Statistical Analysis Plan AC-078A201

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

ClinicalTrials.gov Identifier NCT02839200

The study was sponsored was sponsored by Actelion Pharmaceuticals Ltd.
Study sponsorship was transferred to Idorsia Pharmaceuticals Ltd in July 2018.

STATISTICAL ANALYSIS PLAN FOR CLINICAL STUDY REPORT

Protocol AC-078A201

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

Investigational Drug	ACT-541468
Actelion Document Number	D-17.384
Document Status/Version Number	Final Version 2.0
Date	3 July 2017
Author	[redacted]
Reviewers	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AE	Adverse event
AESI	AEs of special interest
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
CI	Confidence interval
CRO	Contract research organization
CSR	Clinical study report
C-SSRS©	Columbia Suicide Severity Rating Scale
DSST©	Digit Symbol Substitution Test
ECG	Electrocardiogram
ED_{50}	Dose that is expected to give half the maximum change
E_{max}	Maximum change in effect associated with dose
EODB	End of Double-Blind treatment
EOS	End of Study
FAS	Full analysis set
IDMC	Independent Data Monitoring Committee
ISB	Independent Safety Board
ISI©	Insomnia Severity Index
KSS	Karolinska Sleepiness Scale
LLOQ	Lower Limit of Quantification
LPS	Latency to Persistent Sleep
LS	Least squares
MCP-Mod	Multiple Comparison Procedure – Modeling
MedDRA	Medical Dictionary for Regulatory Activities
mFAS	Modified full analysis set
NAW	Number of awakenings
PK	Pharmacokinetic

PPS	Per-protocol analysis set
PSG	Polysomnography
PT	Preferred term
QTcB	QT corrected according to Bazett's formula
QTcF	QT corrected according to Fridericia's formula
REM	Rapid eye movement
RND	Randomized analysis set
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical analysis system
SCR	Screened analysis set
SDS©	Sheehan Disability Scale
sLSO	Subjective Latency to Sleep Onset
sNAW	Subjective NAW
SOC	System organ class
SQ	Sleep quality
SS	Safety set
sTST	Subjective TST
sWASO	Subjective Wake After Sleep Onset
SWS	Slow wave sleep
TD	Target dose
TIB	Time in bed
TST	Total Sleep Time
VAS	Visual Analog Scale
WASO	Wake After Sleep Onset
WHO	World Health Organization
WHO DRL	WHO Drug Reference Listing

1 INTRODUCTION

This statistical analysis plan (SAP) describes in detail the analyses and data presentation for the final clinical study report (CSR).

Obvious corrections to address minor formatting errors or spelling mistakes may be performed at the time of analysis without amending this document.

Data will be analyzed by Actelion and/or designated contract research organizations (CROs) using Statistical Analysis System (SAS) version 9.3, and for the Multiple Comparison Procedure – Modeling (MCP-Mod) analysis, R version 3.1.2 will be used. The analyses for the closed session in the Independent Data Monitoring Committee (IDMC) meetings will be performed by an independent statistical analysis center.

Protocol AC-078A201, Final Version 2, dated 3 August 2016 was used when writing this SAP.

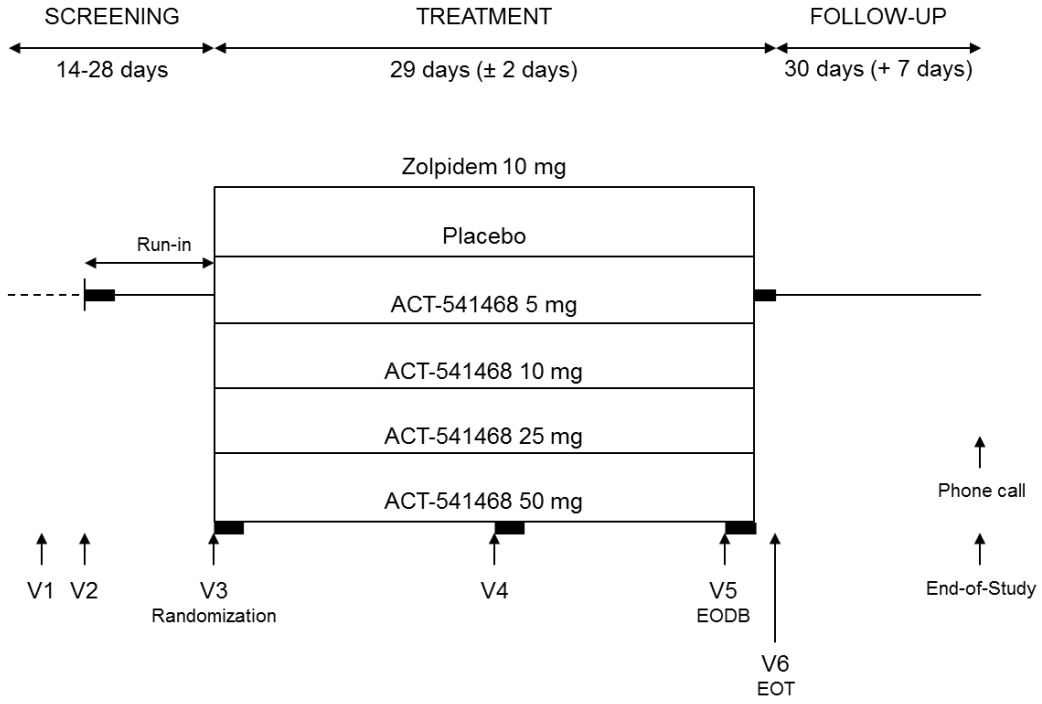
2 STUDY DESIGN

This is a multi-center, double-blind, randomized, placebo-controlled, active-reference, parallel-group, polysomnography (PSG) dose-response Phase 2 study to assess the efficacy and safety of ACT-541468 in adult subjects with insomnia disorder.

Approximately 300 subjects will be randomized (stratified by gender) in a 1:1:1:1:1:1 ratio to either ACT-541468 5 mg, 10 mg, 25 mg, 50 mg, placebo or zolpidem 10 mg.

The study design is shown in Figure 1. Double-blind study treatment will be administered from V3 to the End of Double-Blind treatment (EODB - V5); an expected duration of 29 days. The double-blind treatment period starts with the first dose of double-blind study treatment in the first evening of the Randomization Visit (V3, Day 1). The EODB is reached in the second morning of V5 (Day 30) after the last dose of double-blind study treatment and after all second morning assessments have been performed.

Figure 1 Study design



Note: V = Visit; ▬ = PSG nights; EODB = End of Double-Blind treatment; EOT = End of Treatment.

3 OBJECTIVES

3.1 Primary objective

The primary objective of the study is to evaluate the dose response of ACT-541468 on the change of Wake After Sleep Onset (WASO) assessed by PSG after treatment on Days 1&2.

3.2 Secondary objectives

The secondary objectives of the study are to evaluate the dose response relationship of ACT-541468 on Latency to Persistent Sleep (LPS) on Days 1&2, and subjective Latency to Sleep Onset (sLSO) and subjective WASO (sWASO) at Week 4.

3.3 Other objectives

To explore the effect of ACT-541468 on sleep parameters during the study, sleep parameters will include various objective and subjective measures.

To assess the safety and tolerability of oral administration of ACT-541468 in subjects with insomnia disorder.

To explore the relationship between ACT-541468 exposure (concentration of ACT-541468 approximately 9–10 h post-dose at second morning of V3, V4, and V5) and safety.

4 GENERAL ANALYSIS CONSIDERATIONS

Categorical data will be summarized by frequency counts and percentages. Unless specified otherwise, the percentage will be based on the corresponding treatment group size.

Continuous data will be summarized using descriptive statistics (N, mean, standard deviation, median, minimum and maximum).

Unless noted otherwise, all statistical summaries will be presented by treatment group and in the following order (Placebo, ACT-541468 5 mg, ACT-541468 10 mg, ACT-541468 25 mg, ACT-541468 50 mg, zolpidem 10 mg).

5 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 each analysis set will be summarized and listed.

