A Phase 3 Study to Determine the Antipsychotic Efficacy and Safety of ALKS 3831 in Adult Subjects with Acute Exacerbation of Schizophrenia

Unique Protocol ID:	ALK3831-A305
NCT Number:	NCT02634346
EudraCT Number:	2015-003373-15
Date of Statistical Analysis Plan:	23 June 2017

STATISTICAL ANALYSIS PLAN

ALK3831-A305

Study Title:	A Phase 3 Study to Determine the Antipsychotic Efficacy and Safety of ALKS 3831 in Adult Subjects with Acute Exacerbation of Schizophrenia
Document Status:	Final Version 1.0
Document Date:	23 June 2017
Based on:	ALK3831-A305 Protocol Amendment 1.0 Date: 22 September 2016
Sponsor:	Alkermes, Inc. 852 Winter Street Waltham, MA 02451 USA

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Abbreviation or Term	Explanation or Definition
AE	adverse event
AIMS	Abnormal Involuntary Movement Scale
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical [classification system]
BARS	Barnes Akathisia Rating Scale
BMI	body mass index
CGI-I	Clinical Global Impression - Improvement
CGI-S	Clinical Global Impression - Severity
CI	confidence interval
C-SSRS	Columbia Suicide Severity Rating Scale
ECG	electrocardiogram
eCRF	electronic case report form
EOT	end of treatment
EPS	extra pyramidal symptoms
ET	early termination
FAS	full analysis set
HbA1C	Hemoglobin A1c
IWQOL	Impact of Weight on Quality of Life
LOCF	last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MMRM	mixed model with repeated measurements
MOAS	Modified Overt Aggression Scale
OLZ	olanzapine
PANSS	Positive and Negative Syndrome Scale
PCS	potentially clinically significant
PK	pharmacokinetic
QTcB	QT corrected with Bazett formula
QTcF	QT corrected with Fridericia formula

Abbreviation or Term	Explanation or Definition
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Simpson-Angus Scale
TEAE	treatment-emergent adverse event
ULN	upper limit of normal
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical methods to be used for analyses and presentation of efficacy and safety data for the study ALK3831-A305. This document has been prepared based on Alkermes ALK3831-A305 Study Protocol Amendment 1.0 (dated 22 Sep 2016).

1.1. Study Objectives

1.1.1. Primary objective

The primary objective of this study is to evaluate the antipsychotic efficacy of ALKS 3831 (a fixed-dose combination of olanzapine and samidorphan) in adult subjects with an acute exacerbation of schizophrenia.

1.1.2. Secondary objective

The secondary objective of this study is to evaluate the safety and tolerability of ALKS 3831 in adult subjects with an acute exacerbation of schizophrenia.

1.2 Summary of the Study Design

ALK3831-A305 is a Phase 3, multinational, multicenter, randomized, double-blind, placebo-controlled study in subjects experiencing an acute exacerbation of schizophrenia. Subjects will be screened up to 10 days prior to randomization. Upon completion of screening assessments, subjects meeting eligibility criteria will be admitted to an inpatient unit on Day -1, if they are not already inpatient. Currently prescribed antipsychotics will be discontinued at Screening.

On Day 1 (Visit 2), subjects meeting eligibility criteria will be randomized in a 1:1:1 fashion to one of the following 3 treatment groups: ALKS 3831, olanzapine, or placebo, and will receive double-blind treatment for up to 4 weeks.

On Days 1 and 2, subjects randomized to ALKS 3831 will receive 10/10 [10 mg olanzapine/10 mg samidorphan] and subjects randomized to olanzapine will receive 10 mg. On Day 3, the dose will be increased to 20/10 [20 mg olanzapine/10 mg samidorphan] for subjects randomized to ALKS 3831 or 20 mg for subjects randomized to olanzapine. Following the increase on Day 3, the dose may be decreased to 10/10 (ALKS 3831) or 10 mg (olanzapine) at end of Week 1 (Day 7) or Week 2 (Day 15) if there are tolerability problems based on judgment of the investigator. No further dose adjustments will be allowed Day 15 onward for the remaining 2 weeks of the treatment period. All subjects will receive 1 tablet daily in a blinded manner.

