Cover page

Official Title:

STATISTICAL ANALYSIS PLAN FOR CLINICAL STUDY REPORT
ID-078A301 - Multi-center, double-blind, randomized, placebo-controlled, parallel-group, polysomnography study to assess the efficacy and safety of ACT-541468 in adult and elderly subjects with insomnia disorder

ClinicalTrials.gov Identifier:

NCT03545191

Date of document:

25 March 2020

ACT-541468

STATISTICAL ANALYSIS PLAN

FOR CLINICAL STUDY REPORT

ID-078A301

Version 4

Multi-center, double-blind, randomized, placebo-controlled, parallel-group, polysomnography study to assess the efficacy and safety of ACT-541468 in adult and elderly subjects with insomnia disorder

Document Number: D-20.054

25 March 2020

Revision history

Version date	Version	Implemented change(s)
16 January 2019	1	Initial version
7 August 2019	2	• Section 4.1.4: Clarified definition of treatment received. • Section 5.4.2: Updated the calculation of treatment compliance. • Section 6.1.7: Updated the Markov Chain Monte Carlo (MCMC) imputation model to conform with the multivariate normal distribution assumption (SAS® code updated accordingly in Section 15). • Section 6.3: Added boxplots for sleep architecture parameters, and added measurement for Insomnia Severity Index© (ISI©) Total score. • Section 7.2: Updated the definition of the reporting period for treatment-emergent adverse events (TEAE), and added the summary of cases sent for Independent Safety Board (ISB) adjudication. • Section 7.7: Updated the endpoints used to assess the potential of rebound insomnia.
20 February 2020	3	• Section 5.4.1: Updated summarized categories for the duration of double-blind study treatment. • Section 5.4.2: Updated the calculation of treatment compliance if wallet not returned. • Section 6.1.7: Updated the multiple imputation procedure to avoid imputing nonsensical values (SAS® code updated accordingly in Section 15). • Section 6.3: Added a summary for total sleep time by quarter of the night and sleep onset latency. Added boxplots for total sleep time and % of total sleep time by each quarter of the night. • Section 7: Clarified the handling of End-of-Treatment (EOT) safety assessments in the event of premature study treatment discontinuation. • Section 7.2: Summary of TEAE added back in. • Section 7.4: Updated the definition of subjects at risk of having a marked ECG abnormality during the treatment withdrawal period.

		• Section 12.4: Clarified the definition of baseline. • Section 13.4: Mentioned additional analysis for vital signs and the Columbia-Suicide Severity Rating Scale (C-SSRS).
25 March 2020	4	• Section 5.3.4: Clarified which Anatomical Therapeutic Chemical (ATC) classification level will be presented. • Section 5.4.2: Added a condition for the calculation of the total duration of treatment interruptions. • Section 7.4: Updated definition of subjects at risk of having a marked ECG abnormality for criteria based on change from baseline and clarified how subjects will be counted. • Section 7.9: Added a condition for the assignment of the assessments to a study period based on scheduled visits. • Section 12.2: Added a condition of the double-blind (DB) study period definition. • Section 15.4: Updated SAS® code for the jump to reference (J2R) method.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS 8

1 INTRODUCTION 10

2 STUDY DESIGN AND FLOW 10

3 OBJECTIVES 11

3.1 Primary objectives 12
3.2 Secondary objectives 12
3.3 Safety objectives 12

4 ANALYSIS SETS 12

4.1 Definitions of analysis sets 12
4.1.1 Screened analysis set 12
4.1.2 Full analysis set 12
4.1.3 Per-protocol set 13
4.1.4 Safety set 13
4.1.5 Treatment withdrawal set 13
4.1.6 Pharmacokinetic analysis set 13
4.2 Usage of the analysis sets 14

5 STUDY SUBJECTS' VARIABLES AND ANALYSES 14

5.1 Subject disposition 14
5.1.1 Screening failures 14
5.1.2 Study disposition 15
5.1.3 Study and study treatment completion/discontinuation 16
5.1.4 Study enrollment 16
5.2 Protocol deviations 17
5.3 Subject characteristics 17
5.3.1 Demographics and baseline characteristics 17
5.3.2 Baseline disease characteristics 17
5.3.3 Previous and concomitant diseases at screening 18
5.3.4 Previous and concomitant therapy 18
5.4 Study treatment exposure and compliance 20
5.4.1 Exposure 20
5.4.2 Study treatment compliance 21

6 EFFICACY VARIABLES AND ANALYSES 22

6.1 Analysis of the primary and secondary variables 22

6.1.1 Primary efficacy endpoints 22
6.1.2 Secondary efficacy endpoints 23
6.1.3 Overall testing strategy and hypotheses 23
6.1.4 Statistical model 26
6.1.5 Handling of partially missing data 28
6.1.6 Descriptive analyses 28
6.1.7 Sensitivity analyses 28
6.1.8 Supportive analyses 31
6.1.9 Subgroup analyses 32
6.2 Analysis of other efficacy variables 32
6.3 Analysis of exploratory variables 33
6.4 Estimating minimal clinically important differences 36
6.5 Patient preferences exploratory endpoints 36

7 SAFETY VARIABLES AND ANALYSES 36

7.1 Overview of safety analyses including subgroup analyses 36
7.2 Adverse events 37
7.3 Laboratory tests 39
7.4 Electrocardiography 43
7.5 Vital signs and body weight 44
7.6 Withdrawal symptoms 44
7.7 Rebound insomnia 45
7.8 Next-day residual effect 45
7.9 Columbia-Suicide Severity Rating Scale© 46
7.10 Epworth Sleepiness Scale© 47
7.11 Subgroup analysis of safety variables 48

8 PHARMACOKINETIC VARIABLES AND ANALYSES 48

9 PHARMACOKINETIC/PHARMACODYNAMIC ANALYSES 48

9.1 Exposure-safety analyses 48
9.2 Exposure-efficacy analyses 49

10 GENERAL STATISTICAL METHODOLOGY 49

10.1 Handling of partially missing data 49
10.2 General rules for data presentations 49

11 INTERIM ANALYSES 50

12 GENERAL DEFINITIONS AND DERIVATIONS 50

12.1 Treatment start and end dates 50

12.2 Study periods 50
12.3 Treatment day and Study day 51
12.4 Baseline 52
12.5 Change from baseline 52
12.6 Handling of data when total sleep time is zero 52
12.7 Missing data patterns 52
12.8 Time window definitions for calculating subjective endpoints 53

13 CHANGES OR CLARIFICATIONS TO ANALYSES PLANNED IN THE STUDY PROTOCOL 54

13.1 Changes to the analyses planned in the study protocol 54
13.2 Changes in the conduct of the study / data collection 55
13.3 Clarifications concerning endpoint definitions and related variables or statistical methods 55
13.4 Additional analyses as compared to the study protocol 55

14 REFERENCES 56

15 APPENDICES 57

15.1 SAS® code for Mixed Model for Repeated Measures 57
15.2 SAS® code for multiple imputation with MCMC for monotone missingness mechanism 57
15.3 SAS® code for multiple imputation under MAR assumption 57
15.4 SAS® code for multiple imputation under MNAR assumption for jump to reference method 58
15.5 SAS® code for multiple imputation under MNAR assumption for delta-adjusted (tipping point) 59

LIST OF TABLES

Table 1 Imputation rules for an incomplete or missing concomitant therapy date 20
Table 2 Insomnia Daytime Symptoms and Impacts Questionnaire description 33
Table 3 Imputation rules for an incomplete or missing AE date 39
Table 4 Marked abnormalities in laboratory parameters for reporting 40
Table 5 Marked abnormalities in ECG parameters 44
Table 6 Meaning of the first and last PSG end date at Visits 3, 6 and 8 53
Table 7 Time windows (in days) to calculate weekly averages for certain subjective endpoints 54

LIST OF FIGURES

Figure 1 Study design for ID-078A301 11
Figure 2 Hypothesis testing strategy 25

LIST OF ABBREVIATIONS AND ACRONYMS

AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	Body mass index
BWSQ	Benzodiazepine Withdrawal Symptom Questionnaire
CI	Confidence interval
CSR	Clinical study report
C-SSRS©	Columbia-Suicide Severity Rating Scale©
DB	Double-blind
ECG	Electrocardiogram/graphy
eCRF	Electronic case report form
EODBT	End-of-Double-Blind-Treatment
EOS	End-of-Study
EOT	End-of-Treatment
ESS©	Epworth Sleepiness Scale©
FAS	Full analysis set
IDMC	Independent Data Monitoring Committee
IDSIQ	Insomnia Daytime Symptoms and Impacts Questionnaire
ISAC	Independent Statistical Analysis Center
ISB	Independent Safety Board
ISI©	Insomnia Severity Index©
J2R	Jump to reference
LPS	Latency to Persistent Sleep
LS	Least squares
MAR	Missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MMRM	Mixed model for repeated measures

MNAR	Missing not at random
PD	Protocol deviation
PGA-S	Patient Global Assessment of Disease Severity
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetic
PPS	Per-protocol set
PSG	Polysomnography
PT	Preferred term
REM	Rapid eye movement
SAE	Serious adverse event
SAP	Statistical analysis plan
SCR set	Screened analysis set
SD	Standard deviation
SDS©	Sheehan Disability Scale©
SI	Standard international
sLSO	Subjective Latency to Sleep Onset
SOC	System organ class
SS	Safety set
sTST	Subjective Total Sleep Time
sWASO	Subjective Wake After Sleep Onset
SWS	Slow wave sleep
TBIL	Total bilirubin
TEAE	Treatment-emergent adverse event
TST	Total Sleep Time
TWS	Treatment withdrawal set
ULN	Upper limit of normal
VAS	Visual analog scale(s)
WASO	Wake After Sleep Onset
WHO	World Health Organization

1 INTRODUCTION

This statistical analysis plan (SAP) describes in detail the analyses and data presentation for the clinical study report (CSR) prepared for ID-078A301.

Obvious corrections to address minor formatting errors or spelling mistakes may be performed at the time of analysis without amending this and related documentation (e.g., mock shells).

Data will be analyzed by Idorsia and/or designated contract research organizations supervised by Idorsia using SAS® version 9.4 or higher and R version 3.4.3 or higher.

The analyses for the Independent Data Monitoring Committee (IDMC) closed session meetings will be performed by an Independent Statistical Analysis Center (ISAC) and are detailed in a separate SAP.

Protocol ID-078A301, Final Version 4, dated 30 July 2018 was used when writing this SAP.

A sub-study will be performed to collect data on subject preferences, which aims to enroll at least 360 subjects from the USA and Germany who are part of the approximately 900 subjects enrolled in the ID-078A301. The statistical analyses of the sub-study will be detailed in a separate document and reported separately.

Unless noted otherwise, summaries will be produced by treatment group [see Section 10.2].

2 STUDY DESIGN AND FLOW

This is a 3-month, multi-center, double-blind, randomized, placebo-controlled, parallel-group, Phase 3 study to assess the efficacy and safety of two doses (25 mg and 50 mg) of ACT-541468 in subjects with insomnia disorder.

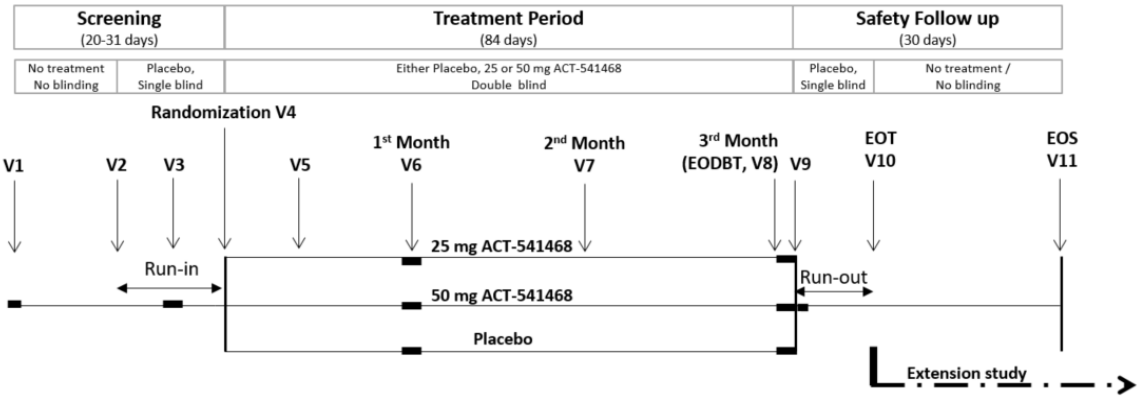
Approximately 900 subjects per study will be randomized to either 25 mg ACT-541468, 50 mg ACT-541468 or placebo, in a 1:1:1 ratio.

Treatment allocation is further stratified by age into two categories: < 65 and ≥ 65 years. Approximately 40% of subjects will be elderly subjects ($\geq$ 65 years old), of which approximately 5% will be above 75 years old.

The study will be conducted at approximately 75 sites in 10 countries.

The study design is shown in Figure 1.

Figure 1 Study design for ID-078A301



▬ = polysomnography nights; EODBT = End-of-Double-Blind-Treatment; EOS = End-of-Study; EOT = End-of-Treatment, V = Visit.

The study comprises the following 3 phases: the screening phase, the treatment phase and the safety follow-up phase.

The **screening phase** starts with the signing of the informed consent form at Visit 1 and ends at Randomization at Visit 4. It includes the *Screening period* (from Visit 1 to Visit 2) and the *Run-in period* (from Visit 2 to Visit 4).