5.1 Screened analysis set

The screened analysis set (SCR set) includes 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).

5.2 Randomized analysis set

The randomized analysis set (RND set) includes all subjects who have been assigned to a double-blind study treatment.

5.3 Full analysis set

The Full analysis set (FAS) includes all subjects from the RND set who received at least one dose of double-blind study treatment.

In order to adhere to the intention-to-treat principle:

- Subjects will be evaluated according to the study treatment they have been assigned to (which may be different from the study treatment they actually receive),
- All available data are included.

5.4 Modified full analysis set

The modified FAS (mFAS) includes all subjects from the FAS and who have at least one WASO assessment at baseline and one at Days 1&2.

5.5 Per-protocol analysis set

The Per-protocol analysis set (PPS) comprises all subjects from the mFAS who have two consecutive WASO values at Baseline and at Days 1&2, and who complied with the protocol sufficiently to allow relevant assessment of treatment effects.

Criteria for sufficient compliance:

- Subject has taken required treatment on both PSG nights during run-in period and Days 1&2,
- Absence of major protocol deviations (i.e., those leading to exclusion from the PPS), as defined in a separate protocol deviation document.

Subjects not meeting the criteria above will be excluded from the PPS.

5.6 Safety set

The Safety set (SS) includes all subjects who received at least one dose of double-blind study treatment. Subjects will be evaluated according to the actual treatment they received (based on dispensed kit number), which may differ from the randomly assigned treatment.

5.7 Withdrawal set

The Withdrawal set comprises all subjects included in the SS who received single-blind placebo treatment (at least one capsule) in the run-out period.

5.8 Pharmacokinetic analysis set

The pharmacokinetic analysis set (PK set) includes all subjects in the SS who have at least one PK sample collected after initiation of study drug. Subjects receiving placebo or zolpidem will be excluded from the PK set.

6 STUDY SUBJECTS VARIABLES AND ANALYSES

6.1 Screening failures

The following summaries will be based on the SCR set:

- Number (%) of subjects who discontinued during screening period (based on ‘Was the subject randomized?’ recorded as ‘No’ in the ‘Randomization’ page),
- Number (%) of subjects who discontinued during screening period, but did not enter the run-in period (based on ‘Was the subject randomized?’ recorded as ‘No’ in the ‘Randomization’ page, missing ‘Date of capsules dispensed’ in the ‘Study Single

Blind Treatment Dispensing & Accountability' page at V2 and no assessments at V2),

- Number (%) of subjects who failed screening during the run-in period (based on 'Was the subject randomized?' recorded as 'No' in the 'Randomization' page, and either a non-missing 'Date of capsules dispensed' in the 'Study Single Blind Treatment Dispensing & Accountability' page at V2, or a non-missing assessment recorded at V2),
- Primary reasons for screening period discontinuation (i.e., inclusion/exclusion criteria not met, subject withdrew consent, or other).

Subjects may be re-screened once. For these subjects, only the timing and reason of the last screen failure will be reported in the summary table. Re-screened subjects that are randomized will be excluded from this summary. However, all reasons for screening failure will be included in the listing.

Demographic data for subjects who fail screening and corresponding reason(s) for screen failure will be listed.

6.2 Subject disposition

The following will be based on the SCR set:

- Number of subjects screened,
- Number of subjects treated in single-blind placebo treatment run-in period (based on non-missing date of capsules dispensed at Visit 2),
- Number of subjects randomized (based on non-missing randomization number),
- Number of subjects treated in double-blind treatment period (based on non-missing treatment start date),
- Number of subjects treated at last PSG treatment nights,
- Number of subjects treated in single-blind placebo treatment run-out period (based on non-missing date of capsules dispensed at Visit 6),
- Number of subjects completing the end of study (EOS) safety follow-up (based on non-missing date of phone call on the 'Safety Follow-up Phone Call Summary - EOS' page).

The following summaries will be based on the FAS:

- Number (%) of subjects who prematurely discontinued double-blind study treatment (based on discontinuation reason entered as 'Premature Discontinuation' in the 'Study Double-blind Treatment Log' page),
- Primary reasons for premature double-blind study treatment discontinuation (based on discontinuation reason entered on the 'Premature Discontinuation of Study Treatment' page).

The following summaries will be based on the RND set:

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

Premature treatment and study discontinuations will be listed based on the RND set.

6.3 Protocol deviations

The RND set will be used for the protocol deviation summary tables and the SCR set for listings. All important and other protocol deviations will be summarized per pre-specified category (i.e., During screening period; During run-in period (V2); During double-blind treatment phase; During follow-up phase; Non study phase specific) by treatment group and overall. A subject with multiple occurrences of a protocol deviation is counted only once per protocol deviation category. A listing of all protocol deviations will be provided.

6.4 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 overall. Data will be listed individually by subject.

6.4.1 Demographics

Demographic data including age, gender, race, ethnicity, height, and baseline weight and body mass index will be listed and summarized. In addition, age categories (< 18, 18–< 50, ≥ 50 years), and BMI categories (< 25, 25–30, > 30 kg/m^2) will be summarized.

6.4.2 Medical history and current medical conditions

Relevant medical history and current medical conditions will be coded using MedDRA terminology. Medical history and current (ongoing) medical conditions, excluding insomnia-related conditions and symptoms, will be summarized and listed. Summaries will be presented by primary system organ class (SOC) and preferred term (PT). The MedDRA version used for reporting will be specified in the footnote of the applicable output.

6.4.3 Previous and concomitant therapies

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

Number (%) of subjects having taken at least one previous or concomitant treatment will be summarized by ATC class and individual PT within each ATC class based on the SS.

Summaries will be provided for previous, study-concomitant and double-blind study treatment-concomitant therapies defined below.

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.

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 30 days after study treatment discontinuation. 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 double-blind study treatment or initiated during the double-blind treatment period until 1 day after the last dose of double-blind study treatment.

All concomitant therapies will be listed using the SS.

An incomplete (day or month missing) or missing concomitant therapy date will be imputed as described in the table below. 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	If the (imputed) concomitant therapy end date is on or after the start of double-blind study treatment, and if the study treatment start falls within the upper and lower limits (inclusive), the study treatment start date is used. 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. If the concomitant therapy resolution date is missing, then: 1. if the double-blind study treatment start falls within the upper and lower limits (inclusive), the double-blind study treatment start date is used else: 2. 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. In all the other cases, the lower limit is used.	Whichever is the earlier of the concomitant therapy end date or double-blind study treatment start date.

6.5 Study treatment exposure and compliance

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

Duration of treatment, reasons for treatment end and compliance data will be listed.

6.5.1 Exposure

Duration of double-blind study treatment (in days) is defined as the difference between the treatment end date and the treatment start date plus one day. This calculation ignores periods of treatment interruption. The duration of double-blind study treatment will be summarized.

6.5.2 Compliance

Compliance will be assessed through study treatment accountability. Compliance will be based on a non-missing date recorded for 'Last Study Treatment Intake date' in the 'Polysomnography (PSG)' page, and an unmarked record for 'I did not take my study medication' in the sleep diary.

Compliance to double-blind treatment, defined as the percentage of capsules taken compared to the planned capsules during the double-blind treatment period (V3 to V5 [EODB]), will be summarized. The number (%) of subjects with compliance to double-blind treatment < 70% will be tabulated.

Compliance to double-blind treatment for each week (cumulatively) will be summarized. The number (%) of subjects with compliance to double-blind treatment < 70% for every cumulative week will be tabulated.

Week 1 (7 consecutive days) includes and extends until 6 days after the start date of double-blind treatment (V3 - 1st PSG evening). Weeks 1&2 includes Week 1, and extends until one day before the date of double-blind treatment at V4 - 1st PSG evening. Weeks 1&2&3 includes Weeks 1&2, and includes and extends until 6 days after the date of double-blind treatment at V4 - 1st PSG evening. Weeks 1&2&3&4 includes Weeks 1&2&3 and extends until one day before the date of double-blind treatment at V5 - 1st PSG evening.