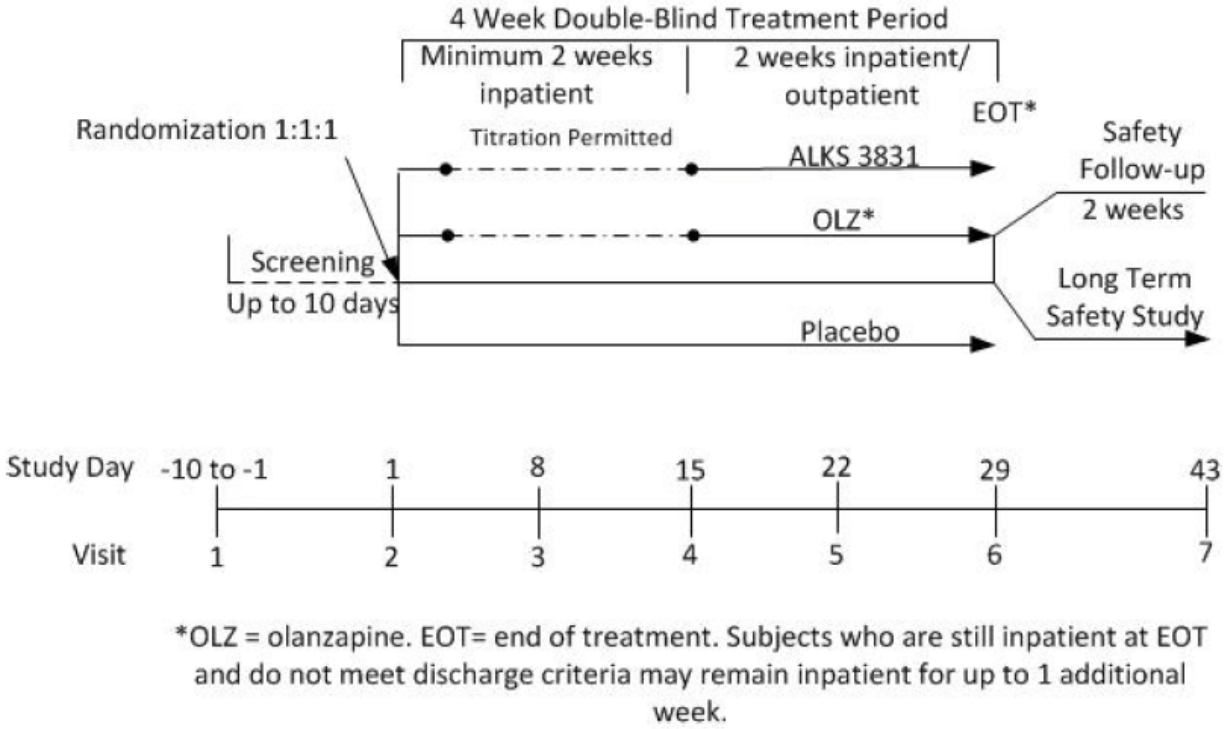
Subjects are required to be inpatient for the first 2 weeks of the treatment period (until Day 15). Following the mandatory 2-week inpatient stay, subjects can either continue the study as inpatients for the full 4-week treatment period or be discharged at the end of Week 2 or end of Week 3 if they meet the discharge criteria specified in Protocol Section 8.1.1.

Visits during the double-blind period will be every week. Each weekly visit will consist of clinical assessments plus drug dispensing.

Subjects completing 4 weeks of treatment with study drug will be eligible to continue in the open-label, long-term safety study (ALK3831-A306) and continue to receive ALKS 3831 for up to 52 weeks. Subjects not continuing in the long term safety study will enter a 2-week Safety Follow-up Period. During this follow-up period subjects may be treated with any antipsychotic medication according to physician recommendations.