The **double-blind (DB) treatment phase** starts with the first dose of DB study treatment in the first evening of the randomization visit (Visit 4) and lasts until the second morning of Visit 8 after all second morning assessments have been performed, i.e., End-of-Double-Blind-Treatment (EODBT).

The **safety follow-up phase** starts from EODBT and lasts until End-of-Study (EOS). It consists of a single-blind placebo *Run-out period* (from the evening of Visit 9 until after all visit assessments have been performed at Visit 10, i.e., End-of-Treatment [EOT]) and a *Safety follow-up period* (from EOT until EOS).

See Section 12.2 for detailed definitions of the study periods.

Subjects who do not prematurely discontinue DB study treatment and who complete the Run-out period will be eligible to enter the ID-078A303 extension study. The statistical analyses for ID-078A303 CSR will be detailed in a separate document and reported separately.

3 OBJECTIVES

The main objectives of the study are to assess efficacy and safety of ACT-541468 in subjects with insomnia disorder.

Efficacy will be evaluated primarily by the drug effects on sleep onset and sleep maintenance.

3.1 Primary objectives

To evaluate the efficacy of 25 mg and 50 mg ACT-541468 on objective sleep parameters in subjects with insomnia disorder.

3.2 Secondary objectives

To evaluate the efficacy of 25 mg and 50 mg ACT-541468 on subjective sleep parameters and next-day functioning in subjects with insomnia disorder.

3.3 Safety objectives

To assess the safety and tolerability of ACT-541468 in subjects with insomnia disorder during treatment and upon treatment discontinuation.

4 ANALYSIS SETS

4.1 Definitions of analysis sets

A subject must have given informed consent before being included in any analysis set.

The number of subjects in each analysis set defined below will be tabulated. Any subject excluded along with reason(s) for exclusion from Per-protocol set (PPS), Treatment withdrawal set (TWS), or Pharmacokinetic (PK) analysis set will be summarized and listed.

4.1.1 Screened analysis set

The Screened analysis set (SCR set) comprises all subjects who entered screening and have a subject identification number.

Summaries based on the SCR set will be presented as one group (i.e., all subjects), unless specified otherwise. For summaries by treatment group, subjects will be evaluated according to the treatment they have been assigned to.

4.1.2 Full analysis set

The Full analysis set (FAS) comprises all subjects assigned (i.e., randomized) to a DB study treatment.

In order to adhere to the intention-to-treat principle as much as possible:

- Subjects will be evaluated according to the treatment and strata they have been assigned to, which may differ from the treatment they received.
- All available data will be included.

4.1.3 Per-protocol set

The PPS comprises all subjects from the FAS who received at least one dose of DB study treatment and who complied with the protocol sufficiently to be likely to exhibit the treatment effects.

Criteria for sufficient compliance:

- Absence of major protocol deviations (i.e., those leading to exclusion from the PPS), as defined in a separate protocol deviation (PD) document that will be finalized before unblinding treatment codes.
- DB study treatment compliance is ≥ 70% (as calculated in Section 5.4.2).

Subjects not meeting the criteria for sufficient compliance will be excluded from the PPS. Subjects will be evaluated according to the treatment they have been assigned and actual stratum they belong to.

4.1.4 Safety set

The Safety set (SS) comprises all subjects who received at least one dose of DB study treatment.

Subjects will be evaluated according to the actual treatment they received, which may differ from the randomly assigned treatment.

Actual treatment received is defined as:

- the assigned DB study treatment when received at least once,
- the first DB study treatment received if the assigned DB study treatment was never received.

4.1.5 Treatment withdrawal set

The TWS comprises all subjects included in the SS who received at least one dose of single-blind placebo treatment in the Run-out period.

Subjects will be evaluated according to the actual treatment they received, which may differ from the randomly assigned treatment.

4.1.6 Pharmacokinetic analysis set

The PK set comprises all subjects in the SS who have at least one PK sample collected after initiation of DB study treatment.

Subjects will be evaluated according to the actual treatment they received. Subjects receiving placebo will be excluded from the PK set.

4.2 Usage of the analysis sets

The analyses of efficacy endpoints as well as demographic and baseline disease characteristics will be performed using the FAS, and the PPS for certain sensitivity analyses.

The SS will be used for the analysis of safety endpoints (including previous and concomitant medications, and study treatment exposure).

The TWS will be used for the analysis of endpoints used to assess withdrawal symptoms (e.g., Benzodiazepine Withdrawal Symptom Questionnaire [BWSQ] Total score) and rebound insomnia (i.e., objective sleep parameters Wake After Sleep Onset [WASO], Latency to Persistent Sleep [LPS] and subjective TST [sTST]).

The PK set will be used for the analysis of ACT-541468 plasma concentrations 9–10 h post-dose.

5 STUDY SUBJECTS' VARIABLES AND ANALYSES

5.1 Subject disposition

5.1.1 Screening failures

Screening failures will be summarized based on the SCR set and will include:

- Number (%) of subjects who discontinued during the screening phase (based on 'Was the subject randomized?' recorded as 'No' in the 'Randomization' page); this summary will be broken down as follows:
  - Number (%) of subjects who discontinued during the Screening period (based on 'Was the subject randomized?' recorded as 'No' in the 'Randomization' page, and missing 'Treatment start date' in the 'Study Single-Blind [Run-In] Treatment Log' page).
  - Number (%) of subjects who discontinued during the Run-in period (based on 'Was the subject randomized?' recorded as 'No' in the 'Randomization' page, and a non-missing 'Treatment start date' in the 'Study Single-Blind [Run-In] Treatment Log' page).
- Primary reason for screening phase discontinuation (based on reason for not randomized entered on the 'Randomization' page); this summary will be broken down as follows:
  - Primary reason for Screening period discontinuation (based on reason for not randomized entered on the 'Randomization' page for subjects missing 'Treatment start date' in the 'Study Single-Blind [Run-In] Treatment Log' page).
  - Primary reason for Run-in period discontinuation (based on reason for not randomized entered on the 'Randomization' page for subjects with a non-missing 'Treatment start date' in the 'Study Single-Blind [Run-In] Treatment Log' page).

– Primary reason for premature single-blind placebo treatment discontinuation in the run-in period (based on reason for treatment discontinuation entered on the 'Study Single-Blind [Run-In] Treatment Log' page).

All reasons for screening failure will be listed using the SCR set.

Of note, subjects are not permitted to be re-screened.

5.1.2 Study disposition

Study disposition will be summarized based on the SCR set, presented by randomized treatment group and all screened subjects combined, and will include:

- Number of subjects screened.
- Number (%) of subjects treated in the single-blind placebo treatment run-in period (based on a non-missing 'Treatment start date' in the 'Study Single-Blind [Run-In] Treatment Log' page).
- Number (%) of subjects who completed single-blind placebo treatment run-in period (based on 'Reason for treatment stop' entered as 'Completion as per protocol' in the 'Study Single-Blind [Run-In] Treatment Log' page).
- Number (%) of subjects randomized (based on non-missing randomization number).
- Number (%) of subjects treated in the DB study period (based on non-missing 'Treatment start date' in the 'Study Double Blind Treatment Log' page).
- Number (%) of subjects who completed the DB study period (based on 'Reason for treatment stop' entered as 'Completion as per protocol' in the 'Study Double Blind Treatment Log' page).
- Number (%) of subjects treated in the single-blind placebo treatment Run-out period (based on a non-missing 'Treatment start date' in the 'Study Single-Blind [Run-Out] Treatment Log' page).
- Number (%) of subjects who completed single-blind placebo treatment Run-out period (based on 'Reason for treatment stop' entered as 'Completion as per protocol' in the 'Study Single-Blind [Run-Out] Treatment Log' page).
- Number (%) of subjects completing the study (based on 'Did the subject complete the study?' recorded as 'Yes' in the 'End of Study Status' page).
- Number (%) of subjects completing the study and not entering the extension study (based on 'Did the subject complete the study?' recorded as 'Yes' in the 'End of Study Status' page, and 'Did the subject sign the informed consent to participate in the extension study?' recorded as 'No' in the 'Study Extension' page).
- Number (%) of subjects entering the extension study (based on 'Did the subject sign the informed consent to participate in the extension study?' recorded as 'Yes' in the 'Study Extension' page).

5.1.3 Study and study treatment completion/discontinuation

The following summary will be based on the FAS and presented by treatment group and all treatment groups combined:

- Number (%) of subjects who prematurely discontinued DB study treatment (based on reason for treatment stop entered as 'Discontinuation' in the 'Study Double Blind Treatment Log' page).
- Primary reason for premature DB study treatment discontinuation (based on reason for treatment discontinuation entered on the 'Study Double Blind Treatment Log' page).

All reasons for premature DB study treatment discontinuation will be listed based on the FAS.

The following summary will be based on the FAS and presented by treatment group and all treatment groups combined:

- Number (%) of subjects who prematurely discontinued from the study (based on non-missing reason for stopping study entered in the 'End of Study Status' page).
- Primary reason for premature discontinuation from the study (based on reason for stopping study entered in the 'End of Study Status' page).

All reasons for premature study discontinuation will be listed based on the FAS.

The following summary will be based on the TWS and presented by treatment group and all treatment groups combined:

- Number (%) of subjects who prematurely discontinued single-blind placebo treatment in the Run-out period (based on reason for treatment stop entered as 'Discontinuation' in the 'Study Single-Blind [Run-Out] Treatment Log' page).
- Primary reason for premature single-blind placebo treatment discontinuation in the Run-out period (based on reason for treatment discontinuation entered on the 'Study Single-Blind [Run-Out] Treatment Log' page).

All reasons for premature single-blind placebo treatment discontinuation in the Run-out period will be listed based on the TWS.

5.1.4 Study enrollment

The number (%) of subjects randomized will be displayed by country and site based on the FAS by treatment group and all treatment groups combined.

The randomization scheme and codes will be listed for randomized subjects only.

5.2 Protocol deviations

The FAS will be used for the summary of PDs. All PDs and Important PDs will be summarized in separate tables as per pre-specified category (i.e., selection criteria, Investigational Medicinal Product, study conduct/procedure, forbidden medication, withdrawal criteria and other) by treatment group and all treatment groups combined.

A subject with multiple occurrences of a protocol deviation is counted only once per protocol deviation category.

A listing of protocol deviations will be provided using the FAS.

5.3 Subject characteristics

Unless noted otherwise, summaries and listings described in this section will be based on the FAS.

Summaries will be produced by treatment group and all treatment groups combined. Data will be listed individually by subject.

5.3.1 Demographics and baseline characteristics

Demographic data at screening, including age, sex, race, ethnicity, height, weight, body mass index (BMI) and region (US, other), will be summarized and listed.

The number (%) of subjects randomized in each age category (< 18, 18 – < 65, 65 – < 75, 75 – < 85, 65 – < 85, ≥ 85 years at screening), in each BMI category (< 25, 25 – 30, > 30 kg/m^2 at screening), and in each region (US, other) will be further summarized.

The number of subjects treated in each age category (< 18, 18 – < 65, 65 – < 85 and ≥ 85 years at screening) by sex will be summarized based on the SS.

5.3.2 Baseline disease characteristics

Baseline disease characteristics include dissatisfaction with sleep quantity or quality (Y/N), difficulty initiating sleep (Y/N), difficulty maintaining sleep characterized by frequent awakenings or problems returning to sleep after awakenings (Y/N), early morning awakening with inability to return to sleep (Y/N) and sleep disturbance causing significant distress or impairment in social, occupational, educational, academic, behavioral, or other important areas of functioning (Y/N).

These baseline disease characteristics, including the time since insomnia diagnosis (in years) at randomization, will be listed and summarized.

An incomplete (day or month missing) or missing insomnia diagnosis date will be imputed using the 15th day of the month (if only the day is missing) or 1st of July (if day and month are missing).

If the insomnia diagnosis imputed date is after the randomization date, then the randomization date –1 day will be used as insomnia diagnosis date.

5.3.3 Previous and concomitant diseases at screening

Relevant medical history and current medical conditions will be coded using MedDRA terminology.

Any disease or diagnosis is defined as previous if 'Ongoing at informed consent signed' is answered as 'No'; all other diseases/diagnoses are considered as study concomitant (where answer is 'Yes').

Medical history and current (ongoing) medical conditions, excluding insomnia-related conditions and symptoms, will be summarized separately and listed. Summaries will be presented for each treatment group by primary system organ class (SOC) and preferred term (PT). Medical history will be sorted by SOC and PT within each SOC by descending frequency based on all treatment groups combined.

The MedDRA version used for reporting will be specified in the footnote of the applicable output.

5.3.4 Previous and concomitant therapy

Therapies collected will be coded using the WHO Drug Global reference dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system. The WHO Drug Global version used for reporting will be specified in the footnote of the applicable output.

Previous therapies are any treatments for which the end date is prior to signing of the informed consent form. A previous therapy is to be recorded in the 'Previous/Concomitant Medication' page if discontinued less than 30 days prior to signing of the informed consent form, however, all previous therapies recorded will be reported.

Study-concomitant therapies are any treatments that are either ongoing at the signing of informed consent, or initiated during the time from the signing of informed consent up to EOS. The use of all study-concomitant therapies (including contraceptives and traditional and alternative medicines, e.g., plant-, animal-, or mineral-based medicines) is to be recorded in the 'Previous/Concomitant Medication' page.

Double-blind study treatment concomitant therapies (a subset of study-concomitant therapies) are any treatments that are either ongoing at the start of DB study treatment or initiated during the DB study period until 1 day after the last dose of DB study treatment.