7 ANALYSIS OF THE PRIMARY VARIABLE

7.1 Endpoint: Objective sleep maintenance

The primary efficacy endpoint relates to objective sleep maintenance and is assessed through the absolute change in WASO from Baseline to Days 1&2.

Baseline is the mean of the two PSG nights during the run-in period (V2). Days 1&2 is the mean of the corresponding two PSG treatment nights at V3.

WASO is the time (in minutes) spent awake after onset of persistent sleep (see definition of LPS in Section 8.2) until lights on as determined by PSG. The PSG recording is centrally scored by independent scorers.

7.2 Handling of missing data

No imputation will be performed for missing WASO values. However, if one of the two values is missing either for Baseline or for Days 1&2, the single value available will be used as the mean for this time point. This implies implicit imputation: the missing data point is given the same value as the non-missing data point.

7.3 Statistical hypothesis, model and method of analysis

The primary statistical analysis will be performed on the mFAS.

To evaluate the dose response relationship of ACT-541468 on the change from baseline in WASO, this study uses the generalized MCP-Mod methodology [Bretz 2005, Pinheiro 2006, Pinheiro 2014]. This approach combines a Multiple Comparison Procedure (MCP) to assess the efficacy of ACT-541468 versus placebo followed by a modeling (Mod) step to characterize the dose response relationship and to identify a dose (or dose range) that has shown signs of a clinically relevant effect.

The MCP-Mod analysis will be performed using the R *DoseFinding* package [Bornkamp 2016].

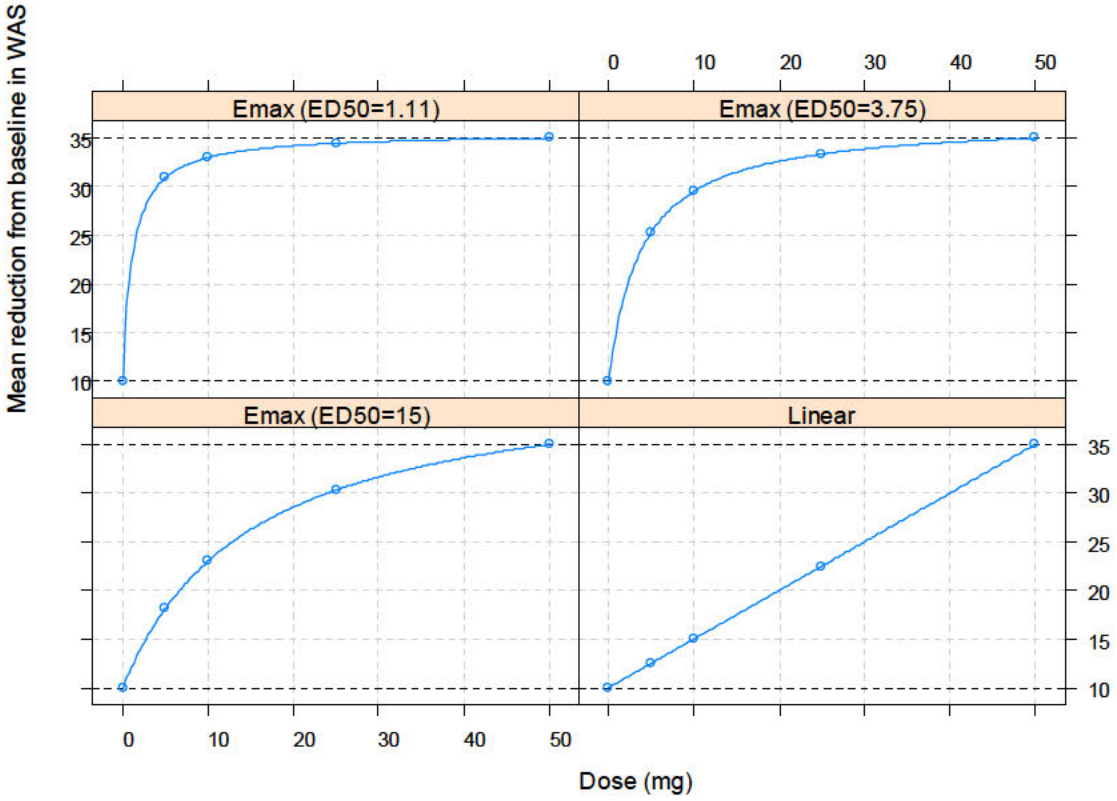
Details on how the primary endpoint will be analyzed within the MCP-Mod framework are provided below.

7.3.1 Design

7.3.1.1 Candidate dose response models

To begin, the set of pre-specified candidate parametric dose response models that could possibly represent the underlying dose response shape are determined. For this study, four pre-specified dose response models are considered: three E_{max} models and one linear model. A graphical display of these models are shown in Figure 2. Here, the maximum mean reduction in WASO from baseline to Days 1&2 with ACT-541468 50 mg dose is assumed to be 25 minutes longer than with placebo.

Figure 2 Candidate dose response models



The following describes how the parameters (i.e., dose that is expected to give half the maximal change [ED_{50}]) of these models are calculated.

For the purpose of determining optimal contrasts, only the standardized form $f^0(d,\theta^0)$ of the dose response model $f(d,\theta)=\theta_0+\theta_1 f^0(d,\theta^0)$ needs to be considered [Bretz 2005], i.e., the scale parameters (θ_0 and θ_1) are removed from the functional form.

The functional form of the E_{max} model is defined as $f(d,\theta)=E_0+E_{max}\frac{d}{ED_{50}+d}$, where d = dose; E_0 = placebo (basal) effect; E_{max} = maximum change in effect associated with dose; ED_{50} = dose that is expected to give half of the maximum change.

The standardized form of the E_{max} model is therefore $f^0(d,\theta^0)=\frac{d}{ED_{50}+d}$, which represents the percentage (p) of the maximum change from the basal effect associated with dose (d) (since the standardized E_{max} model is the same as the functional form when

E_0 = 0 and E_{max} = 1). Rearranging the formula above gives $ED_{50} = \frac{d(1-p)}{p}$. The specification of ED_{50} will fully parameterize the standardized form of the model.

The ED_{50} for each E_{max} model (i.e., 1.11, 3.75, 15) was calculated assuming a p of 90%, 80% and 40% at an ACT-541468 dose (d) of 10 mg, 15 mg, and 10 mg, respectively.

The calculation of each optimal contrast [see Section 7.3.2.1] depends on the pre-defined estimates of ED_{50} and dose levels (including placebo as 0 mg). Since three E_{max} models are pre-specified, three different sets of testing contrasts will be determined for the E_{max} model (family).

The functional form of the linear model is defined as $f(d,\theta) = E_0 + \delta d$, where d = dose; E_0 = placebo (basal) effect. Hence, the standardized form of the linear model is a function of dose only $f^0(d) = d$. Since no parameter estimates θ^0 are required, the determination of the appropriate single testing contrast is derived based on dose alone.

To ensure no inflation of type I error, the pre-specified parameter values θ^0 (i.e., ED_{50}) are not to be changed after the blind is broken.

7.3.2 Analysis

Key components of MCP-Mod at the analysis stage include:

- MCP step: Testing for the presence of a dose response signal,
- Mod step: Selecting the best dose response model for the observed data out of a pre-specified set of candidate models, and
- Mod step: Estimating target dose(s; TDs) of interest via modeling.

7.3.2.1 Implementation of the MCP step

The MCP step consists of testing a set of optimal contrasts: each testing contrast representing one of the pre-specified dose response models. At the same time, this step addresses the issue of possible model misspecifications (the reason why multiple candidate dose response shapes are pre-defined) and also includes the associated adjustments for multiplicity within a rigorous hypothesis testing framework.

