Int. 2000;17:405–418.

Mishima K, Okawa M, Hishikawa Y, Hozumi S, Hori H, Takahashi K. Morning bright light therapy for sleep and behavior disorders in elderly patients with dementia. Acta Psychiatrica Scandinavica. 1994; 89: 1–7.

12 APPENDICES

Appendix 12.1 Sponsor's Grading for Determining Markedly Abnormal Laboratory Results

The following table of Sponsor's Grading for Laboratory Values is taken from the protocol.

Sponsor's Grading for Laboratory Values

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

Sponsor's Grading for Laboratory Values

	Grade 1	Grade 2	Grade 3	Grade 4
Glucose, serum-low (hypoglycemia)	<LLN – 55 mg/dL <LLN – 3.0 mmol/L	<55 – 40 mg/dL <3.0 – 2.2 mmol/L	<40 – 30 mg/dL <2.2 – 1.7 mmol/L	<30 mg/dL <1.7 mmol/L life-threatening consequences; seizures
Phosphate, serum-low (hypophosphatemia)	<LLN – 2.5 mg/dL <LLN – 0.8 mmol/L	<2.5 – 2.0 mg/dL <0.8 – 0.6 mmol/L	<2.0 – 1.0 mg/dL <0.6 – 0.3 mmol/L	<1.0 mg/dL <0.3 mmol/L life-threatening consequences
Potassium, serum-high (hyperkalemia)	>ULN – 5.5 mmol/L	>5.5 – 6.0 mmol/L	>6.0 – 7.0 mmol/L hospitalization indicated	>7.0 mmol/L life-threatening consequences
Potassium, serum-low (hypokalemia)	<LLN – 3.0 mmol/L	<LLN – 3.0 mmol/L; symptomatic; intervention indicated	<3.0 – 2.5 mmol/L hospitalization indicated	<2.5 mmol/L life-threatening consequences
Sodium, serum-high (hypernatremia)	>ULN – 150 mmol/L	>150 – 155 mmol/L	>155 – 160 mmol/L hospitalization indicated	>160 mmol/L life-threatening consequences
Sodium, serum-low (hyponatremia)	<LLN – 130 mmol/L	N/A	<130 – 120 mmol/L	<120 mmol/L life-threatening consequences
Triglyceride, serum-high (hypertriglyceridemia)	150 – 300 mg/dL 1.71 – 3.42 mmol/L	>300 – 500 mg/dL >3.42 – 5.7 mmol/L	>500 – 1000 mg/dL >5.7 – 11.4 mmol/L	>1000 mg/dL >11.4 mmol/L life-threatening consequences
Uric acid, serum-high (hyperuricemia)	>ULN – 10 mg/dL ≤0.59 mmol/L without physiologic consequences	N/A	>ULN – 10 mg/dL ≤0.59 mmol/L with physiologic consequences	>10 mg/dL >0.59 mmol/L life-threatening consequences

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

Appendix 12.2 **SAS Code Example for Dose Response Analysis**

PROC MCPMOD DATA = XXXX OUT = XXXXX COMPMETHOD = pval

ALPHA = 0.025;

RESPONSE Resp;

DOSE Dose;

MODELDOSES **0 2.5 5 10 15/** PLACEFF = **50** MAXEFF **= 90;**

EmaxS: MODEL SigEmaxS(**20.46,0.5);**

Linear: MODEL LinearS;

Log: MODEL Log(3.3);

EXP: MODEL EXP(6.2)

Logistic: MODEL LogitS(0.13,0.732);

Quadratic: MODEL QuadS(-0.029);

Beta: MODEL Beta_mods **(0.339,0.05);**

Emax: MODEL EmaxS **(0.66);**

DOSESEL TD (DELTA = **10);**

SELMODEL AIC;

DOSSUMRY OUT = DOSSUMRY_d;

OPTCONT OUT = OPTCONT_d ;

CORMAT OUT = CORMAT_d ;

run;

Note to Programmers: Please use initial points as in the table of sample size determination section.

SIGNATURE PAGE

Author:

PPD

PPD

Date

Approval:

PPD

PPD

Date

PPD

PPD

Date