Subjects who are inpatient at the end of the 4-week treatment period and are not continuing in the extension study (ALK3831-A306) may continue as inpatients for up to 1 additional week of the 2-week safety follow-up period. These subjects may be discharged from the inpatient unit at any time during this week based on Investigator's judgment. For subjects continuing in the safety extension study (ALK3831-A306), details regarding the additional inpatient week will be provided in the ALK3831-A306 protocol.

Figure 1: Study Design Schematic



*OLZ = olanzapine. EOT= end of treatment. Subjects who are still inpatient at EOT and do not meet discharge criteria may remain inpatient for up to 1 additional week.

2. SAMPLE SIZE CONSIDERATION

The planned sample size is 390 subjects in total, 130 subjects per treatment group. This sample size will provide at least 90% power to show superiority of the ALKS 3831 group compared to placebo group at a 2-sided alpha level of 0.05, assuming a 10-point improvement of PANSS total score at week 4, a standard deviation (SD) of 20, and a dropout rate of 30%.

3. DATA ANALYSIS

3.1. General Statistical Methodology

Baseline for efficacy or safety analysis is defined as the last non-missing efficacy or safety assessment on or before the first dose of study drug in the Double-blind Treatment Period.

In general, descriptive statistics: n, mean (SD), median, minimum, and maximum, for continuous variables and number and percentage of subjects in each category for categorical variables will be provided by treatment group for all variables.

All statistical tests and confidence intervals (CIs), unless stated otherwise, will be 2-sided and will be set at an alpha level of 0.05.

Unless specified otherwise, all formal comparisons will be between ALKS 3831 and placebo groups. An exploratory analysis will be conducted between olanzapine and placebo groups.

All source data will be presented as subject data listings.

3.2. Definitions of Analysis Populations

3.2.1. Safety Population

The Safety Population will include all randomized subjects who receive at least 1 dose of study drug during the double-blind treatment period. The Safety Population will be used for the safety analyses.

3.2.2. Efficacy Population

The full analysis set (FAS) will include all subjects in the Safety Population who have at least one post-baseline PANSS assessment. The FAS will be used for the efficacy analyses.

3.3. Disposition

The number and percentage of subjects completing or prematurely discontinuing the study including reasons for discontinuation will be summarized overall and by treatment group for the following:

- Subjects who were randomized
- Subjects in the Safety Population
- Subjects in the FAS Population
- Subjects who completed the double-blind treatment period
- Subjects who discontinued the study during the double-blind treatment period along with reasons for discontinuation

3.4. Protocol Deviation

Subjects with major protocol deviations using the following categories will be summarized by treatment group along with supportive listings for each category.

- Did not meet the inclusion/exclusion criteria
- Received prohibited medications
- Lack of adherence with study medication, as defined by subjects taking less than 70% of the protocol specified amount of study medication
- Randomization or dosing error

3.5. Demographics and Baseline Characteristics

Demographics and baseline characteristics such as sex, age, race, ethnicity, weight, and body mass index (BMI) will be summarized overall and by treatment group for the Safety and FAS Population.

Baseline efficacy characteristics, eg, PANSS and CGI-S, will be summarized by treatment group for the FAS Population to assess the comparability of the treatment groups.

Medical and psychiatric history will be summarized for the Safety Population by treatment group and overall.

3.6. Prior and Concomitant Medication

Prior medications are defined as medications taken prior to the first dose of study drug. Concomitant medications are defined as medications taken on or after the first dose of study drug. All medications as documented by the investigator will be coded using the World Health Organization Anatomical Therapeutic Chemical (WHO-ATC) drug dictionary Enhanced Extended with Herbal version WHODRUG_C / 161E+H.

Prior and concomitant medications will be summarized by the preferred drug name for the Safety population. Concomitant medications that are taken during the double-blind treatment period will be included in the summary table. For the summary tables, if a subject has taken a prior or concomitant medication more than once, the subject will be counted only once for the medication. All reported medications (including those initiated after the last dose of double-blind study medication) will be included in the listing.