For a single given candidate dose response model, the optimal contrast for testing the hypothesis of a flat dose-response profile with maximal power is given by:

$$\boldsymbol{c}_m^{opt} \propto \hat{\boldsymbol{S}}^{-1}\left(\boldsymbol{\mu}_m^0 - \frac{\boldsymbol{\mu}_m^{0\prime}\hat{\boldsymbol{S}}^{-1}\boldsymbol{1}}{\boldsymbol{1}'\hat{\boldsymbol{S}}^{-1}\boldsymbol{1}}\right)$$

Where $\boldsymbol{\mu}_m^0$ is the standardized mean response at each dose for a given candidate dose response model (m) and $\hat{\boldsymbol{S}}$ denotes the estimated variance-covariance of $\hat{\boldsymbol{\mu}}_d$ [Pinheiro 2014] (i.e., the estimated mean response per individual dose level from the observed data; further details below).

In the calculation of the optimal contrasts (one for each candidate dose response model $\boldsymbol{c}^{opt}{}_{1,\ldots,}\ \boldsymbol{c}^{opt}{}_4$), $\boldsymbol{\mu}_m^0$ relies on the standardized form of the dose response model (not the complete functional form).

For convenience, the contrast coefficients are normalized such that $\|\boldsymbol{c}^{opt}\| = 1$.

Once data are available, the optimal contrasts are applied to $\hat{\boldsymbol{\mu}}_d$.

For each candidate dose response model (m), the null hypothesis H_0: $(\boldsymbol{c}^{opt}{}_m)'\boldsymbol{\mu} = 0$ (absence of a dose-response relationship) will be tested against the two-sided alternative H_1: $(\boldsymbol{c}^{opt}{}_m)'\boldsymbol{\mu} \neq 0$ (existence of a dose response). $\boldsymbol{\mu}$ is the true, but unknown vector of means.

The t-test statistic associated with each of the candidate dose response models is given by $T_m = \dfrac{(\boldsymbol{c}_m^{opt})'\hat{\boldsymbol{\mu}}}{\left[\boldsymbol{C}'\hat{\boldsymbol{S}}\boldsymbol{C}\right]_{m,m}^{1/2}}$ with $[A]_{m,m}$ denoting the m^{th} diagonal element of the matrix A, and $\boldsymbol{C} = [\boldsymbol{c}^{opt}{}_{1,\ldots,}\ \boldsymbol{c}^{opt}{}_4]$ representing the matrix containing the optimal contrasts.

The test statistic used for establishing an overall dose-response signal will be the maximum of the individual model contrast test statistics $T_{max} = \max_m(T_m)$, or alternatively, the associated multiplicity adjusted *p*-value associated with T_{max}.

Analysis will only proceed to the modeling (Mod) step if a statistically significant dose-response signal is established, i.e., the null hypothesis is rejected under any candidate dose response model (adjusted two-sided *p*-value < 0.05).

7.3.2.2 Implementation of the Mod step

The Mod step involves model selection, fitting the univariate dose response profile, and estimating TD(s) based on the fitted model.

A two-stage approach will be used to fit the dose-response model to the observed data. This approach utilizes generalized least squares (LS; [Pinheiro 2014]) and relies on asymptotic results, but has the appeal of being a general purpose application as it depends only on $\hat{\boldsymbol{\mu}}_d$ and $\hat{\boldsymbol{S}}$. For example, the obtained (adjusted) estimates for $\hat{\boldsymbol{\mu}}_d$ and $\hat{\boldsymbol{S}}$ inherit the missing data assumptions of the underlying first stage model.

First stage

An analysis of covariance (ANCOVA) model (first stage model) will be used to calculate estimates for $\hat{\boldsymbol{\mu}}_d$ (i.e., LS mean estimates for each ACT-541468 dose level and placebo) and $\hat{\boldsymbol{S}}$ (i.e., variance-covariance matrix of $\hat{\boldsymbol{\mu}}_d$).

The change from baseline in WASO (assumed to be normally distributed) will be modeled using a fixed effect for treatment group (placebo and each ACT-541468 dose level), an additional fixed effect for gender, and a linear slope to adjust for WASO at baseline.

The variance of the change from baseline in WASO is assumed to be constant across each treatment group.

The LS mean for each treatment group 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.

Second stage

Among the candidate dose response models with a significant contrast test (i.e., adjusted two-sided *p*-value < 0.05) identified in the MCP step, the dose response model parameters (e.g., ED_{50}) are estimated by fitting the dose response model (family) to the estimates obtained from the mixed effects model ($\hat{\boldsymbol{\mu}}_d$ and $\hat{\boldsymbol{S}}$) in the first stage. The best fitting dose response model (E_{max} or linear) will be selected based on the gAIC (smaller is better). This chosen dose response model will then be used to estimate the TD.

Based on the selected dose response model, the predicted mean (95% CI) dose response profile for the change from baseline in WASO will be graphically displayed.

7.3.2.3 Dose selection based on WASO

Estimation of a TD will be performed on the basis of the fitted dose response model (E_{max} or linear) chosen in the second stage. The estimated TD will be the lowest dose strength anywhere within the studied dose range (0 mg to 50 mg) which shows a clinically meaningful effect (i.e., 15 minutes better than predicted mean value of placebo) and a statistically significant effect (i.e., lower 95% CI of the predicted mean value at TD > observed mean value of placebo). Note: the choice of 95% CI is not driven by the purpose of controlling type I error rates.

To aid understanding of the accuracy/variation associated with the selected dose response model, the following parameters of interest will be derived using a nonparametric bootstrap technique.

- The 0.025, 0.5 (median) and 0.975 bootstrap percentiles of the predicted mean (on the change in reduction in WASO of ACT-541468 relative to placebo) for the chosen dose response model,
- The 0.025, 0.5 (median) and 0.975 bootstrap percentiles of the TD defined as 15 minutes better than predicted mean value of placebo,
- The 0.025, 0.5 (median) and 0.975 bootstrap percentiles of the TD defined as 20 minutes better than predicted mean value of placebo,
- If the E_{max} model is selected, the 0.025, 0.5 and 0.975 bootstrap percentiles of the E_0, ED_{50} and E_{max} will be presented.

The nonparametric bootstrap technique will proceed as follows:

1. Sample, with replacement, the data available at each period per subject until the resampled dataset contains the same number of subjects as in the original dataset,
2. From the resampled dataset, calculate and store the parameter estimates that define the dose response model,
3. Repeat steps 1 and 2 numerous times (say 1000) to derive the bootstrap distribution for each parameter,
4. Use the bootstrap distribution to calculate the parameters of interest described above.

7.3.3 Other

The primary endpoint will also be summarized by gender and overall using descriptive statistics.

7.4 Supportive/sensitivity analyses

Sensitivity analyses will be conducted on the PPS if this analysis set differs from the mFAS significantly (e.g., > 10% reduction in number of subjects).

Change from baseline to Days 1&2 in mean WASO will also be analyzed using an ANCOVA model with the same factors as in the main analysis with the addition of region, race and age.

Each of the four doses of ACT-541468 (5, 10, 25 and 50 mg) will be compared to placebo using the single-step Dunnett procedure. For each comparison, the placebo-adjusted LS mean will be displayed along with associated standard error, 95% CI and multiplicity adjusted two-sided *p*-value.

The active comparator (zolpidem) will be compared separately versus placebo using an ANCOVA model with a factor for treatment group (placebo or active comparator) and a covariate for Baseline mean WASO and gender. This analysis is considered exploratory and no formal multiplicity adjustment will be made.

8 ANALYSIS OF THE SECONDARY EFFICACY VARIABLES

The analysis of the secondary efficacy variables will be performed on the FAS.

(Also applicable to Section 9.) Questionnaire data based on the nights in the sleep laboratory (i.e., the mornings following any PSG night) will be excluded from the calculation of the weekly averages since:

- Data may be influenced by the subject being at the sleep laboratory compared to the outpatient environment,
- Time in bed (TIB) is limited to 8 hours in the sleep laboratory.

(Also applicable to Section 9.) For objective endpoints, subjects must have at least one value at a given time point (e.g., Days 1&2) to calculate a mean. For subjective endpoints, subjects must have at least 3 days of data during each week (time windows defined in Section 13.8) to calculate a weekly mean. Otherwise, the mean value will be considered missing for that time point or week.