Number (%) of subjects having taken at least one previous or concomitant therapy will be summarized by ATC classification level 4 (or next highest available level) and individual preferred name within each ATC classification based on the SS. Previous and concomitant

therapy will be sorted by ATC class and preferred name within each ATC class by descending frequency based on all treatment groups combined.

Summary tables will be provided for previous, study-concomitant and DB study treatment-concomitant therapies separately. An additional summary table will be provided for DB study treatment concomitant therapies at Baseline (i.e., therapy started on, or is ongoing at, the DB study treatment start date). All concomitant therapies will be listed using the SS. The period (e.g., run-in, DB, run-out [Section 12.2]) for which the concomitant therapy started will be displayed in the listing and those related to an adverse event (AE) will be flagged.

The number (%) of subjects currently on cognitive behavioral therapies and reasons for not using cognitive behavioral therapy will be summarized and listed using both the FAS and the SCR set. Cognitive behavioral therapy is not coded.

An incomplete (day or month missing) or missing concomitant therapy date will be imputed as described in Table 1. The 'lower limit' and 'upper limit' refer to the earliest and latest possible dates, respectively. As an example: If concomitant therapy start date is MAR2017 (day missing), the lower limit is 01MAR2017 and the upper limit is 31MAR2017; If concomitant therapy start date is 2017 (day and month missing), the lower limit is 01JAN2017 and the upper limit is 31DEC2017.

Table 1 Imputation rules for an incomplete or missing concomitant therapy date

Field	Incomplete date	Missing date
Concomitant therapy end date	The upper limit.	No imputation; the therapy is considered as ongoing.
Concomitant therapy start date	The rules below apply in the order presented: 1. If the (imputed) concomitant therapy end date is on or after the start of DB study treatment, and if the DB study treatment start falls within the upper and lower limits (inclusive), the DB study treatment start date is used. 2. If the (imputed) concomitant therapy end date is on or after the start of run-in single-blind placebo treatment, and if the run-in single-blind placebo treatment start falls within the upper and lower limits (inclusive), the run-in single-blind placebo treatment start date is used. 3. If the concomitant therapy resolution date is missing, and: a. if the DB study treatment start falls within the upper and lower limits (inclusive), the DB study treatment start date is used; else: b. if the run-in single-blind placebo treatment start falls within the upper and lower limits (inclusive), the run-in single-blind placebo treatment start date is used. 4. In all the other cases, the lower limit is used.	Whichever is the earlier of the concomitant therapy end date or DB study treatment start date.

DB = double-blind.

The purpose of imputing concomitant therapy dates is only to assign a concomitant therapy to a specific treatment phase for the summary tables. No imputed date is considered in the medical evaluation of a concomitant therapy causal relationship to an individual AE.

5.4 Study treatment exposure and compliance

Unless noted otherwise, summaries and listings described in this section will be based on the SS.

5.4.1 Exposure

The duration of single-blind placebo run-in treatment, DB study treatment, and single-blind placebo run-out treatment (including categories: ≤ 1, > 1 – 2, > 2 – 3, > 3 – 4, > 4 – 6,

> 6 – 8, > 8 – 10, > 10 – 11, > 11 – 12, > 12 – 13, > 13 weeks for DB study treatment; 0, > 0 – 2, > 2 – 5, > 5 – 9, > 9 – 14, > 14 – 20, > 20 days for single-blind placebo run-in treatment; and 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, ≥ 10 days for single-blind placebo run-out treatment) and duration in subject-years will be summarized. The definition of a year is 365.25 days.

The duration of treatment (in days) is defined as the difference between the respective (run-in/DB/run-out) treatment end date and the respective (run-in/DB/run-out) treatment start date plus one day. This calculation ignores periods of treatment interruption. The summary for run-out will be based on the TWS.

The duration of treatment along with reason for treatment stop will be listed.

5.4.2 Study treatment compliance

Study treatment compliance will be assessed through study treatment accountability.

Study treatment compliance per wallet dispensed is entered in the electronic case report form (eCRF) for the Run-in period, the DB study period, and the Run-out period and will only be listed.

The following formula will be used to calculate compliance for each of these periods:

Study treatment compliance (%) = [(number of tablets dispensed – number of tablets returned*) / (treatment duration – total duration of treatment interruptions)] × 100

* If a subject did not return his wallet (e.g., it is lost), compliance will not be calculated and the compliance for the period will be set to missing.

Treatment duration and total duration of treatment interruptions in Run-in, DB, or Run-out periods are calculated using the dedicated eCRF page 'Study Treatment Log' as follows:

- Treatment duration (days) = date of last drug intake of DB study treatment (run-in / run-out single-blind placebo treatment) – date of first drug intake of DB study treatment (run-in / run-out single-blind placebo treatment) + 1 day.
- Total duration of treatment interruptions (days) is the sum of all the treatment interruptions' durations. A treatment interruption duration = date restarted DB study treatment (run-in / run-out single-blind placebo treatment) – end date of DB study treatment (run-in / run-out single-blind placebo treatment) due to an interruption –1 day. For example, if a subject stopped his/her treatment on 19MAR2019 and started again on 23MAR2019, as the treatment is taken on the start and end dates, the calculation will be 23MAR2019–19MAR2019 – 1 day = 3 days of interruption. Should the subject interrupt the treatment again for 7 days, the total duration of interruptions will be 10 days. For subjects who complete the study treatment for a given study period but end on an interruption before continuing study treatment in the next study period,

the start date of the next study period will be used as the date restarted study treatment of the given study period. E.g., if a subject completed DB study treatment with an interruption on 23JUN2019 and then started the treatment withdrawal period on 26JUN2019, the duration of that treatment interruption will be (26JUN2019 – 23JUN2019) – 1 day = 2 days.

The percentage of days treated is defined as follows:

- Percentage of days treated = 100 × (treatment duration – total duration of treatment interruptions) / treatment duration.

Study treatment compliance (including categories 0%, > 0% – < 50%, 50% – < 70%, 70% – < 80%, 80% – < 90%, 90% – < 100%, 100%, > 100%), number of treatment interruptions per subject by category (0, 1, 2, 3, 4, > 4 interruptions), duration of treatment interruptions and percentage of days treated (including categories 0%, > 0% – < 50%, 50% – < 70%, 70% – < 80%, 80% – < 90%, 90% – < 100%, 100%) will be summarized for Run-in, DB and Run-out periods, separately. The summaries of the Run-out period will be based on the TWS.

The number (%) of subjects who have treatment interruptions, and the corresponding reasons will be tabulated.

Study treatment compliance, dispensing and accountability data will be listed.

6 EFFICACY VARIABLES AND ANALYSES

Unless noted otherwise, the analyses of efficacy endpoints will be performed using the FAS. Efficacy data described below will be listed using the FAS.

6.1 Analysis of the primary and secondary variables

6.1.1 Primary efficacy endpoints

The primary efficacy endpoints of this study are measured during polysomnography (PSG) nights and defined as:

- Change from baseline to Month 1 in WASO (sleep maintenance).
- Change from baseline to Month 3 in WASO.
- Change from baseline to Month 1 in LPS (sleep onset).
- Change from baseline to Month 3 in LPS.

LPS is the time (in minutes) from start of recording to the beginning of the first continuous 20 epochs (i.e., 10 min) scored as non-awake, i.e., epochs scored as either sleep stage 1 (S1), sleep stage 2 (S2), sleep stage 3 (slow wave sleep [SWS]), or rapid eye movement (REM), as determined by PSG.

WASO is the time (in minutes) spent awake after onset of persistent sleep until lights on, as determined by PSG.

Baseline is the mean of the two PSG nights at Visit 3 during the Run-in period. Month 1 and Month 3 are the mean of the two PSG nights at Visit 6 and Visit 8, respectively.

6.1.2 Secondary efficacy endpoints

The secondary efficacy endpoints of this study are measured on patient questionnaires and defined as:

- Change from baseline[a] to Month 1[b] in sTST.
- Change from baseline[a] to Month 3[c] in sTST.
- Change from baseline[a] to Month 1[b] in IDSIQ Sleepiness domain score.
- Change from baseline[a] to Month 3[c] in IDSIQ Sleepiness domain score.

sTST (in minutes) is based on the time reported by the subject in item 9 of the sleep diary i.e., the answer to the question "In total, how long did you sleep last night?".

The Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ) is a patient-reported outcome measure structured in 3 domains (Alertness/Cognition; Mood; Sleepiness). The IDSIQ Sleepiness domain score is based on the sum of the responses reported by the subject for 4 items: item 8, 11, 12, and 13. Item 8 score is reversed prior to summing [see Section 6.2]. The IDSIQ Sleepiness domain score can range from 0 to 40 (whole numbers only) with higher scores indicating greater burden during the daytime.

[a] Baseline is the mean value based on the screening sleep diary / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 3.

[b] 'Month 1' is the mean value based on the sleep diary / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 6.

[c] 'Month 3' is the mean value based on the sleep diary / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 8.

6.1.3 Overall testing strategy and hypotheses

The Type I error rate will be controlled for the testing of multiple null hypotheses associated with the primary and secondary endpoints assessed at 1 and 3 months of treatment and the two dose levels included in this study (i.e., 25 mg and 50 mg).

The eight statistical null hypotheses associated with the primary efficacy endpoints are:

Sleep maintenance:

- $H1_{WASO}$: Higher Dose – Placebo = 0 for WASO at Month 1.
- $H2_{WASO}$: Higher Dose – Placebo = 0 for WASO at Month 3.
- $H3_{WASO}$: Lower Dose – Placebo = 0 for WASO at Month 1.

- $H4_{WASO}$: Lower Dose – Placebo = 0 for WASO at Month 3.

Sleep onset:

- $H1_{LPS}$: Higher Dose – Placebo = 0 for LPS at Month 1.
- $H2_{LPS}$: Higher Dose – Placebo = 0 for LPS at Month 3.
- $H3_{LPS}$: Lower Dose – Placebo = 0 for LPS at Month 1.
- $H4_{LPS}$: Lower Dose – Placebo = 0 for LPS at Month 3.

The eight statistical null hypotheses associated with the secondary efficacy endpoints are:

Sleep quantity:

- $H1_{sTST}$: Higher Dose – Placebo = 0 for sTST at Month 1.
- $H2_{sTST}$: Higher Dose – Placebo = 0 for sTST at Month 3.
- $H3_{sTST}$: Lower Dose – Placebo = 0 for sTST at Month 1.
- $H4_{sTST}$: Lower Dose – Placebo = 0 for sTST at Month 3.

Next-day performance:

- $H1_{IDSIQ}$: Higher Dose – Placebo = 0 for IDSIQ Sleepiness domain score at Month 1.
- $H2_{IDSIQ}$: Higher Dose – Placebo = 0 for IDSIQ Sleepiness domain score at Month 3.
- $H3_{IDSIQ}$: Lower Dose – Placebo = 0 for IDSIQ Sleepiness domain score at Month 1.
- $H4_{IDSIQ}$: Lower Dose – Placebo = 0 for IDSIQ Sleepiness domain score at Month 3.

where 'Higher Dose', 'Lower Dose' and 'Placebo' represent the mean change from baseline for the given endpoint (WASO, LPS, sTST or IDSIQ Sleepiness domain score) and time point (Month 1 or Month 3) for the 50 mg, 25 mg and placebo treatment group, respectively.

Each null hypothesis will be tested against the alternative hypothesis that ACT-541468 improves the respective endpoint at the given dose and time point compared to placebo.

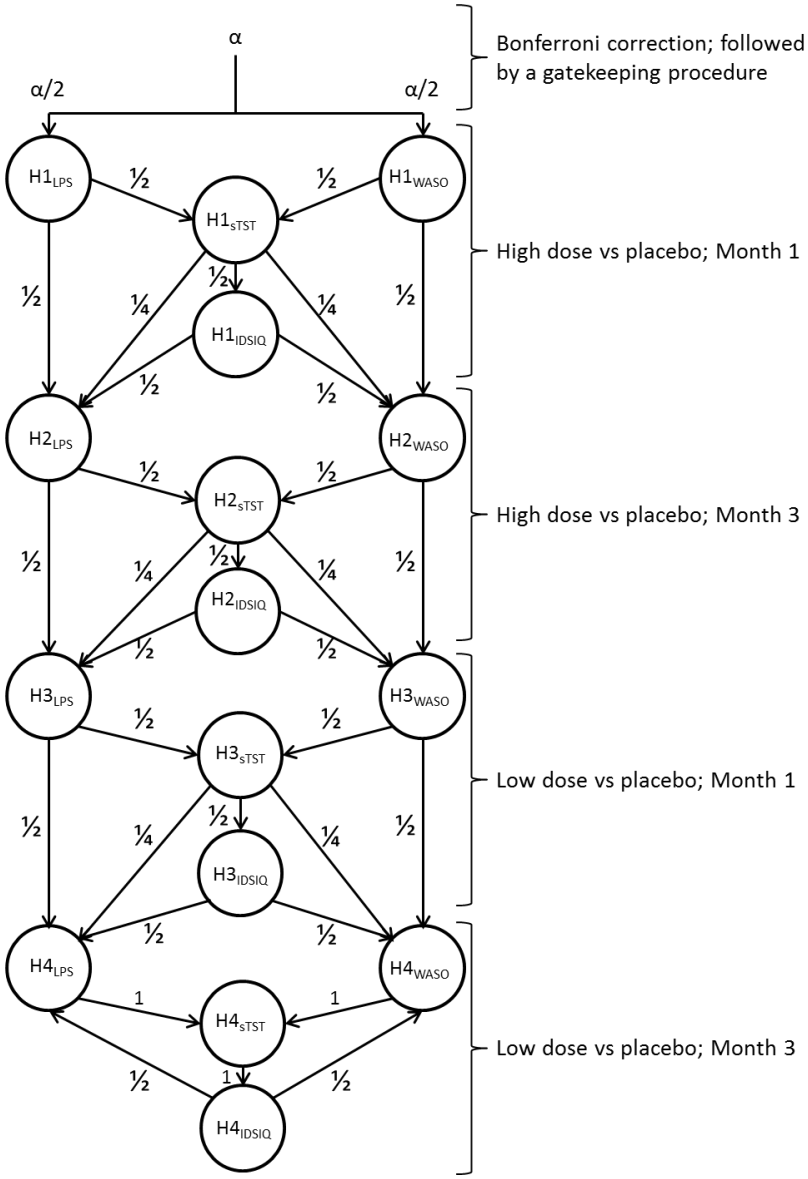
The order of testing and the alpha level applied to each null hypothesis will be based on the Bonferroni-based gatekeeping procedure [Bretz 2009] (shown in Figure 2 as a directed graph) that will control the study-wise Type I error at a two-sided 5% significance level.