3.7. Treatment Adherence Rate and Extent of Exposure

3.7.1. Treatment Adherence Rate

Treatment adherence to the daily dosing schedule during the double-blind treatment period will be summarized by treatment group for the Safety Population. Treatment adherence will be calculated as follows:

$$100 \times \frac{\text{Total tablets dispensed} - \text{total tablets returned} - \text{total tablets lost}}{\text{Total tablets scheduled to be taken}}$$

3.7.2. Duration of Study Drug Administration

Duration of study drug administration is defined as the number of days from the date of the first dose of study drug taken to the date of the last dose taken, inclusive (ie, last dose date – first dose

date + 1). Duration of study drug administration will be summarized for the Safety Population by treatment group.

The overall prescribed mean and modal dose of olanzapine will be summarized by treatment group for relevant treatment groups. Number and percentage of subjects will be summarized by their final dose level.

3.8. Efficacy Analyses

3.8.1. General Considerations

All statistical tests will be 2-sided hypothesis tests performed at the 5% significance level. All confidence intervals will be 2-sided 95% confidence intervals.

Baseline is defined as the last non-missing observation on or before the first dose of study drug in the double-blind treatment period.

Unless specified otherwise, all formal comparisons will be between ALKS 3831 and placebo groups. An exploratory analysis will be conducted between olanzapine and placebo groups.

3.8.2. Primary Efficacy Endpoint

The primary efficacy endpoint is change from baseline in PANSS total score at Week 4.

3.8.2.1. Primary Analysis

For the primary efficacy endpoint, change from baseline in PANSS total score at Week 4, primary analysis will be carried out using a mixed model with repeated measurements (MMRM) with an unstructured variance-covariance matrix based on the observed data for the FAS population. The model will include region (US vs. non-US), visit, treatment, and interaction term of visit and treatment as categorical variables, and baseline PANSS total score as a covariate. The Kenward-Roger approximation[2] will be used to adjust the denominator degree of freedom. The change from baseline in PANSS total score at Week 4 in the ALKS 3831 group will be compared to that in the placebo group using superiority testing at two-sided alpha level of 0.05. The least squares (LS) mean, standard error (SE), and LS mean difference between ALKS 3831 and placebo along with the 95% CI, will be summarized.

3.8.2.2. Sensitivity Analysis

To assess the robustness of the primary analysis of the primary endpoint, the following sensitivity analyses will be performed:

- Delta-adjusted Pattern Mixture Model[3] will be conducted to assess the impact of missing data. It incorporates the clinical assumption that ALKS 3831 subjects who discontinue at a given time point would have, on average, their unobserved PANSS worsened by some amount δ compared with the observed PANSS of subjects on the same treatment arm who continue to the next timepoint. Subjects who discontinue from the placebo arm would have the same PANSS trajectory as the placebo subjects who stay on the study. A sequential regression-based multiple imputation (MI) procedure will be used to incorporate the assumption and to allow uncertainty in the imputations to be reflected appropriately in the analysis. The imputation model will

include the measurement at the current timepoint as the response variable, and region (US vs. non-US) as factor, the measurements at the previous timepoints and the baseline as covariates. Twenty imputations will be carried out. For each of the 20 imputed data sets, the ANCOVA model with region (US vs. non-US) and treatment group as factors and baseline PANSS as covariate will be fitted to the change from baseline at Week 4 to obtain the treatment effect estimate and standard error. Rubin's rule will be used to combine the treatment effect estimates and standard errors across imputations. The shift parameters to be used in the sensitivity analysis will accounts for 10%, 20%, 30%, 40%, and 50% of the observed treatment difference between ALKS 3831 and placebo.

- The MMRM model may adjust for additional covariates and/or factors.

3.8.2.3. Examination of Subgroups

Subgroup analyses of the primary endpoint will be performed for each of the following subgroup.

- Sex (male, female)
- Age (<55 years, ≥55 years)
- Race (white, black or African American, other)
- Region (US, non-US)
- Baseline PANSS total score (<95 points, ≥95 points)

The forest plot of the least squares mean difference along with 95% CI between ALKS 3831 and placebo will be provided by the subgroup factors listed above.

3.8.3. Key Secondary Endpoint

The key secondary efficacy endpoint is change from baseline in CGI-S at Week 4.