The same MCP-Mod analysis as performed for the primary endpoint will be conducted to evaluate the dose response of ACT-541468 on each of the secondary endpoints (separately) defined below. Three E_{max} models with an ED_{50} of 1.11, 3.75, 15, respectively, and one linear model will be considered. Summary statistics will also be provided.

The active comparator (zolpidem) will be compared separately versus placebo using an ANCOVA model with a factor for treatment group (placebo or active comparator) and a covariate for Baseline mean WASO and gender.

8.1 Subjective sleep maintenance

Subjective sleep maintenance is assessed through the absolute change in sWASO from Baseline to Week 4.

sWASO is the self-reported time (min) spent awake after sleep onset as reported in the sleep diary.

8.2 Objective sleep onset

Objective sleep onset is assessed through the absolute change in LPS from Baseline to Days 1&2.

Baseline is the mean of the two PSG nights during the run-in period (V2), and Days 1&2 is the mean of the corresponding two PSG treatment nights (V3).

LPS is the time (min) from start of recording to the beginning of the first continuous 20 epochs (i.e., 10 min) scored as non-wake, 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.

8.3 Subjective sleep onset

Subjective sleep onset is assessed through the absolute change in sLSO from Baseline to Week 4.

sLSO is the self-reported time (min) to fall asleep as reported in the sleep diary.

9 ANALYSIS OF OTHER EFFICACY VARIABLES

The analysis of other efficacy variables will be performed on the FAS.

For endpoints based on objective assessments (e.g., WASO, LPS and Total Sleep Time [TST] as determined by PSG):

- Baseline is the mean of the two PSG nights during the run-in period (V2),
- Days 1&2 is the mean of the corresponding two PSG treatment nights at V3,
- Days 15&16 is the mean of the corresponding two PSG treatment nights at V4,
- Days 28&29 is the mean of the corresponding two PSG treatment nights at V5.

For endpoints based on subjective assessments (e.g., sWASO, sLSO, subjective TST [sTST], Sleep Quality [SQ]):

- Weeks 1&2&3&4 are the mean values of each week separately based on the sleep diary entries at home while on double-blind study treatment during Week 1, Week 2, Week 3 and Week 4 as defined in Section 13.8.

9.1 Objective and subjective sleep maintenance

Summary statistics will be provided for:

- Absolute values of WASO at Baseline, Days 1&2, Days 15&16 and Days 28&29 by gender and overall,
- Absolute change in WASO from Baseline to Days 15&16, and to Days 28&29 by gender and overall,
- WASO over time (by hour of the night and by quarter of the night),
- Absolute values of sWASO at Baseline and Weeks 1&2&3&4 by gender and overall,
- Absolute change in sWASO from Baseline to Weeks 1&2&3 by gender and overall.

9.2 Objective and subjective sleep onset

Summary statistics will be provided for:

- Absolute values of LPS at Baseline, Days 1&2, Days 15&16 and Days 28&29 by gender and overall,
- Absolute change in LPS from Baseline to Days 15&16, and to Days 28&29 by gender and overall,
- Absolute values of sLSO at Baseline and Weeks 1&2&3&4 by gender and overall,
- Absolute change in sLSO from Baseline to Weeks 1&2&3 by gender and overall.

9.3 Objective and subjective sleep onset and sleep maintenance

The effect of ACT-541468 treatment versus placebo on the changes from baseline over time (i.e., repeated measures) in WASO, sWASO, LPS, and sLSO (separately) will be estimated by fitting a linear mixed model adjusted for baseline, with treatment group, time point, treatment group by time point interaction and gender as fixed effects, and subject as a random effect. A compound symmetry structure will be considered since the correlation of the changes from baseline over time is assumed to be constant and similar for each treatment group.

9.4 Objective and subjective Total Sleep Time

TST is the amount of actual sleep time (min), i.e., time spent in epochs scored as non-wake as determined by PSG.

sTST is the self-reported time (min) spent asleep as reported in the sleep diary.

Summary statistics will be provided for:

- Absolute values of TST at Baseline, Days 1&2, Days 15&16 and Days 28&29 by gender and overall,
- Absolute change in TST from Baseline to Days 1&2, Days 15&16, and Days 28&29 by gender and overall,
- Absolute values of sTST at Baseline and Weeks 1&2&3&4 by gender and overall,
- Absolute change in sTST from Baseline to Weeks 1&2&3&4 by gender and overall.

9.5 Sleep Quality

SQ is the sleep quality as determined by scores on the Visual Analog Scale (VAS; mm).

Summary statistics will be provided for:

- Absolute values of SQ at Baseline and Weeks 1&2&3&4,
- Absolute change in SQ from Baseline to Weeks 1&2&3&4.

9.6 Sleep architecture

Absolute change from Baseline to Days 1&2, Days 15&16, and Days 28&29 in the duration (min) and percentage of TST in each sleep stage (S1, S2, SWS and REM) will be summarized over the whole night, and for each quarter of the night. Absolute values will also be summarized.

Absolute change from Baseline to Days 1&2, Days 15&16, and Days 28&29 in the latency (min) to each sleep stage (S1, S2, SWS, and REM) will be summarized. Absolute values will also be summarized.

9.7 Objective and subjective sleep continuity

Mean values are the mean of the two PSG treatment nights at each visit.

Absolute change from Baseline to Days 1&2, Days 15&16, and Days 28&29 in mean number of shifts from S2, SWS or REM to S1 or wake for the whole night will be summarized. Absolute values will also be summarized.

Absolute change from Baseline to Days 1&2, Days 15&16, and Days 28&29 in mean wake time during sleep (defined as the time [min] spent in epochs scored as wake between LPS and last epoch not scored wake) for the whole night will be summarized. Absolute values will also be summarized.

Absolute change from Baseline to Days 1&2, Days 15&16, and Days 28&29 in mean number of awakenings (NAW; defined as the NAW between first epoch and last epoch not scored wake) will be summarized for the whole night, for each quarter of the night, and for each hour of the night. Absolute values will also be summarized.

Absolute change from Baseline to Weeks 1&2&3&4 in the mean self-reported NAW will be summarized. Absolute values will also be summarized. Baseline and Weeks 1&2&3&4 are defined in Section 9.

9.8 Sleep efficiency

Sleep efficiency (%) is defined as 100 × (TST [min] / total TIB [min]) where total TIB is fixed to 480 min during the PSG nights.

Absolute change from Baseline to Days 1&2, Days 15&16 and Days 28&29 in sleep efficiency will be summarized. Absolute values will also be summarized.

9.9 Insomnia severity

The Insomnia Severity Index (ISI©) is a validated instrument to quantify perceived insomnia severity [Bastien 2001]. The ISI is composed of seven items that evaluate (1) severity of sleep-onset (initial), (2) sleep maintenance (middle), (3) early morning awaking (terminal) problems, (4) satisfaction with current sleep pattern, (5) interference with daily functioning, (6) noticeability of impairment attributed to the sleep problem, and (7) level of distress caused by the sleep problem. Each item is scored as 0 (not at all), 1, 2, 3 or 4 (extremely) and the time interval considered is the last 2 weeks. Total scores range from 0–28 with high scores indicating greater insomnia.

Absolute change from Baseline (V1) to the second morning of V5 (Day 30) in ISI scores. Absolute values will also be summarized.

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

- ≥ 6 point reduction in ISI.

9.10 Next-day performance

Absolute change from Baseline to Weeks 1&2&3&4 in next-day performance assessed at home by scores on the VAS (mm) assessing morning sleepiness, daytime alertness and daytime ability to function. Absolute values will also be summarized.

10 SAFETY ANALYSIS

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

Unless noted otherwise, the safety summary tables and figures will include assessments/events occurring during the on-treatment period (i.e., from day of first dose of double-blind study treatment to EOS), also defined as being treatment-emergent.

All safety data will be listed, and those collected during the post-treatment period will be flagged.

10.1 Adverse events

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

AEs will be summarized (frequency counts and percentages) by SOC and/or PT, and maximum intensity.