To account for the concurrent evaluation (multiple comparison) of two distinct endpoint categories (i.e., sleep maintenance and sleep onset) a Bonferroni correction is applied. Both endpoint categories will be tested at half of the two-sided 5% significance level. This supports the sponsor's intention to make a superiority claim vs placebo for efficacy in either a sleep maintenance and/or sleep onset indication.

The remaining hypotheses will be tested following the gatekeeping strategy moving from Month 1 to Month 3 for higher dose ACT-541468 vs placebo and then from Month 1 to Month 3 for lower dose ACT-541468 vs placebo. The pre-specified proportion of alpha (weight) that will be distributed once a given null hypothesis (node) is rejected is shown

on the arrow in the directed graph. If a certain null hypothesis cannot be rejected, then the alpha level used for that test is absorbed at that node and not distributed further.

Figure 2 Hypothesis testing strategy



The graph is sequentially updated after rejection of a hypothesis following the algorithm below (algorithm 1 presented in Bretz 2009):

0. Set $I = M$.
1. Let j=arg $\min_{i \in I} p_i / \alpha_i$
2. $p_j \leq \alpha_j$, reject H_j ; otherwise stop.
3. Update the graph:

$$I \rightarrow I \setminus \{j\}$$

$$\alpha_l \rightarrow \begin{cases} \alpha_l + \alpha_j g_{jl}, & l \in I \\ 0 & otherwise \end{cases}$$

$$g_{lk} \rightarrow \begin{cases} \frac{g_{lk} + g_{lj} g_{jk}}{1 - g_{lj} g_{jl}}, & l, k \in I, \; l \neq k \\ 0 & otherwise \end{cases}$$

4. If $|I| \geq 1$, go to step 1; otherwise stop.

Where g are the weights, α the local significance, p the observed unadjusted two-sided p-values from the statistical model [see Section 6.1.4] for M null hypotheses and I denotes the associated index set (i.e., the given set of non-rejected null hypotheses to be tested).

The initial graph and the algorithm unequivocally define the multiple test strategy because the test decisions are independent of the rejection sequence [Bretz 2009]. In other words, the final result (i.e., the rejection or non-rejection of a given null hypothesis) will always be the same regardless of the order in which the hypotheses are tested.

6.1.4 Statistical model

All available data, regardless of the occurrence of any type of potential intercurrent event (e.g., study treatment discontinuation for any reason, the use of prohibited medication for any reason) will be included in the analyses. This approach (labeled in ICH E9[R1] as the treatment policy strategy) aligns, as close as possible, to the intention-to-treat principle, and is believed to provide an estimate of the treatment effect that is appropriate for regulatory and clinical decision-making.

The occurrence of intercurrent events in ID-078A301 is expected to be low. In the Phase 2 study (AC-078A201) of 1-month duration, no subjects (out of 356) took prohibited medication during treatment, and 3.9% (14 subjects) discontinued treatment prematurely. In 4 subjects, treatment discontinuation was due to AEs: 2 subjects on 10 mg ACT-541468, 1 subject on 50 mg ACT-541468, and 1 subject on 10 mg zolpidem. 7 subjects discontinued treatment for 'Other' reasons (withdrawal of consent, 2 subjects; subject randomized in error / randomization criteria not met, 4 subjects; change in subject's work schedule, 1 subject). In addition, 3 subjects discontinued DB study treatment with no

reason provided. No differences across treatment groups in the reasons or rates of premature treatment discontinuation was evident. However, it is acknowledged that the rate of intercurrent events is likely to be slightly higher in the Phase 3 studies due to the increased number of centers and the longer duration (3 months) for some of the endpoints.

It is expected that the type and occurrence of intercurrent events are similar across treatment groups. Therefore, the efficacy data collected following the occurrences of a intercurrent events are, on average, likely to be similar across treatment groups, diluting the treatment effect slightly. However, this seems reasonable since occurrence of intercurrent events are more likely due to unfavorable conditions, again providing an estimate of the treatment effect that is appropriate for regulatory and clinical decision-making.

Moreover, to ensure the estimand based on the treatment policy strategy can be estimated with minimal assumptions about missing data, efforts are being made to retain subjects in the trial and adhere to the schedule of assessments regardless of the occurrence of an intercurrent event.

A longitudinal data analysis method (i.e., linear mixed effects model) will be used for the analysis of change from baseline in WASO, LPS, sTST and IDSIQ Sleepiness domain score, separately.

The analysis model will adjust for the baseline value of the relevant response variable (either WASO, LPS, sTST or IDSIQ Sleepiness domain score), age group (< 65; ≥ 65 years), treatment (higher dose; lower dose; placebo), visit (Month 1; Month 3), and the interaction of treatment by visit, and baseline by visit [see SAS® code in Section 15.1].

An unstructured covariance matrix, shared across treatment groups, will be used to model the correlation among repeated measurements. A restricted maximum likelihood approach will be used to derive (unbiased) estimates of variance components. The Kenward-Roger approximation will be used to compute the denominator degrees of freedom and adjust standard errors [Kenward 1997].

To evaluate the efficacy hypotheses, appropriate contrasts will be used to test the treatment differences of interest (i.e., the difference in least squares [LS] mean change from baseline between higher dose ACT-541468 vs placebo and lower dose ACT-541468 vs placebo, both at Month 1 and Month 3 [see Section 15.1]).

The LS mean for each treatment group per time point will be displayed along with associated standard errors and 95% confidence intervals (CIs). For each ACT-541468 dose level comparison with placebo, the placebo-adjusted LS mean will be displayed along with associated standard error, 95% CI and unadjusted two-sided p-value.

6.1.5 Handling of partially missing data

Section 10.1 explains how the primary and secondary efficacy endpoints are derived in the event of partially missing data.

The (partially) missing data patterns of each primary and secondary efficacy endpoint will be shown with a shift table for each treatment group. The table will show the shifts in the number of available WASO, LPS, sTST or IDSIQ Sleepiness domain score values from baseline to Month 1, and from baseline to Month 3, respectively.

6.1.6 Descriptive analyses

Empirical cumulative distribution function plots of the observed cases of the primary and secondary efficacy endpoints for each treatment group will be provided. The primary and secondary efficacy endpoints along with their observed values will be summarized by age group (< 65; ≥ 65 years) and overall using descriptive statistics. A summary will be provided showing the cumulative number and percentage of subjects meeting a certain threshold for WASO, LPS, sTST and IDSIQ Sleepiness domain score, separately, for the observed range of values.

A plot of the mean change from baseline over time (per week, as defined in Table 7) for sTST and IDSIQ Sleepiness domain score will be provided together with a summary table.

6.1.7 Sensitivity analyses

The incidence and pattern of missing values will be explored to assess the appropriateness of the statistical analysis and possible impact on results. The observed missing data patterns for WASO, LPS, sTST and IDSIQ Sleepiness domain score at Baseline, Month 1 and Month 3 will be presented for each treatment group. The missing data patterns will be sorted from completely missing to completely observed by the order of the first occurrence of missing data.

The mean change from baseline of the primary and secondary efficacy endpoints will be plotted over time (to Month 1 and to Month 3) for study completers and for premature study withdrawals by treatment group and by reason for study discontinuation together with a summary table. A cumulative incidence plot will be provided showing the timing (relative to randomization date) and reasons for premature DB study treatment discontinuation and premature study discontinuation, separately, by treatment group. Based on the exploration of the observed missing data patterns, further sensitivity analyses to those described below may be performed.

The linear mixed effects model (or mixed model for repeated measures [MMRM]) relies on the missing at random (MAR) assumption. That is, conditional on the statistical model and observed values of the outcome (e.g., change from baseline WASO), the probability of missing data does not depend on the unobserved values of that outcome. Consequently,

the subject's missing values are estimated based on similar subjects who remain in the study.

Performing the main analysis under MAR is not likely to bias results in favor of ACT-541468 to an important degree, especially if the proportion of withdrawals are similar across treatment groups (or not significantly greater in the ACT-541468 groups), and subjects who withdraw early for a given reason tend to have similar outcomes before withdrawal (e.g., less favorable than completers) across treatment groups. In addition, assuming missing data are MAR in the placebo group is considered conservative; subjects who withdraw are expected to have worse efficacy over time compared to those placebo subjects that remain, whose data will constitute the basis for the imputed values for the withdrawn subjects. Furthermore, a placebo effect is expected to increase (improve) over time, particularly in those subjects on placebo who remain in the study.

Sensitivity analysis will be performed to assess the robustness of the conclusions of the main analysis to departures from the MAR assumption. For this, models that assume missing not at random (MNAR) will be used [Mallinckrodt 2013]. These models will be based on the multiple imputation (MI) methodology [Rubin 1987] and will be specifically used to assess the bias that can result when the outcomes for subjects who discontinue differ from those who complete. Sensitivity analyses will be performed for the analysis of change from baseline in WASO, LPS, sTST and IDSIQ Sleepiness domain score, separately.

As a general approach, imputations will be performed on observed values (baseline, Month 1 and Month 3) specifying a minimum and a maximum threshold to avoid imputation of nonsensical values. The boundaries will range from 0 to 480 for WASO, LPS and sTST, and from 0 to 40 for IDSIQ Sleepiness domain score. When an intended imputed value is outside the defined boundaries, this value will be discarded, and the procedure will draw another value for imputation. The maximum number of discarded values will be set to 100 for each imputed dataset. The imputed dataset will not be produced if more than 100 values have been discarded. If more than 20% of the planned imputed datasets are not produced, the results for that analysis will not be shown as this would likely lead to unreliable conclusions. Once all the multiply-imputed datasets are produced, the change from baseline to Month 1 and Month 3 will be calculated.

In all the MI analyses described below a preliminary step will be performed to obtain a monotone missing data pattern [see Section 12.7]. That is, any non-monotone (or intermediate) missing data will be imputed using the Markov Chain Monte Carlo (MCMC) methodology [Schafer 1997] under MAR assumptions within each treatment group. This assumption is considered reasonable given that subjects are expected to miss intermediate visits due to reasons unrelated to their insomnia disorder (e.g., scheduling conflicts), or have missing intermediate values due to technical problems (e.g., PSG not fully recorded due to power outage). In addition, the proportion of non-monotone missing data is expected

to be infrequent and imputing these few values using the MCMC method is not expected to compromise the validity of the sensitivity analyses. The MCMC imputation will be performed by treatment group and the variables included in the model are age (as continuous), baseline, Month 1 values and Month 3 values. The MCMC methodology assumes a multivariate normal distribution over these variables [see SAS® code in Section 15.2].

As a reference for the sensitivity analyses under MNAR (described below), the main analyses will be repeated using MI methodology under MAR. Imputation for monotone missing data will be performed sequentially, one visit at a time using the linear regression method [Ratitch 2013]. A linear regression model will be fitted sequentially for each visit (i.e., Month 1 and then Month 3) with treatment (higher dose; lower dose; placebo), age group (< 65; ≥ 65 years), baseline and any prior visit (i.e., Month 1 when fitting the linear regression model to Month 3 values) as predictors. Missing values are then replaced for each visit sequentially (i.e., Month 1 and then Month 3) by predictions from the respective linear regression model [Rubin 1987]. Note that values imputed at Month 1 are included when fitting the linear regression model to Month 3 values. The resulting multiple data sets are analyzed using the same statistical model as for the primary analysis (i.e., linear mixed effects model with the same factors). Results from these analyses are combined using Rubin's methodology [Rubin 1987; see SAS® code in Section 15.3]. The results obtained from this MI analysis are expected to be similar to the main MMRM analysis since both methods rely on the same MAR assumption.

Two sensitivity analyses assuming MNAR will be performed: A control-based imputation method called jump to reference (J2R) [Carpenter 2013], and a delta-adjusted imputation method [Yan 2009]. Each method differs in its assumption about the monotone missing data of the ACT-541468 groups. However, for both methods, subjects from the placebo group that discontinue the study are assumed to evolve in the same way as placebo subjects who remain in the study (i.e., MAR is assumed for the placebo group). As explained above, this assumption is considered conservative.

The J2R imputation method can be considered conservative (and more conservative than the methods above based on the MAR assumption) as it assumes that subjects discontinuing study from the ACT-541468 groups exhibit the same future evolution as placebo subjects. The same sequential procedure using linear regression models for imputing monotone missing data as described above will be used. However, predictions to replace the missing values in the ACT-541468 groups will be based on data from the placebo subjects only. As such, the factor for treatment is omitted from the sequential regression models. After imputation, the multiply-imputed datasets are then analyzed via standard MI analysis. Note that the same linear mixed effects model as in the main analysis will be used [see SAS® code in Section 15.4].

A delta-adjusted imputation method will be used to progressively stress test the MAR assumption of the main analysis. That is, a tipping-point analysis will be performed to see how severe departures from the MAR assumption must be to overturn the conclusions of the main analysis (i.e., the point where significant results become non-significant). The assumption here is that subjects from the ACT-541468 groups who discontinue from study would have, on average, their observed efficacy score worsened by some amount (delta) compared with the efficacy observed from subjects of the same ACT-541468 group that remain in the study. A series of analyses with progressively increasing deltas ($\delta_{WASO/LPS}$ = 15, 30, 45, 60; δ_{sTST} = –20, –40, –60, –80; δ_{IDSIQ} = 2.5, 5, 7.5, 10; from less conservative to more conservative) will be performed. The delta adjustment will only be made to the multiply-imputed values (based on the MAR assumption) for the subjects in the ACT-541468 groups. Note that a delta of zero corresponds to the standard MAR-based MI analysis that is described above to obtain a reference for the sensitivity analyses. Again, the same linear mixed effects model as in the main analysis will be used [see SAS® code in Section 15.5].

For all the analyses above using MI methodology, the seed is set to 708695 and 100 multiply-imputed datasets are created. The decision of imputing 100 datasets was taken from O'Kelly [O'Kelly 2014], where it is stated that number of imputations should typically be between 20 and 100 (simulations performed by Graham [Graham 2007]). Similar recommendations were provided by others [Bodner 2008, Von Hippel 2009, White 2011].

For the main analysis and each planned sensitivity analysis, the LS mean estimate of the difference in change from baseline between higher dose ACT-541468 vs placebo and lower dose ACT-541468 vs placebo both at Month 1 and Month 3 will be provided, together with its estimated standard error and associated p-value. The LS mean estimate of the change from baseline for each treatment group at Month 1 and Month 3 will also be provided, together with its estimated standard error and associated p-value.

A plot over time for the primary and secondary efficacy endpoints showing the LS mean of each treatment group per imputation method (i.e., MAR; J2R; delta-adjusted including $\delta = 0$) will be provided.

6.1.8 Supportive analyses

The main analysis performed for the primary and secondary efficacy endpoints [Section 6.1.4] will be repeated using the PPS to assess the effect protocol deviations may have on the results. Of note, these analyses rely on the MAR assumption.

The impact of any deviation from the assumption that the covariance structure is similar across treatment groups will be investigated by re-running the main analyses using two additional types of covariance structures. Firstly, the assumption of a common within-group standard deviation (SD) will be relaxed by using an unstructured covariance

matrix that is specific to each treatment group. Secondly, the more restrictive Toeplitz structure, shared across treatment groups, will be used. The Toeplitz structure requires fewer parameters to be estimated and will be considered as the main statistical analysis model if the model described in Section 6.1.4 fails to converge.

6.1.9 Subgroup analyses

Subgroup analyses for the primary and secondary efficacy endpoints will be performed to investigate the consistency of the treatment effect (i.e., ACT-541468 vs placebo) across subgroups defined by:

Age:	< 65, ≥ 65 years
Sex:	Male, female
Region:	US, other (non-US)

The same model as for the main analysis (MMRM) will be fitted per subgroup. Of note, age group will not be included as a covariate in the model for subgroup analysis by age. Treatment effect estimates (LS mean plus 95% CI) will be presented via forest plots.

6.2 Analysis of other efficacy variables

The following efficacy endpoints along with their observed values will be summarized by age group (< 65; ≥ 65 years) and overall using descriptive statistics:

- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in Total Sleep Time (TST).
- Change from baseline[a] to Month 1[b] and Month 3[c] in subjective WASO (sWASO).
- Change from baseline[a] to Month 1[b] and Month 3[c] in subjective Latency to Sleep Onset (sLSO).
- Change from baseline[a] to Month 1[b] and Month 3[c] in IDSIQ scores (i.e., Total score; Alert/Cognition and Mood domain scores).

[a] Baseline is the mean value based on the screening sleep diary / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 3.

[b] 'Month 1' is the mean value based on the sleep diary / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 6.

[c] 'Month 3' is the mean value based on the sleep diary / IDSIQ entries performed at home in the 7 days immediately preceding the first PSG at Visit 8.

sWASO is the self-reported time spent awake after sleep onset as reported in item 7 of the sleep diary, i.e., the answer to the question "In total, how long did these awakenings last?".

sLSO (in minutes) is the self-reported time to fall asleep as reported in item 5 of the sleep diary, i.e., the answer to the question "How long did it take you to fall asleep?".

Section 10.1 explains how these endpoints are derived in the event of partially missing data.

The IDSIQ has 3 domains assessing Alert/Cognition, Mood, and Sleepiness [Table 2]. A Total score can also be obtained. For each domain, the item responses in the respective domains are summed (each item is based on a categorical scale ranging from 0 to 10). Prior to summation, items 1, 2, 8, 10 and 14 must be reversed [see Table 2 for details regarding scoring]. For the Total score, all item responses are summed.

Table 2 Insomnia Daytime Symptoms and Impacts Questionnaire description

Domain	Scoring	Minimum/Maximum Score
Alert/Cognition	Daily: Sum of Item 1*, Item 2*, Item 3, Item 9, Item 10*, Item 14*	Minimum score: 0 Maximum score: 60 Higher score: greater burden of illness
	Weekly Average: Mean of the daily domain score over 7 days	
Mood	Daily: Sum of Item 4, Item 5, Item 6, Item 7	Minimum score: 0 Maximum score: 40 Higher score: greater burden of illness
	Weekly Average: Mean of the daily domain score over 7 days	
Sleepiness	Daily: Sum of Item 8*, Item 11, Item 12, Item 13	Minimum score: 0 Maximum score: 40 Higher score: greater burden of illness
	Weekly Average: Mean of the daily domain score over 7 days	
Total score	Daily: Sum of all domains above	Minimum score: 0 Maximum score: 140 Higher score: greater burden of illness
	Weekly Average: Mean of the daily domain score over 7 days	

* Item 1, Item 2, Item 8, Item 10, and Item 14 scores are reverse scored prior to summation.

A plot of the mean change from baseline over time (per week, as defined in Table 7) for sWASO, sLSO and IDSIQ scores (i.e., Total score; Alert/Cognition and Mood domain scores) will be provided together with a summary table.

The same model as for the main analysis, MMRM [Section 6.1.4], will be fitted for these efficacy endpoints. The LS mean for each treatment group will be displayed along with associated standard errors and 95% CIs. For each ACT-541468 dose level comparison with placebo, the placebo-adjusted LS mean will be displayed along with associated standard error, 95% CI and unadjusted two-sided p-value.

6.3 Analysis of exploratory variables

The following exploratory endpoints will be summarized along with their observed values at the respective time points:

- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in WASO over time (by hour of the night and by quarter of the night).
- Change from baseline[a] (Visit 3) to Month 1[b] and Month 3[c] in the following visual analog scale (VAS) scores collected in the sleep diary: quality of the sleep, depth of the sleep, daytime alertness and ability to function.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in duration of TST in each sleep stage (S1, S2, SWS and REM) over the whole night, and for each quarter of the night.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in sleep architecture assessed as percentage of TST in each sleep stage (S1, S2, SWS, and REM) over the whole night, and for each quarter of the night.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in sleep onset latency.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in numbers of shifts from S2, SWS or REM to S1 or awake.
- Change from baseline (Visit 3) to Month 1 and Month 3 in Insomnia Severity Index© (ISI©) scores.
  - The number (%) of subjects with ≥ 6 points decrease in ISI© Total score from baseline (Visit 3) to Month 1 (Visit 6) and Month 3 (Visit 8) will be tabulated.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in number of awakenings (defined as the number of awakenings between first epoch and last epoch not scored wake) as measured by PSG (will be summarized for the whole night, for each quarter of the night, and for each hour of the night).
- Change from baseline[a] (Visit 3) to Month 1[b] and Month 3[c] in number of self-reported awakenings.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in Sleep Efficiency (defined as 100 × [TST (min) / time in bed (min)]) where time in bed is fixed to 480 min during the PSG nights.
- Change from baseline (Visit 3) to Month 1 (Visit 6 or, if that is missing, week 4 of the questionnaire) and Month 3 (Visit 8 or, if that is missing, week 12 of the

questionnaire) in Patient Global Assessment of Disease Severity (PGA-S) scores (daytime symptoms).

- Change from baseline (Visit 3) to Month 1 (Visit 6 or, if that is missing, week 4 of the questionnaire) and Month 3 (Visit 8 or, if that is missing, week 12 of the questionnaire) in Patient Global Impression of Change (PGI-C) scores (daytime symptoms).
- Change from baseline (Visit 3) to Month 1 (Visit 6 or, if that is missing, week 4 of the questionnaire) and Month 3 (Visit 8 or, if that is missing, week 12 of the questionnaire) in PGI-C scores (night-time symptoms).
- Change from baseline (Visit 3) to Month 1 (Visit 6 or, if that is missing, week 4 of the questionnaire) and Month 3 (Visit 8 or, if that is missing, week 12 of the questionnaire) in Patient Global Impression of Severity (PGI-S) scores (night-time symptoms).

[a] Baseline is the mean value based on the screening sleep diary entries performed at home in the 7 days immediately preceding the first PSG at Visit 3.

[b] 'Month 1' is the mean value based on the sleep diary entries performed at home in the 7 days immediately preceding the first PSG at Visit 6.

[c] 'Month 3' is the mean value based on the sleep diary entries performed at home in the 7 days immediately preceding the first PSG at Visit 8.

Section 10.1 explains how these endpoints, except for ISI©, are derived in the event of partially missing data.

Additionally, the following PSG measurements will be summarized:

- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in duration (min) of TST in each sleep stage (S1, S2, SWS and REM) by quarter of the night.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in duration (min) from lights off to the first epoch of each sleep stage (S1, S2, SWS and REM).
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in duration of time (min) from onset of persistent sleep (LPS) to the first epoch of REM sleep.

- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in duration of time (min) from sleep onset (LSO) to the first epoch of REM sleep.
- Change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) in duration of wakefulness (min) from onset of persistent sleep (LPS) till the last epoch of sleep prior lights on.

Boxplots of the change from baseline (mean of the 2 PSG nights at Visit 3) to Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) will be provided for the following:

- Duration of TST (min) in each sleep stage over the whole night, and by quarter of the night.
- % of TST spent in each sleep stage over the whole night, and by quarter of the night.
- Latency (min) to each sleep stage over the whole night.
- Duration of REM sleep (min) by quarter of the night.
- Duration of S1 sleep (min) by quarter of the night.
- Duration of the S2 sleep (min) by quarter of the night.
- Duration of SWS sleep (min) by quarter of the night.
- WASO (min) by quarter of the night.

Boxplots will also be provided for the observed values at Baseline (mean of the 2 PSG nights at Visit 3), Month 1 (mean of the 2 PSG nights at Visit 6) and Month 3 (mean of the 2 PSG nights at Visit 8) for the % of TST spent in each sleep stage over the whole night.

6.4 Estimating minimal clinically important differences

This analysis will be described in a separate document and reported separately.

6.5 Patient preferences exploratory endpoints

This analysis will be described in a separate document and reported separately.

7 SAFETY VARIABLES AND ANALYSES

7.1 Overview of safety analyses including subgroup analyses

Unless noted otherwise, the SS will be used for tables and listings of safety data.

Safety data described below will be listed.

In the event of permanent study treatment discontinuation, it is recommended to perform the safety assessments of Visit 8 within 7 days of the last DB study treatment intake. Although some subjects may be considered as being off study treatment, the time from the

last DB study treatment intake to the Visit 8 assessments is expected to be short. Therefore, for subjects who prematurely discontinue treatment, the safety data (laboratory tests, ECG, vital signs) measured within 7 days of the last DB study treatment intake will be included as being taken during DB study treatment. A subject who prematurely discontinues treatment may continue the planned study visits until EOS, without the run-out period. Therefore, for subjects who prematurely discontinue treatment, the safety data (laboratory tests, ECG, vital signs) measured more than 7 days from the last DB study treatment intake will be excluded from analyses.

7.2 Adverse events

Unless noted otherwise, the AE summary tables will include treatment-emergent AEs (TEAEs): AEs that started or worsened on or after DB study treatment start date up to the earlier of 30 days after DB study treatment end date or the date of enrollment into the ID-078A303 extension study. An AE that started or worsened during any planned or unplanned study treatment interruption (e.g., run-out, safety follow-up) that is falling in the treatment-emergent time window defined above, will be considered as treatment-emergent.

AEs will be coded using MedDRA. The MedDRA version used for reporting will be specified in the footnote of the applicable output.

The number (%) of subjects experiencing a TEAE (including serious AEs [SAEs], AEs of special interest [AESIs] after adjudication by the Independent Safety Board [ISB], and AEs leading to premature discontinuation or temporary interruption of DB study treatment) will be summarized by SOC and/or PT, and/or maximum intensity.

Each AESI category (i.e., excessive daytime sleepiness, cataplexy, complex sleep behavior including hallucinations / sleep paralysis, suicide/self-injury, narcolepsy-like symptoms) will be summarized by PT separately.

A subject with multiple intensities reported for an AE will be summarized under the maximum intensity recorded for the event.

Apart from the summaries of occurrences, where each event is counted, a subject with multiple occurrences of an AE is counted only once in the AE category (e.g., SOC, PT). If a single AE worsens on the same treatment group, then this event will be considered as one occurrence. AEs will be sorted by descending frequency, first in the highest ACT-541468 dose level, then in the next highest ACT-541468 dose level, and finally in the placebo treatment group. After this sorting, SOCs are presented in alphabetical order with PT sorted within SOC in alphabetical order.