The following AE summaries will be provided:

- Treatment-emergent AEs,
- AEs emerging during the run-in period [see Section 13.2],
- AEs emerging during the withdrawal period [see Section 13.2],
- Treatment-emergent AEs related to study treatment,
- Treatment-emergent AEs leading to premature discontinuation of double-blind study treatment,
- Treatment-emergent AEs leading to temporary interruption of double-blind study treatment,
- Occurrence of treatment-emergent non-serious AEs,
- Treatment-emergent serious AEs (SAEs),
- Occurrence of treatment-emergent SAEs,
- Treatment-emergent SAEs related to study treatment,
- Treatment-emergent AEs with fatal outcome,
- Treatment-emergent AEs related to study treatment with fatal outcome,
- On-treatment deaths with cause of death.

All AEs will be listed as well as all deaths with cause of death. In addition, separate listings will be provided for AEs leading to discontinuation of double-blind study treatment, and SAEs.

Treatment-emergent AEs of special interest (AESI) both according to investigator (based on the question 'Do you consider this AE an Adverse Event of Special Interest (AESI)?' recorded as 'Yes' on the 'Adverse Event' page), and after adjudication by an Independent Safety Board (ISB), will be summarized. The following summaries will be provided:

- Treatment-emergent AESI as per investigator,
- Treatment-emergent AESI after ISB adjudication.

AESI will be listed separately.

A subject with multiple intensities reported for an AE will be summarized under the maximum intensity recorded for the event. A subject with multiple occurrences of an AE is counted only once in the AE category (e.g., system organ class, preferred term).

An incomplete (day or month missing) or missing AE date will be imputed as described in the table below. 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 2 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	If the (imputed) AE resolution date is on or after the start of double-blind study treatment, and if the double-blind study treatment start falls within the upper and lower limits (inclusive), the double-blind study treatment start date is used. If the (imputed) AE resolution 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. If the AE resolution date is missing then: 1. if the double-blind study treatment start falls within the upper and lower limits (inclusive), the double-blind study treatment start date is used else: 2. 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. In all the other cases, the lower limit is used.	Whichever is the earlier of the AE resolution date or double-blind study treatment start date.

AE = Adverse event

10.2 Laboratory data

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 unscheduled assessments.

Descriptive summary statistics by scheduled visit will be provided for observed values and absolute changes from baseline in both hematology and blood chemistry laboratory parameters. Only scheduled assessments will be considered for this summary.

The absolute change from baseline to last value in the double-blind treatment period for hematology and blood chemistry parameters will be summarized.

Marked laboratory abnormalities are defined in table below. The number (%) of subjects with treatment-emergent [defined in Section 10] marked laboratory abnormalities 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: those not meeting the criterion at baseline (or having a missing baseline value) and having at least one post-baseline value for a given parameter.

All laboratory data for subjects with at least one treatment-emergent marked laboratory abnormality will be listed.

Table 3 Marked abnormalities in laboratory parameters for reporting

Laboratory parameter	Criteria for marked laboratory abnormalities
Hematology	
Hemoglobin (g/L)	< 100 (LL) < 80 (LLL) > 20 above ULN or > 20 above baseline if baseline > ULN (HH) > 40 above ULN or > 40 above baseline if baseline > ULN (HHH)
Hematocrit (%)	< 32% (Male); < 28% (Female) (LL) < 20% (LLL) > 60% (Male); > 55% (Female) (HH) > 65% (HHH)
Platelets (10^9/L)	< 75 (LL) < 50 (LLL) > 600 (HH) > 999 (HHH)
Leucocytes (10^9/L)	< 3.0 (LL) < 2.0 (LLL) > 20.0 (HH) > 100.0 (HHH)
Neutrophils (10^9/L)	< 1.5 (LL) < 1.0 (LLL)
Eosinophils (10^9/L)	> 5.0 (HH)
Eosinophils (%)	> 5% (HH)
Lymphocytes (10^9/L)	< 0.8 (LL) < 0.5 (LLL) > 4.0 (HH) > 20.0 (HHH)
WBC count (10^9/L)	< 3.0 (LL) < 2.0 (LLL) > 20.0 (HH) > 100.0 (HHH)

Laboratory parameter	Criteria for marked laboratory abnormalities
Prothrombin INR	> 1.5 × ULN (HH) > 2.5 × ULN (HHH)
Reticulocyte (%)	> 2.5% (HH)
Blood chemistry	
ALT (U/L)	> 3 × ULN (HH) > 5 × ULN (HHH) > 10 × ULN (HHHH)
AST (U/L)	> 3 × ULN (HH) > 5 × ULN (HHH) > 10 × ULN (HHHH)
ALP (U/L)	> 2.5 × ULN (HH) > 5 × ULN (HHH)
TBIL (µmol/L)	> 2 × ULN (HH) > 5 × ULN (HHH)
Creatinine (µmol/L)	> 1.5 × ULN or > 1.5 × baseline if baseline > ULN (HH) > 3 × ULN or > 3 × baseline if baseline > ULN (HHH)
Albumin (g/L)	< 30 (LL) < 20 (LLL)
Calcium (mmol/L)	< 2.0 (LL) < 1.75 (LLL) > 2.9 (HH) > 3.1 (HHH)
Potassium (mmol/L)	< 3.2 (LL) < 3.0 (LLL) > 5.5 (HH) > 6.0 (HHH)
Sodium (mmol/L)	< 130 (LL) > 150 (HH) > 155 (HHH)
Chloride (mmol/L)	< 74 (LL) > 131 (HH)
Creatine kinase (µg/L)	> 5 × ULN (HH) > 10 × ULN (HHH)
Gamma-glutamyl transferase (U/L)	> 2.5 × ULN (HH) > 5 × ULN (HHH)
Blood urea nitrogen (mmol/L)	> 2.5 × ULN (HH) > 5 × ULN (HHH)
Uric acid (µmol/L)	> 590 (HH) > 720 (HHH)

Laboratory parameter	Criteria for marked laboratory abnormalities
Glucose (mmol/L)	< 3 (LL)
	< 2.2 (LLL)
	> 8.9 (HH)
	> 13.9 (HHH)
TSH (mIU/L)	< 0.28 (LL)
	> 6.6 (HH)
Free T_3 (pmol/L)	< 2.8 (LL)
	> 7.8 (HH)
Total T_3 (nmol/L)	< 0.9 (LL)
	> 3.0 (HH)
Free T_4 (pmol/L)	< 7.2 (LL)
	> 30 (HH)
Total T_4 (nmol/L)	< 45 (LL)
	> 160 (HH)

ALP = Alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; INR = International Normalized Ratio; TBIL = Total bilirubin; TSH = thyroid-stimulating hormone; T_3 = thyroxine; T_4 = triiodothyronine; ULN = upper limit of normal; WBC = white blood cell. LL, LLL, HH, HHH, HHHH = flags for low and high marked abnormalities.

Elevated liver parameters during and up to 30 days after the EODB will be summarized: the number (%) of subjects meeting the criteria defined below within a given central laboratory sample will be tabulated by treatment group, including all ACT-541468 dose levels combined. Percentages will be based on the number of subjects at risk: those not meeting the criterion at baseline (or having a missing baseline value) and having at least one post-baseline value for a given parameter.

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

10.3 Vital signs, weight and physical examination

Each summary and listing will include unscheduled assessments. Assessments from supine and sitting position will be presented together.

Absolute changes from baseline (mean of the two PSG nights at V2, run-in period) to last value on double-blind study treatment in vital signs (systolic and diastolic BP, pulse rate

and body temperature) will be summarized. The absolute values at baseline and last value on double-blind study treatment will also be summarized.

The changes from baseline (V1 or last value before double-blind treatment) to V5 (or last value on or following double-blind study treatment) in body weight will be summarized.

Physical examination abnormalities will be listed.

10.4 Electrocardiograms

Each electrocardiogram (ECG) summary and listing will include unscheduled assessments.

The absolute change from baseline to last value in the double-blind treatment period for ECG parameters (QT corrected according to Fridericia's formula, QT corrected according to Bazett's formula, heart rate, PR, QRS) will be summarized. The absolute values at baseline and last value on double-blind study treatment will also be summarized.