The following AE summary tables will be provided:

- Treatment-emergent AEs.
- Treatment-emergent AEs during the DB study period[a].

- AEs occurring (i.e., that started or worsened) during the Run-in period.
- AEs occurring (i.e., that started or worsened) during the treatment withdrawal period (based on the TWS).
- TEAEs during the DB study period[a] related to study treatment.
- AEs leading to premature discontinuation of DB study treatment.
- AEs leading to temporary interruption of DB study treatment.
- Treatment-emergent SAEs (including occurrences).
- Treatment-emergent SAEs related to study treatment (including occurrences).
- Disclosure to public database purposes: most frequent (threshold to be defined at time of reporting) non-serious TEAEs (including occurrences).
- Treatment-emergent AESIs after ISB adjudication.
- TEAEs during the DB study period[a] related to abuse (as defined in the document *Search specifications for safety monitoring)*.
- TEAEs during the DB study period[a] based on selected Standardised MedDRA Queries or pre-defined search criteria (as defined in the document *Search specifications for safety monitoring)*.
- TEAEs with fatal outcome.
- TEAEs related to study treatment with fatal outcome.
- Total number of deaths (those occurring on or after DB study treatment start date up to the earlier of 30 days after DB study treatment end date or the date of enrollment into the ID-078A303 extension study).

[a] Includes only those TEAEs occurring (i.e., that started or worsened) during the DB study period [see Section 12.2].

All AEs will be listed as well as all deaths with cause of death. Any death will be listed using the SCR set. In addition, separate listings will be provided for SAEs.

An AE listing will be provided also for subjects who were not randomized (i.e., screen failures) and subjects who were randomized but did not take DB study treatment to report any subjects who discontinued run-in single-blind placebo treatment or discontinued the study, respectively, due to AE.

AEs submitted for adjudication will be listed flagging, separately, AEs that were adjudicated as an AESI and AEs that were considered an AESI by the investigator.

The number of subjects (including occurrences) with at least one case submitted for ISB adjudication will be summarized and listed.

An incomplete (day or month missing) or missing AE date will be imputed as described in the Table 3. The 'lower limit' and 'upper limit' refer to the earliest and latest possible dates, respectively. As an example: If AE onset date is MAR2017 (day missing), the lower limit is 01MAR2017 and the upper limit is 31MAR2017; If AE onset date is 2017 (day and month missing), the lower limit is 01JAN2017 and the upper limit is 31DEC2017.

Table 3 Imputation rules for an incomplete or missing AE date

Field	Incomplete date	Missing date
AE resolution date	The upper limit.	No imputation; the AE is considered as ongoing.
AE onset date	The rules below apply in the order presented: 1. If the (imputed) AE end date is on or after the start of DB study treatment, and if the DB study treatment start falls within the upper and lower limits (inclusive), the DB study treatment start date is used. 2. If the (imputed) AE end date is on or after the start of run-in single-blind placebo treatment, and if the run-in single-blind placebo treatment start falls within the upper and lower limits (inclusive), the run-in single-blind placebo treatment start date is used. 3. If the AE resolution date is missing, and: a. if the DB study treatment start falls within the upper and lower limits (inclusive), the DB study treatment start date is used; else: b. if the run-in single-blind placebo treatment start falls within the upper and lower limits (inclusive), the run-in single-blind placebo treatment start date is used. 4. In all the other cases, the lower limit is used.	Whichever is the earlier of the AE resolution date or DB study treatment start date.

AE = adverse event; DB = double-blind.

The purpose of imputing AE dates is only to assign an AE to a specific treatment phase for the summary tables. No imputed date is considered in the medical evaluation of an AE.

7.3 Laboratory tests

Laboratory analyses are based on data received from the central laboratory. Laboratory data will be converted into SI units. Unless noted otherwise, summaries and listings will include scheduled, re-test and unscheduled assessments.

Descriptive summary statistics for each scheduled visit (re-test excluded) will be provided for observed values and change from baseline in both hematology and blood chemistry laboratory parameters.

The change from baseline to last value (re-test and unscheduled assessments excluded) in the DB study period for hematology and blood chemistry parameters will be summarized.

Marked laboratory abnormalities are defined in Table 4 below. The number (%) of subjects with marked laboratory abnormalities during DB study treatment will be tabulated. A subject will be counted only once, but may be reported in more than one marked laboratory abnormality criterion of a given parameter. Percentages will be based on the number of subjects at risk, i.e., those having at least one post-baseline value for a given parameter.

Shift from baseline to worst value post-baseline based on laboratory normal ranges will be provided for the following laboratory parameters:

- Liver enzymes: alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase and total bilirubin (TBIL).
- Thyroid hormones: Thyroid-stimulating hormone, free T_3 and free T_4.
- Lipids: total cholesterol and triglycerides.
- Hematology: leukocytes, neutrophils, lymphocytes, eosinophils.

All laboratory data for subjects with at least one marked laboratory abnormality during DB study treatment will be listed. Any local laboratory data collected will be listed separately.

Table 4 Marked abnormalities in laboratory parameters for reporting

Laboratory parameter	Criteria for marked laboratory abnormalities
Hematology (SI unit)	
Hemoglobin (g/L)	< 100 < 80 > 20 above ULN or > 20 above baseline if baseline > ULN > 40 above ULN or > 40 above baseline if baseline > ULN
Hematocrit (%)	< 32% (Male); < 28% (Female) < 20% > 60% (Male); > 55% (Female) > 65%
Platelets (10^9/L)	< 75 < 50 > 600 > 999
Leukocytes (10^9/L)	< 3.0 < 2.0

Laboratory parameter	**Criteria for marked laboratory abnormalities**
	> 20.0 > 100.0
Neutrophils (10^9/L)	< 1.5 < 1.0
Eosinophils (10^9/L)	> 5.0
Eosinophils (%)	> 5%
Lymphocytes (10^9/L)	< 0.8 < 0.5 > 4.0 > 20.0
Reticulocyte (%)	> 2.5%
Blood chemistry (SI unit)	
ALT (U/L)	> 3 × ULN > 5 × ULN > 10 × ULN
AST (U/L)	> 3 × ULN > 5 × ULN > 10 × ULN
Alkaline phosphatase (U/L)	> 2.5 × ULN > 5 × ULN
Total bilirubin (µmol/L)	> 2 × ULN > 5 × ULN
Creatinine (µmol/L)	> 1.5 × ULN or > 1.5 × baseline if baseline > ULN > 3 × ULN or > 3 × baseline if baseline > ULN
Creatinine clearance (ml/min/ 1.73 m^2)	< 30 < 60
Albumin (g/L)	< 30 < 20
Calcium (mmol/L)	< 2.0 < 1.75 > 2.9 > 3.1
Potassium (mmol/L)	< 3.2 < 3.0 > 5.5 > 6.0
Sodium (mmol/L)	< 130 > 150

Laboratory parameter	Criteria for marked laboratory abnormalities
	> 155
Chloride (mmol/L)	< 74 > 131
Creatine kinase (µg/L)	> 5 × ULN > 10 × ULN
Gamma-glutamyl transferase (U/L)	> 2.5 × ULN > 5 × ULN
Blood urea nitrogen (mmol/L)	> 2.5 × ULN > 5 × ULN
Uric acid (µmol/L)	> 590 > 720
Glucose (mmol/L)	< 3 < 2.2 > 8.9 > 13.9
TSH (mIU/L)	< 0.28 > 6.6
Free T_3 (pmol/L)	< 2.8 > 7.8
Total T_3 (nmol/L)	< 0.9 > 3.6
Free T_4 (pmol/L)	< 7.2 > 24
Total T_4 (nmol/L)	< 53 > 210
Triglycerides	> 5.7 mmol/L
Total cholesterol	> 2 × ULN

ALT = alanine aminotransferase; AST = aspartate aminotransferase; SI = standard international; TSH = thyroid-stimulating hormone; T_3 = triiodothyronine; T_4 = thyroxine; ULN = upper limit of normal.

Elevated liver parameters during DB study treatment will be summarized: the number (%) of subjects meeting the criteria defined below within a given central laboratory sample will be tabulated by treatment group. Percentages will be based on the number of subjects at risk, i.e., those having at least one post-baseline value for all parameters within the given criteria.

- AST or ALT > 3 × upper limit of normal (ULN)
- AST or ALT > 5 × ULN
- AST or ALT > 10 × ULN

- TBIL > 1.5 × ULN and > 2 × ULN
- ALP > 1.5 × ULN
- (ALT or AST > 3 × ULN) and (TBIL > 1.5 × ULN)
- (ALT or AST > 3 × ULN) and (TBIL > 2 × ULN)

7.4 Electrocardiography

Unless noted otherwise, ECG summaries and listings will include scheduled, re-test and unscheduled assessments.

Descriptive summary statistics for each scheduled visit (re-test excluded) will be provided for observed values and change from baseline in ECG parameters (QT corrected according to Fridericia's formula, QT corrected according to Bazett's formula, heart rate, PR, QRS).

Marked ECG abnormalities are defined in Table 5. The following summaries will be provided for each ECG parameter:

- Number (%) of subjects with a marked ECG abnormality during the DB study period.
- Number (%) of subjects with a marked ECG abnormality during the treatment withdrawal period (based on the TWS).

Percentages will be based on the number of subjects at risk, i.e., those having at least one post-baseline value in the study period under consideration per ECG parameter for criteria based on post-baseline values only, or those having a baseline value and at least one post-baseline value in the study period under consideration per ECG parameter for criteria based on change from baseline. For non-mutually exclusive criteria (i.e., HR, PR, and QRS), a subject will be counted only once in case of multiple occurrences but may be reported in more than one marked ECG abnormality criterion of a given parameter. For mutually exclusive criteria (i.e., QTc), only the subject's worst post-baseline value (or worst change from baseline value) will be counted.

ECG findings during the DB study period and during the treatment withdrawal period (based on the TWS) will be summarized separately and listed. ECG abnormality categories and findings will be sorted by descending frequency, first in the highest ACT-541468 dose level, then in the next highest ACT-541468 dose level, and finally in the placebo treatment group. After this sorting, ECG abnormality categories will be presented in alphabetical order with ECG finding sorted within ECG abnormality category in alphabetical order. All ECG values for subjects with at least one marked ECG abnormality during the DB study period or during the treatment withdrawal period will be listed.

Table 5 Marked abnormalities in ECG parameters

ECG parameter	Criteria for marked ECG abnormalities
QTcF, QTcB (ms)	> 450 and ≤ 480 > 480 and ≤ 500 > 500 > 30 and ≤ 60 increase from baseline > 60 increase from baseline
HR (bpm)	< 45 < 50 > 10 and ≤ 20 decrease from baseline > 20 decrease from baseline
PR (ms)	> 200
QRS (ms)	> 110

bpm = beats per minute; ECG = electrocardiogram; HR = heart rate; QTcB = QT interval corrected according to Bazett's formula; QTcF = QT interval corrected according to Fridericia's formula.

7.5 Vital signs and body weight

Each summary will include only scheduled assessments (re-test excluded) and listings will include both scheduled and unscheduled assessments.

The change from baseline (mean of the two PSG nights at Visit 3, run-in period) to Month 1 (mean of the two PSG nights at Visit 6), Month 2 (Visit 7), Month 3 (mean of the two PSG nights at Visit 8) and Run-out period (Visit 9 and Visit 10) in vital signs (systolic and diastolic blood pressure, and pulse rate) will be summarized. The observed values at baseline and each scheduled post-baseline visit will also be summarized.

The change from baseline (Visit 1 or last value before DB study treatment) to Visit 8 (or last value on study) in body weight will be summarized.

7.6 Withdrawal symptoms

The TWS will be used to assess the potential for withdrawal symptoms.

The BWSQ assesses the main symptoms which might be experienced by subjects during withdrawal from benzodiazepines.

The questionnaire consists of 20 items with each item rated by the subject as either 0 (No), 1 (Yes-moderate) to 2 (Yes-severe). The BWSQ Total score (possible range: 0 to 40) for the change from the last available assessment on DB study treatment to the beginning and the end of the treatment withdrawal period (in the morning at Visit 9 and at Visit 10, respectively) will be summarized. Observed values will also be summarized.

The number (%) of subjects with a BWSQ Total score above 20 will be tabulated by visit.

The number (%) of subjects with one or more BWSQ symptom scored as 'severe' will be tabulated by visit.

In addition, withdrawal symptoms after DB study treatment discontinuation will be assessed through the incidence of AEs, and marked ECG abnormalities, occurring during the treatment withdrawal period [see Section 12.2]. As described in their respective section, the incidence of AEs [Section 7.2] and marked ECG abnormalities [Section 7.4] occurring during the treatment withdrawal period (between Visit 9 and Visit 10) will be summarized.

7.7 Rebound insomnia

The TWS will be used to assess the potential for rebound insomnia. The change from baseline (Visit 3) to the treatment withdrawal period (Visit 9, run-out) in objective sleep parameters (WASO and LPS) will be summarized. The change from baseline (mean value based on the screening sleep diary entries performed at home in the 7 days immediately preceding the first PSG at Visit 3) to the treatment withdrawal period (after PSG night at Visit 9 [see Section 12.8]) in the subjective sleep parameter sTST will be summarized. Empirical cumulative distribution function plots of these objective and subjective sleep parameters for each treatment group will be provided, together with a summary showing the cumulative number and percentage of subjects meeting a certain threshold for the observed range of values.

7.8 Next-day residual effect

Coding sub-test©

The Coding sub-test© is a measure of attention, perceptual speed, motor speed, visual scanning and memory. The total Coding sub-test© score is the number of correct symbols entered.

The change from baseline (mean of the two PSG morning assessments at Visit 3, run-in period) to Month 1 and Month 3 (mean of the two PSG morning assessments at Visit 6 and Visit 8, respectively), and Run-out period (Visit 9) in total Coding sub-test© score will be summarized. Observed values will also be summarized.

Sheehan Disability Scale©

The Sheehan Disability Scale© (SDS©) consists of three questions on impairment of work, social life, and family life / home responsibilities each on a 10-point scale. The three items will be summed into a single dimensional measure of global functional impairment that ranges from 0 (unimpaired) to 30 (highly impaired). Number of days lost (school, work or normal daily responsibilities) and number of days underproductive (school, work or normal daily responsibilities) are also collected.

The change from baseline (mean of the two PSG morning assessments at Visit 3, run-in period) to Month 1 and Month 3 (mean of the two PSG morning assessments at Visit 6 and

Visit 8, respectively), and Run-out period (Visit 9) in each SDS© item (impairment of work, social life, family life / home responsibilities, number of days lost and number of days underproductive) and Total score will be summarized. The Total score will be calculated only for subjects with all the three items available: impairment of work, social life and family life / home responsibilities. Observed values will also be summarized.

Visual analog scales

The morning VAS consists of three questions: quality of subjects' sleep, depth of sleep and sleepiness in the morning. The evening VAS consists of two questions: daytime alertness and ability to function. The VAS scores are collected in the sleep diary and answered on a continuous bipolar (or unipolar for sleepiness in the morning) scale ranging from 0 to 100 points; for all questions, a higher score reflects a better outcome.

The change from baseline[a] (Visit 3) to Month 1[b] and Month 3[c] in morning sleepiness VAS score will be summarized. Observed values will also be summarized.

[a] Baseline is the mean value based on the screening sleep diary entries performed at home in the 7 days immediately preceding the first PSG at Visit 3.

[b] 'Month 1' is the mean value based on the sleep diary entries performed at home in the 7 days immediately preceding the first PSG at Visit 6.

[c] 'Month 3' is the mean value based on the sleep diary entries performed at home in the 7 days immediately preceding the first PSG at Visit 8.

Neurological examination

The time from treatment until the subject is considered safe to leave the center the next day will be summarized for each scheduled PSG visit.

7.9 Columbia-Suicide Severity Rating Scale©

The Columbia-Suicide Severity Rating Scale© (C-SSRS©) is an instrument that evaluates suicidal ideation and behaviors.

The C-SSRS© outcome categories are provided below. Each category has a binary response (yes/no) and are numbered and ordered below for convenience.

1 – Wish to be Dead
2 – Non-specific Active Suicidal Thoughts
3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
5 – Active Suicidal Ideation with Specific Plan and Intent
6 – Preparatory Acts or Behavior
7 – Aborted Attempt
8 – Interrupted Attempt

9 – Actual Attempt (non-fatal)
10 – Completed Suicide

Self-injurious behavior without suicidal intent is also a C-SSRS© outcome (although not suicide-related) and has a binary response (yes/no).

Categories 1–5 relate to suicidal ideation and a score of 0 is assigned if no suicidal ideation is present. Categories 6–10 relate to suicidal behavior.

Based on the C-SSRS©, the number (%) of subjects with suicidal ideation by category, suicidal behavior by category, suicidal ideation or suicidal behavior, and/or self-injurious behavior without suicidal intent 1) during DB study treatment and 2) during the treatment withdrawal period (based on the TWS) will be tabulated separately. Percentages will be based on the number of subjects with at least one post-baseline C-SSRS© assessment, in the corresponding period. The assessment will be assigned to a study period based on a scheduled visit as follows: Visit 1 = not assigned to any study period; Visit 2 = Screening period; Visit 3 and Visit 4 = Run-in period; Visit 6, Visit 7 and Visit 8 = DB study period; Visit 9 and Visit 10 = Treatment withdrawal (or Run-out) period.

Shift from baseline showing any change in suicidal ideation and suicidal behavior 1) during the DB study period and 2) during the treatment withdrawal period (based on the TWS) will be provided separately. Subjects will be summarized under the worst of the following three categories, shown here in the order from best to worst: 1) No suicidal ideation or behavior, 2) Suicidal ideation, and 3) Suicidal behavior (subjects with both suicidal ideation and suicidal behavior are also included in the suicidal behavior category). Suicidal ideation includes any one of the five suicidal ideation events (categories 1–5). Suicidal behavior includes any one of the five suicidal behavior events (categories 6–10).

7.10 Epworth Sleepiness Scale©

The Epworth Sleepiness Scale© (ESS©) is a validated questionnaire designed to provide a subjective measure of daytime sleepiness. The ESS© consists of eight situations on chances of dozing; each situation is evaluated with a score that ranges from 0 (never dozing) to 3 (high chance of dozing). The eight items will be summed into a Total score of daytime sleepiness.

The change from baseline (Visit 3) to Month 1 (Visit 6), Month 3 (Visit 8) and end of run-out period (Visit 10) in ESS© Total score will be summarized. Observed values will also be summarized.

The number (%) of subjects meeting the criteria defined below will be tabulated:

- ESS© Total score > 16 at Baseline (Visit 3), Month 1 (Visit 6), Month 3 (Visit 8) and end of Run-out period (Visit 10).

- ESS© Total score > 4 points increase from previous visit at Month 1 (Visit 6), Month 3 (Visit 8) and end of Run-out period (Visit 10).

7.11 Subgroup analysis of safety variables

Summaries of TEAEs occurring (i.e., that started or worsened) during the DB study period [see Section 12.2] by SOC and PT will be performed on the following subgroups:

Age: $< 65, \geq 65$ years

BMI: < 25, 25–30, > 30 kg/m^2

8 PHARMACOKINETIC VARIABLES AND ANALYSES

PK analyses will be performed using the PK set.

Descriptive statistics (n, mean, SD, coefficient of variation [CV%], m [number of non-zero concentrations], geometric mean, geometric SD, geometric CV%, median, minimum and maximum) of the ACT-541468 plasma concentrations collected approximately 9–10 h post-dose in the morning after the second PSG night at Visit 6 and Visit 8 by age categories will be provided. These concentration values will also be displayed graphically by visit and ACT-541468 treatment group: scatter plots vs BMI (with a linear regression line) by sex; boxplots for age categories ($< 65, \geq 65$ years) and all ages combined; empirical cumulative distribution functions, together with a summary showing the cumulative number and percentage of subjects meeting a certain threshold for the observed range of values; and probability density functions.

Concentration values below the lower limit of quantification will be displayed in listings as "BLQ" and handled as zero in the calculations for mean, CV%, SD, median, minimum, and maximum, but handled as missing for the calculation of the geometric mean, geometric SD and geometric CV%.

All individual ACT-541468 plasma concentration data will be listed.

9 PHARMACOKINETIC/PHARMACODYNAMIC ANALYSES

9.1 Exposure-safety analyses

The exposure-safety relationship will be explored using $C_{9–10h}$ (plasma concentrations of ACT-541468 in the morning after the second PSG night at Visit 6 and Visit 8) and will be based on the PK set. Scatter plots will be provided to explore the relationship between the concentration of ACT-541468 and Coding sub-test© and VAS score assessing morning sleepiness. These safety endpoints include only values assessed in the morning or evening after the second PSG night at Visit 6 and Visit 8.

Descriptive statistics [as per Section 10.1] of the ACT-541468 plasma concentrations collected approximately 9–10 h post-dose in the morning after the second PSG night at

Visit 6 and Visit 8 will be provided for subjects with and without clinically significant findings for neurological examination at the matching time point.

9.2 Exposure-efficacy analyses

The exposure-efficacy relationship will be explored using C_{9-10h} (plasma concentrations of ACT-541468 in the morning after the second PSG night at Visit 6 and Visit 8) and will be based on the PK set. Scatter plots will be provided to explore the relationship between the concentration of ACT-541468 and WASO, LPS, sTST, IDSIQ Sleepiness domain score and VAS scores assessing sleep quality, sleep depth, daytime alertness and daytime ability to function. These efficacy endpoints include only values assessed during or in the morning after the second PSG night at Visit 6 and Visit 8.

10 GENERAL STATISTICAL METHODOLOGY

10.1 Handling of partially missing data

Partially missing data for primary endpoint WASO and LPS values, and for variables like TST, number of shifts from S2, SWS or REM to S1 or awake, number of awakenings, Coding sub-test©, SDS©, or neurological examination, will be handled as follows: if one of the two values is missing either for baseline, Month 1 or Month 3, the single value available will be used as the mean for this time point. Otherwise, the mean value will be considered missing for that time point.

For sTST, sWASO, sLSO, IDSIQ total and domain scores, VAS scores and number of self-reported awakenings, subjects must have at least 2 days of data during each week to calculate a weekly mean. Otherwise, the mean value will be considered missing for that week.

The approaches above imply implicit imputation: missing data points are given the same value as the mean of the non-missing data points of that same time point or week.

10.2 General rules for data presentations

Data are listed and summarized as described below.

The tables will use the following header structure (label and order):

ACT-541468 25 mg *N = xxx*	*ACT-541468 50 mg* *N = xxx*	*Placebo* *N = xxx*

Where N indicates the total number randomized or treated appropriate to the analysis set in the corresponding treatment group, unless otherwise specified.

All listings will be sorted by randomized treatment or actual treatment received (appropriate to the analysis set), country, subject number (ascending) and, when appropriate, by visit / date of assessment (ascending). Listings related to the SCR set will present a treatment group label “screening failure” to indicate subjects who were not

randomized and will be listed after the subjects who were randomized or received DB study treatment, as relevant.

Unless noted otherwise, the following descriptive statistics will be used to summarize data: number (%) of subjects for categorical variables or descriptive statistics (number of non-missing values, mean, SD, median, Q1, Q3, minimum, maximum) for continuous variables.

11 INTERIM ANALYSES

No formal interim analysis will be performed for determining whether to stop (or modify) the study early (i.e., no hypothesis testing will be conducted *ad interim*). Therefore, no adjustment for multiple testing is required. This study includes an IDMC and ISB that will assess safety of ACT-541468 on a regular basis as per the IDMC and ISB charters, respectively. Safety and efficacy data supporting the review by the IDMC will be provided by Idorsia for the part of analyses that are blinded and by an ISAC for the unblinded part.

12 GENERAL DEFINITIONS AND DERIVATIONS

12.1 Treatment start and end dates

Run-in single-blind placebo treatment start or end date is the earliest or latest, respectively, date of dose intake recorded on the 'Study Single-Blind (Run-In) Treatment Log' page.

DB study treatment start or end date is the earliest or latest, respectively, date of dose intake recorded on the 'Study Double Blind Treatment Log' page.

Run-out single-blind placebo treatment start and end date is the earliest and latest date of dose intake recorded on the 'Study Single-Blind (Run-Out) Treatment Log' page.

12.2 Study periods

The Screening period is defined as the time from the informed consent date until one day before the Run-in period start date.

The Run-in period is defined as the time from the start of the run-in single-blind placebo treatment to one day before start of DB study treatment or, for those not starting DB study treatment, to run-in single-blind placebo treatment end date.

The DB study period is defined as the time from the day of DB study treatment start until the day of EODBT (Visit 8, 2nd morning), or, for those who prematurely discontinued DB study treatment (i.e., those without an EODBT date), until one day after DB study treatment end date (if the treatment is taken before midnight) or the day of DB study treatment end date (if the treatment is taken after midnight). Time of treatment intake between 00:00 and 11:59 is considered as 'after midnight', while a time between 12:00 and 23:59 is considered as 'before midnight'. For any missing or incomplete time, it is presumed that the treatment was taken before midnight. For subjects who complete DB study treatment with an

interruption before continuing the run-out single-blind placebo treatment, the end of DB study period is the run-out single-blind placebo treatment start date – 1 day.

For those subjects in the TWS, the treatment withdrawal period (or Run-out period) is defined as the time from one day after the end of the DB study period to the latter of seven days after the end of the DB study period or Visit 10 date.

The safety follow-up period is defined as the time from one day after the end of the treatment withdrawal period for those starting the treatment withdrawal period (i.e., subjects in the TWS) until EOS date; or from one day after the end of the DB study period for those not starting the treatment withdrawal period until EOS date.

12.3 Treatment day and Study day

The Treatment day for an assessment or event will be calculated using the DB study treatment start date as reference.

For assessments/events occurring on or after the start date of DB study treatment, Treatment day will be positive and will be calculated as:

Treatment day (days) = Date of assessment/event – Start date of DB study treatment + 1 day

The first day of DB study treatment is Treatment day 1.

For all assessment/events occurring prior to the start date of DB study treatment, Treatment day will be negative and will be calculated as:

Treatment day (days) = Date of assessment/event – Start date of DB study treatment

The Study day for an assessment or event will be calculated using the randomization date as reference.

For assessments/events occurring on or after the randomization date, Study day will be positive and will be calculated as:

Study day (days) = Date of assessment/event – randomization date + 1 day

The day of randomization date is Study day 1.

For all assessment/events occurring prior to the randomization date, Study day will be negative and will be calculated as:

Study day (days) = Date of assessment/event – randomization date

Treatment day and/or Study day will be displayed in the data listings as appropriate.