Marked ECG abnormalities are defined in the table below.

The following summaries will be provided for each ECG parameter:

- Number (%) of subjects having a marked ECG abnormality during the double-blind treatment period,
- Number (%) of subjects having a marked ECG abnormality during the withdrawal period [see Section 13.2].

Percentages will be based on the number of subjects at risk: those not meeting the criterion at baseline (or having a missing baseline value) and having at least one post-baseline value for a given parameter.

All ECG values for subjects with at least one marked ECG abnormality will be listed.

Table 4 Marked abnormalities in ECG parameters

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

HR = Heart rate; QTcB = QT corrected according to Bazett's formula; QTcF = QT corrected according to Fridericia's formula; LL, LLL, H, HH, HHH = flags for low and high marked abnormalities.

10.5 Withdrawal symptoms

The following analyses will be based on the Withdrawal set.

The Benzodiazepine Withdrawal Symptom Questionnaire (BWSQ) consists of 20 items. Response to each item is scored as 'No' = 0, 'Yes - moderate' = 1, 'Yes - severe' = 2 meaning the BWSQ total score can range from 0 to 40.

The BWSQ total score will be summarized using descriptive statistics for the absolute values and changes from V5 (2nd morning) to V6 (morning), and changes from baseline to each scheduled visit.

The generalized MCP-Mod methodology will be used to explore the dose response relationship for the change from V5 to V6 in BWSQ total score. One linear model and various sigmoid E_{max} dose response models covering a wide range of monotonic shapes will be used to determine the different sets of testing contrasts. An ANCOVA model (using a fixed effect for treatment group [placebo and each ACT-541468 dose level], and a linear slope to adjust for BWSQ total score at V5) will be used to calculate estimates for $\hat{\boldsymbol{\mu}}_d$ (i.e., LS mean estimates for each ACT-541468 dose level and placebo) and $\hat{\boldsymbol{S}}$ (i.e., variance-covariance matrix of $\hat{\boldsymbol{\mu}}_d$). The LS means for each treatment group will be displayed along with associated standard errors and 95% CIs. For each ACT-541468 dose level comparison to placebo, the placebo-adjusted LS means will be displayed along with associated standard error, 95% CI and unadjusted two-sided *p*-value. The best fitting dose response model (linear or sigmoid E_{max}) will be selected based on the gAIC (smaller the better). For the selected dose response model, the predicted mean (95% CI) dose response profile for the change from V5 to V6 in BWSQ total score will be graphically displayed.

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 double-blind study treatment discontinuation will be assessed through the incidence of AEs, and marked ECG abnormalities, occurring during the withdrawal period [see Section 13.2]. As described in their respective section, the incidence of AEs [Section 10.1] and marked ECG abnormalities [Section 10.4] occurring during the withdrawal period (between V5 and V6) will be summarized.

The following analyses will be based on the SS.

The generalized MCP-Mod methodology will be used to explore the dose response relationship on the incidence of AEs, and marked ECG abnormalities, separately, occurring during the withdrawal period. One linear model and various sigmoid E_{max} dose response models covering a wide range of monotonic shapes will be used to determine the different sets of testing contrasts. A logistic regression model using a fixed effect for treatment group (placebo and each ACT-541468 dose level) will be used to calculate estimates for $\hat{\boldsymbol{\mu}}_d$ (i.e., LS mean estimates for each ACT-541468 dose level and placebo on the logit scale) and $\hat{\boldsymbol{S}}$ (i.e., variance-covariance matrix of $\hat{\boldsymbol{\mu}}_d$). The best fitting dose response model (linear or sigmoid E_{max}) will be selected based on the gAIC (smaller the better). For the selected dose response model, the predicted mean (95% CI) dose response profile (on the logit scale) for incidence of AEs, and marked ECG abnormalities, occurring during the withdrawal period will be graphically displayed.

10.6 Insomnia rebound effect

The following will be based on the Withdrawal set.

In order to assess the potential rebound insomnia, descriptive statistics are provided for the change from the worst PSG screening night (V2) to the first value after double-blind treatment (V6) in WASO, TST, and LPS.

The generalized MCP-Mod methodology will be used to explore the dose response relationship for the occurrence of worsening from the worst PSG screening night (V2) to V6 in WASO, TST, and LPS, separately. One linear model and various sigmoid E_{max} dose response models covering a wide range of monotonic shapes will used to determine the different sets of testing contrasts. A logistic regression model using a fixed effect for treatment group (placebo and each ACT-541468 dose level) will be used to calculate estimates for $\hat{\boldsymbol{\mu}}_d$ (i.e., LS mean estimates for each ACT-541468 dose level and placebo on the logit scale) and $\hat{\boldsymbol{S}}$ (i.e., variance-covariance matrix of $\hat{\boldsymbol{\mu}}_d$). The best fitting dose response model (linear or sigmoid E_{max}) will be selected based on the gAIC (smaller the better). For the selected dose response model, the predicted mean (95% CI) dose response

profile (on the logit scale) for the occurrence of worsening from the worst PSG screening night (V2) to V6 in WASO, TST, and LPS will be graphically displayed.

An occurrence of worsening in WASO/LPS is when the value at V6 is greater than the worst (i.e., highest) PSG screening night (V2) value. An occurrence of worsening in TST is when the value at V6 is less than the worst (i.e., lowest) PSG screening night (V2) value.

10.7 Other safety parameters of interest

10.7.1 Digit Symbol Substitution Test©

The Digit Symbol Substitution Test (DSST)© is a measure of attention, perceptual speed, motor speed, visual scanning and memory. The total DSST score is the number of correct symbols entered.

Absolute change from baseline (mean of the two PSG morning assessments at V2, run-in period) to Days 1&2, 15&16 and 28&29 (mean of the two PSG morning assessments at V3, V4 and V5, respectively) in total DSST score will be summarized by gender and overall. Absolute values will also be summarized.

10.7.2 Karolinska Sleepiness Scale

The Karolinska Sleepiness Scale (KSS) is a 9-point scale. The scale has a range from 1 being 'very alert' to 9 being 'very sleepy, great effort to keep awake, fighting sleep'.

Absolute change from baseline (mean of the two PSG morning assessments at V2, run-in period) to Days 1&2, 15&16 and 28&29 (mean of the two PSG morning assessments at V3, V4 and V5, respectively) in KSS will be summarized by gender and overall. Absolute values will also be summarized.

10.7.3 Sheehan Disability Scale©

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).

Absolute change from baseline (mean of the two PSG morning assessments at V2, run-in period) to Days 1&2, 15&16 and 28&29 (mean of the two PSG morning assessments at V3, V4 and V5, respectively) in SDS will be summarized by gender and overall. Absolute values will also be summarized.

10.7.4 Neurological Examination

The number (%) of subjects answering 'Yes' to the question 'Is excessive sleepiness observed one hour after lights on?' on the 'Neurological Examinations' page will be tabulated for each scheduled visit by gender and overall.

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

10.7.5 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.

Number (%) of subjects with suicidal ideation, suicidal behavior, and/or self-injurious behavior without suicidal intent based on the C-SSRS during double-blind treatment will be tabulated. Percentages will be based on the number of subjects with at least one post-baseline C-SSRS assessment.

Shifts from baseline showing any changes in suicidal ideation and suicidal behavior during double-blind treatment will also be provided. 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 only, and 3) Suicidal ideation and behavior. 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).

10.8 Pharmacokinetic

PK analyses will be performed based on the PK set.

Descriptive statistics (n, mean, standard deviation, CV%, m [number of non-zero concentrations], geometric mean, geometric standard deviation, geometric CV%, median,

minimum and maximum) of ACT-541468 plasma concentration collected approximately 9–10 h post-dose at V3, V4, and V5 will be provided for each time point (morning after the first and second PSG nights), by gender and overall. Note that the protocol recommends to draw a PK sample in the morning after the first PSG night (at V3, V4, and V5) only in subjects where excessive sleepiness is observed (according to investigator opinion) one hour after lights on.

Concentration values below the lower limit of quantification (< 1.0 ng/mL) will be displayed in listings as zero with a flag and handled as zero in the calculations for mean, CV%, standard deviation, median, minimum, and maximum, but handled as missing for the calculation of the geometric mean, geometric standard deviation and geometric CV%.

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

10.9 Exposure-safety analysis

The exposure-safety relationship will be explored using $C_{9\text{-}10h}$ (plasma concentrations of ACT-541468 in the morning of the second PSG night at V3, V4, and V5) and will be based on the PK set. Safety parameters considered for this analysis will include selected AEs (e.g., somnolence and/or excessive sleepiness), changes from baseline in DSST, KSS, SDS, and morning sleepiness (VAS for 'the way you feel this morning'). $C_{9\text{-}10h}$ will be assessed against each safety parameter using graphical presentations and/or statistical models with dose, time, gender, age and BMI as possible factors.

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 IDMC and ISB charter, respectively. Safety and efficacy data supporting the review by IDMC will be provided by Actelion for the part of analyses that are blinded and by an ISAC for the unblinded part.

12 CHANGES OR CLARIFICATIONS TO ANALYSES PLANNED IN THE STUDY PROTOCOL

12.1 Changes to the analyses planned in the study protocol

The AEs and SAEs with onset between randomization and first day of study treatment will not be summarized. However, AEs emerging during the run-in period will be summarized instead.

The change from baseline (second PSG morning at V2) to the last value on double-blind study treatment in C-SSRS© will not be summarized. However, the changes (shifts) from baseline to worst reported outcome in suicidal ideation and suicidal behavior will be summarized instead.

Analyses of secondary and exploratory efficacy endpoints will be based on the FAS (a broader analysis set) instead of the modified FAS which is linked only to the primary endpoint by requiring at least one WASO assessment at baseline and one at Days 1&2.

12.2 Changes in the conduct of the study / data collection

Not applicable.

12.3 Clarifications concerning endpoint definitions and related variables or statistical methods

In case not all subjects in the SS participate in the withdrawal period, a withdrawal set was added as an analysis set to analyze assessments (e.g., withdrawal symptoms and rebound effect) taken during the withdrawal period.

13 GENERAL DEFINITIONS AND DERIVATIONS

13.1 Treatment start and end dates

Run-in single-blind placebo treatment start date is the earliest date of dose intake (i.e., date of capsules dispensed) recorded on the 'Study Single Blind Treatment Dispensing & Accountability' page at V2.

Double-blind study treatment start (end) date is the earliest (latest) 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 date of dose intake (i.e., date of capsules dispensed) recorded on the 'Study Single Blind Treatment Dispensing & Accountability' page at V6.

13.2 Study periods

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 double-blind study treatment.

The withdrawal period is defined as the time from two days after double-blind study treatment end date to one day after run-out single-blind placebo treatment end date.

13.3 Study day

The study day for an assessment or event will be calculated using the start date of double-blind study treatment as reference.

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

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

The first day of double-blind study treatment is study day 1.

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

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

Study day will be displayed in the data listings.

13.4 Baseline

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

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.

13.5 Absolute change from baseline

The absolute change from baseline is defined as post-baseline value minus baseline value. A positive number indicates an increase as compared to baseline.

13.6 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 and NAW is set as missing since subject is never at risk of either, and
- LPS is set as 480 minutes which is equal to the duration of TIB.

If the sTST is recorded as zero, then similarly:

- sWASO and subjective NAW (sNAW) is set as missing,
- sLSO is set as 480 minutes.

13.7 Handling of missing/incomplete date and time fields

This section describes some general principles to be followed in the case of missing or incomplete date/time.

Missing or incomplete dates are handled as follows:

- Dates are split into 3 parts: year, month and day. Year is the top-level, month is medium level and day is low level. If a part that is expected to contain a number is numeric but the value is outside a valid range, the complete date is handled as missing. For example, if date = 44Nov2000 the whole date is considered to be missing.
- If a part that is expected to contain a number is not numeric, i.e., contains values such as ND, NA, --, ??, 2?, that part is considered to be missing.
- If a part is missing, all lower level parts are considered to be missing. This means that a ddmmyy date ‘21ND99’ is considered as ‘----99’.

- If year is missing, then whole date is considered to be missing, unless reasonable assumptions support an imputation (imputation rule and rationale for assumption will be provided).

Imputation of missing or incomplete dates/times are further specified in the associated analysis section. Unless noted otherwise, the non-imputed date/time will be reported in listings.

13.8 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 (V2, V3, V4 or V5) are shown in the table below [Table 5].

Table 5 Meaning of the first and last PSG end date at visit V2, V3, V4 or V5

V2, V3, V4 or V5			
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:

Screening: the 7 consecutive days immediately preceding the first PSG at V2 (i.e., 7–day window = [End date of first PSG at V2 – 7 days; End date of first PSG at V2 – 1 day]).

Baseline (run-in period): the 7 consecutive days immediately preceding the first PSG at V3, but after V2 (i.e., [Max(End date of first PSG at V3 – 7 days; End date of last PSG at V2); End date of first PSG at V3 – 1 day]). The time window for Baseline may be less than 7 days. Baseline will be missing for any subjects without a PSG at V3.

Week 1: the remaining days between Baseline and Week 2 time windows (i.e., [End date of last PSG at V3 + 1 day; End date of first PSG at V4 – 8 days]). The time window for Week 1 may be greater or less than 7 days.

Week 2: the 7 consecutive days immediately preceding the first PSG at V4 (i.e., 7–day window = [End date of first PSG at V4 – 7 days; End date of first PSG at V4 – 1 day]). In the unlikely event that this time window overlaps with a PSG at V3 (i.e., time window is less than 7 days), then this time window will begin from the end date of the last PSG at V3 + 1 day, and Week 1 will be missing. Week 2 will be missing for any subjects who discontinue the study before V4.

Week 3: the remaining days between Week 2 and Week 4 time windows (i.e., [End date of last PSG at V4 + 1 day; End date of first PSG at V5 – 8 days]). The time window for Week 3 may be greater or less than 7 days.

Week 4: the 7 consecutive days immediately preceding the PSG at V5 (i.e., 7–day window = [End date of first PSG at V5 – 7 days; End date of first PSG at V5 – 1 day]). In the unlikely event that this time window overlaps with a PSG at V4 (i.e., time window is less than 7 days), then this time window will begin from the end date of the last PSG at V4 + 1 day, and Week 3 will be missing. Week 4 will be missing for any subjects who discontinue the study before V5.

For any subjects who discontinue the study between V3 and V4 (i.e., without a PSG at V4)

Week 1: all days from the last PSG at V3 until last entry date in sleep diary (i.e., [End date of last PSG at V3 + 1 day; Last entry date in sleep diary]).

For any subjects who discontinue the study between V4 and V5 (i.e., without a PSG at V5)

Week 3: all days from the last PSG at V4 until last entry date in sleep diary (i.e., [End date of last PSG at V4 + 1 day; Last entry date in sleep diary]).

14 REFERENCES

[Bastien 2001] Bastien C, Vallieres A, Morin C. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med. 2001;2:297–307.

[Bretz 2005] Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. Biometrics. 2005;61(3):738–48.

[Bornkamp 2016] Bornkamp B, Pinheiro J, Bretz F. Documentation of the R Package 'DoseFinding'. 2016, Version 0.9–14.

[Pinheiro 2006] Pinheiro J, Bornkamp B, Bretz F. Design and analysis of dose-finding studies combining multiple comparisons and modeling procedures. J Biopharm Stat. 2006;16(5):639–56.

[Pinheiro 2014] Pinheiro J, Bornkamp B, Glimm E, Bretz F. Model-based dose finding under model uncertainty using general parametric models. Statistics in Medicine 2014;33(10):1646–61.

DOCUMENT HISTORY

Version	Effective Date	Reason
Final 1.0	19 December 2016	Initial release
Final 2.0	3 July 2017	Comments from dry-run review