12.4 Baseline

Baseline is the last non-missing assessment performed or value measured before or on the day of first dose of DB study treatment, unless otherwise defined in the specific analysis section.

An assessment done on the day of the DB study treatment start date is considered a baseline assessment and is not considered a post-baseline assessment in the DB study period.

Subjects with no data on a particular parameter before the first treatment administration will have a missing baseline (and change from baseline) for this parameter.

12.5 Change from baseline

The change from baseline is defined as post-baseline value (any assessment performed after baseline and up to EOS) minus baseline value. A positive number indicates an increase as compared to baseline.

12.6 Handling of data when total sleep time is zero

If the TST is recorded as zero (i.e., subject did not fall asleep during the night), then:

- WASO is set as missing since subject is never at risk, and
- LPS is set as 480 minutes which is equal to the duration of time in bed.

If the sTST is recorded as zero, then similarly:

- sWASO is set as missing, and
- sLSO is set as 480 minutes.

12.7 Missing data patterns

Monotone missing data pattern: Given a dataset with variables, Y1, Y2, …, Yt (in that order) and Vj is missing for a particular subject, then all subsequent variables Vk, $k > j$, are missing for that subject. In addition, given the same dataset and Vj is observed for a particular subject, then all previous variables Vk, $k < j$, are also observed for that subject. Monotone missing data patterns are typically encountered when a subject prematurely withdraws from a study and is expected to be more likely seen than non-monotone missing data patterns.

In the examples below, an 'X' indicates that the variable is observed in the corresponding group and a '.' means that the variable is missing.

Example of monotone missing data patterns:

Subject	Y1	Y2	Y3
1	X	X	X
2	X	X	.
3	X	.	.

When missing data are non-monotone, data can also be missing at intermediate visits (i.e., missing a given visit, but observed again at the next visit). Non-monotone missing data patterns are commonly turned into monotone missing data patterns using MI (based on MCMC), which are then subsequently analyzed using MI methods based on monotone missing data.

Example of non-monotone missing data patterns:

Subject	Y1	Y2	Y3
1	X	X	X
2	.	X	X
3	.	.	X
4	.	X	.
5	X	.	X

12.8 Time window definitions for calculating subjective endpoints

The meaning of the first PSG end date and the last PSG end date at a given visit (Visit 3, Visit 6 and Visit 8) are shown in Table 6.

Table 6 Meaning of the first and last PSG end date at Visits 3, 6 and 8

Visit 3, Visit 6 and Visit 8			
First PSG		**Last PSG**	
Start date (evening)	**End date (morning)**	Start date (evening)	**End date (morning)**

Values falling in the following time windows will be used to calculate weekly averages for the subjective endpoints; values from PSG night will not be included:

Screening: the 7 consecutive days immediately preceding the Run-in single-blind placebo treatment start date but after PSG end date at Visit 1 (morning) (i.e., 7-day window = [Max. (Run-in single-blind placebo treatment start date – 7 days; PSG end date at Visit 1 [morning]); Run-in single-blind placebo treatment start date − 1 day]).

Baseline (Run-in period): the 7 consecutive days immediately preceding the first PSG at Visit 3, but after Run-in single-blind placebo treatment start date (i.e., 7-day window = [Max. (End date of first PSG at Visit 3 – 7 days; Run-in single-blind placebo treatment start date); End date of first PSG at Visit 3 − 1 day]). The time window for Baseline may be less than 7 days. Baseline will be missing for any subjects without a PSG at Visit 3.

Month 1: the 7 consecutive days immediately preceding the first PSG at Visit 6 (i.e., 7-day window = [End date of first PSG at Visit 6 – 7 days; End date of first PSG at Visit 6 – 1 day]). For subjects missing PSG at Visit 6, the 7-day window = [Study day 23; Study day 29].

Month 3: the 7 consecutive days immediately preceding the first PSG morning at (Visit 8 i.e., 7-day window = [End date of first PSG at Visit 8 – 7 days; End date of first PSG at

Visit 8 – 1 day]). For subjects missing PSG at Visit 8, the 7-day window = [Study day 79; Study day 85].

Run-out: the 7 consecutive days immediately following the PSG morning at Visit 9 (i.e., 7-day window = [Date of PSG morning at Visit 9 + 1 day; Min. (Date of PSG morning at Visit 9 + 7 days; Visit 10 date)]). For subjects missing PSG at V9, the 7-day window = [Treatment withdrawal period start date; Min. (Treatment withdrawal period start date + 6 days; Visit 10 date)].

Table 7 defines the time windows used to calculate weekly averages for certain subjective measures that will be summarized per week. These subjective measures include sTST, sWASO, sLSO and IDSIQ scores (i.e., Total score; Alert/Cognition, Mood and Sleepiness domain scores).

Table 7 Time windows (in days) to calculate weekly averages for certain subjective endpoints

Week	Study day (7-day window)
1	2–8
2	9–15
3	16–22
4*	23–29
5	30–36
6	37–43
7	44–50
8	51–57
9	58–64
10	65–71
11	72–78
12*	79–85

* These days might not be the same as in the definition above (Month 1, Month 3).
Values from PSG nights and Run-out period will not be included.

13 CHANGES OR CLARIFICATIONS TO ANALYSES PLANNED IN THE STUDY PROTOCOL

13.1 Changes to the analyses planned in the study protocol

The baseline for PGA-S scores (daytime symptoms) is defined from Visit 3 and not from Visit 1 as stated in the protocol.

Assessments for potential rebound insomnia will be based only on WASO, LPS and sTST for consistency with the primary and secondary endpoints. TST, sWASO and sLSO have been removed.

13.2 Changes in the conduct of the study / data collection

Not applicable.

13.3 Clarifications concerning endpoint definitions and related variables or statistical methods

The VAS scores related to depth of sleep, daytime alertness and ability to function will be analyzed as efficacy variables [Section 6.3] and not as safety variables. The VAS score remaining as a safety variable is the question from the sleep diary related to morning sleepiness [Section 7.8].

TEAEs that started or worsened during the safety follow-up period will only be listed. Summary tables for TEAEs, TEAEs related to study treatment, TEAEs related to abuse and TEAEs based on selected Standardised MedDRA Queries or pre-defined search criteria will include only those TEAEs occurring during the DB study period.

The handling of the missing data for the weekly calculation of the subjective endpoints, which includes the secondary endpoints, will be aligned: subjects must have at least 2 days of data during each week to calculate a weekly mean [Section 10.1].

13.4 Additional analyses as compared to the study protocol

A summary of the number (%) of subjects with ≥ 6 points decrease in ISI© Total score from baseline (Visit 3) to Month 1 (Visit 6) and Month 3 (Visit 8) is added to the analysis of exploratory variables [Section 6.3].

The change from baseline to Month 2 (Visit 7) and to Run-out period (Visit 9 and 10) is added to the vital signs analyses [Section 7.5].

A summary of the neurological examination results is added to the next-day residual effect analysis [Section 7.8].

The occurrence of suicidal ideation and/or behavior during the treatment withdrawal period is added to C-SSRS© analysis [Section 7.9].

A summary of the ESS© [Section 7.10] is added to the safety analyses.

Exposure-efficacy analysis is added to the pharmacokinetic/pharmacodynamic analyses [Section 9.2].

14 REFERENCES

[Bodner 2008] Bodner TE. What Improves with Increased Missing Data Imputations? Structural Equation Modeling 2008;15(4):651–75.

[Bretz 2009] Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. Stat Med 2009;28(4):586–604.

[Carpenter 2013] Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant accessible assumptions and inference via multiple imputation. J Biopharm Stat 2013;23(6):1352–71.

[Graham 2007] Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007;8(3):206–13.

[Kenward 1997] Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997;53(3):983–97.

[Mallinckrodt 2013] Mallinckrodt C, Roger J, Chuang-stein C, Molenberghs G. Missing data: turning guidance into action. Stat Biopharm Res 2013;4:369–82.

[O'Kelly 2014] O'Kelly M, Ratitch B. Clinical Trials with Missing Data: A Guide for Practitioners. 1st ed. Wiley; 2014.

[Ratitch 2013] Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. Pharm Stat 2013;12:337–47.

[Rubin 1987] Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons; 1987.

[Schafer 1997] Schafer JL. Analysis of incomplete multivariate data. New York: Chapman & Hall; 1997.

[Von Hippel 2009] Von Hippel PT. How to impute interactions, squares, and other transformed variables. Sociological Methodology 2009;39(1):265–91.

[White 2011] White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011;30(4):377–399.

[Yan 2009] Yan X, Lee S, Li N. Missing data handling methods in medical device clinical trials. J Biopharm Stat 2009;19:1085–98.

15 APPENDICES

15.1 SAS® code for Mixed Model for Repeated Measures

```
proc mixed data=data;
        class SUBJID AGEGR TRT VIS;
        model CHG = BASE AGEGR TRT VIS TRT*VIS BASE*VIS / ddfm=kr;
        lsmeans TRT*VIS / diff cl;
        repeated VIS / subject=SUBJID type=un;
        estimate '25 mg vs placebo at month1' TRT 1 0 -1 TRT*VIS 1 0 0 0 -1 0 / cl;
        estimate '25 mg vs placebo at month3' TRT 1 0 -1 TRT*VIS 0 1 0 0 0 -1 / cl;
        estimate '50 mg vs placebo at month1' TRT 0 1 -1 TRT*VIS 0 0 1 0 -1 0 / cl;
        estimate '50 mg vs placebo at month3' TRT 0 1 -1 TRT*VIS 0 0 0 1 0 -1 / cl;
run;

/* contrasts are built considering the following treatment order: low dose, high dose, placebo */
```

15.2 SAS® code for multiple imputation with MCMC for monotone missingness mechanism

```
ods graphic on;
proc mi data=data_wide nimpute=100 seed=708695 minimum=. 0 0 0 maximum=. 480 480 480 out=data_mono;
        by _TRT_;
        var AGE BASE OBSM1 OBSM3; /* age is continuous */
        mcmc impute= monotone plots=(trace(mean(OBSM1 OBSM3)));
run;
ods graphic off;
```

15.3 SAS® code for multiple imputation under MAR assumption

```
/* multiple imputation */
proc mi data=data_mono nimpute=1 seed=708695 minimum=. . 0 0 0 maximum=. . 480 480 480 out=data_mi noprint;
        by _IMPUTATION_; /* flag variable for imputation from proc mi */
        class TRT AGEGR;
        var TRT AGEGR BASE OBSM1 OBSM3;
        monotone regression;
run;

/* analyze and combine the results */
ods output lsmeans=mixparms diffs=diffparms(where=(_TRT='PLACEBO' and VIS=_VIS));
proc mixed data=data_mi_long;
```

```
        by _IMPUTATION_; /* flag variable for imputation from proc mi */
        class SUBJID AGEGR TRT VIS;
        model CHG = BASE AGEGR TRT VIS TRT*VIS BASE*VIS / ddfm=kr;
        lsmeans TRT*VIS / diff cl;
        repeated VIS / subject=SUBJID type=un;
run;
proc mianalyze parms=mixparms;
        class TRT VIS;
        modeleffects TRT*VIS;
run;
proc mianalyze parms =diffparms;
        class TRT VIS;
        modeleffects TRT*VIS;
run;
```

15.4 SAS® code for multiple imputation under MNAR assumption for jump to reference method

```
/*** multiple imputation ***/

/* OBSM1 */
proc mi data=data_mono nimpute=1 seed=708695 minimum=. . 0 0 maximum=. . 480 480
out=data_mi2 noprint;
        by _IMPUTATION_; /* flag variable for imputation from proc mi */
        class TRT AGEGR;
        var AGEGR BASE OBSM1;
        monotone regression;
        mnar model (OBSM1 / modelobs=(TRT='placebo'));
run;

/* OBSM3 */
proc mi data=data_mi2 nimpute=1 seed=708695 minimum=. . 0 0 maximum=. . 480 480
out=data_mi3 noprint;
        by _IMPUTATION_; /* flag variable for imputation from proc mi */
        class TRT AGEGR;
        var AGEGR BASE OBSM3;
        monotone regression;
        mnar model (OBSM3 / modelobs=(TRT='placebo'));
run;
```

15.5 SAS® code for multiple imputation under MNAR assumption for delta-adjusted (tipping point)

```
%macro midata(data=, smin=, smax=, sinc=, out=);
data &out;
        set _null_;
run;
/*------------ # of shift values ------------*/
%let ncase= %sysevalf( (&smax-&smin)/&sinc, ceil );
/*------- Imputed data for each shift -------*/
%do jc=0 %to &ncase;
%let sj= %sysevalf( &smin + &jc * &sinc);
proc mi data=&data seed=708695 nimpute=1 minimum=. . 0 0 0 maximum=. . 480 480
480 out=outmi;
        by _IMPUTATION_; /* flag variable for imputation from proc mi */
        class TRT AGEGR;
        monotone regression;
        mnar adjust(OBSM1 / shift=&sj adjustobs=(TRT = 'ACT-541468 25mg'));
        mnar adjust(OBSM1 / shift=&sj adjustobs=(TRT = 'ACT-541468 50mg'));
        mnar adjust(OBSM3 / shift=&sj adjustobs=(TRT = 'ACT-541468 25mg'));
        mnar adjust(OBSM3 / shift=&sj adjustobs=(TRT = 'ACT-541468 50mg'));
        var TRT AGEGR BASE OBSM1 OBSM3;
run;
data outmi;
        set outmi;
        Shift= &sj;
run;
data &out;
        set &out outmi;
run;
%end;
%mend midata;

%midata(data=data_mono, smin=0, smax=60, sinc=15, out=data_mi4);   /*  shifts  for
WASO */
```