Primary Analysis

An MMRM analysis, similar to the primary analysis on primary endpoint (Section 3.8.2.1) will be conducted to analyze the difference between ALKS 3831 and placebo groups.

Sensitivity Analysis

To assess the impact of missing data, a Delta-adjusted Pattern Mixture Model[3] similar to the sensitivity analysis of the primary endpoint in Section 3.8.2.2 will be conducted.

3.8.4. Multiple Comparison / Multiplicity

If the primary analysis of the primary endpoint is statistically significant between ALKS 3831 and placebo, the statistical test will be performed on the key secondary efficacy endpoint between ALKS 3831 and placebo at two-sided alpha of 0.05.

3.8.5. Other Endpoints

- Change from baseline in PANSS total score and PANSS subscales (positive, negative, general psychopathology) by visit
- Change from baseline in CGI-S by visit
- CGI-I at each post-Baseline visit
- PANSS responders (≥30% improvement from baseline in PANSS total score) by visit
- CGI-I responders (CGI-I score of 2 [much improved] or 1 [very much improved]) by visit
- Overall Response defined as: PANSS total score ≥30% improvement from baseline or CGI-I score of 1 or 2 by visit
- Change from baseline in PANSS-Excited Component (EC) score (comprised of 5 items, excitement, tension, hostility, uncooperativeness, and poor impulse control) by visit
- Change from baseline in MOAS score by visit

The CGI-I score at each postbaseline visit will be summarized by descriptive statistics.

For proportion of subjects exhibiting improvement on the CGI-I assessment (CGI-I score of ≤2), a logistic regression based on LOCF imputation for missing data will be used to compare ALKS 3831 with placebo. The model will include the region (US vs. non-US) and treatment group as factors, and baseline PANSS total score as covariates.

All other endpoints will be summarized by treatment group for all visits. Each active group (ALKS 3831 or olanzapine) will be compared to placebo. For continuous outcome, a similar approach to that for the primary endpoint will be used (Section 3.8.2.1). For dichotomous outcomes, a similar approach to that used for CGI-I improvement scores will be used, using LOCF to impute for missing data and adjusting for their corresponding baseline values.

For subjects with MOAS total score of at least 1 at baseline, the change from baseline in MOAS score and subscales will be analyzed using a similar MMRM model as Section 3.8.2.1, adjusting for baseline value as covariate.

3.9. Safety Analysis

3.9.1. Adverse Events

Adverse events will be coded by System Organ Class and Preferred Term using the Medical Dictionary for Regulatory Activities (MedDRA®) Version 19.0 or higher. The verbatim term will be included in the AE listings.

An AE (classified by preferred term) will be considered as treatment-emergent AE (TEAE) if the event is newly occurring or worsening on or after the date of first dose of study drug.

The number and percentage of subjects reporting TEAEs in each treatment group during the double-blind period will be presented by treatment group for the following categories:

- System organ class and preferred term
- Preferred terms in decreasing frequency, and also including the following subsets:
  - Experienced by ≥2% of subjects in any treatment group
  - Experienced by ≥5% of subjects in any treatment group and ≥2 times of placebo group
- System organ class, preferred term, and severity
- System organ class, preferred term for severe TEAEs
- System organ class, preferred term, and relationship
- System organ class, preferred term for drug-related TEAEs

If more than one AE is coded to the same preferred term for the same subject, the subject will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by severity and by relationship to study drug.

In addition, the number and percentage of subjects reporting AEs during the Safety Follow-up Period will be tabulated by the system organ class, preferred term and treatment group.

The number and percentage of subjects who have serious adverse events (SAE) and AEs leading to premature discontinuation from the treatment will be summarized by preferred term and treatment group and will be sorted by decreasing frequency for ALKS 3831 group.

3.9.1.1. Other significant AEs

In addition, incidence of a selected subset of relevant AEs in this class of drugs (eg, extrapyramidal symptoms (EPS) TEAEs, AEs associated with abuse potential, AEs associated with drug withdrawal, and suicide related events, etc.) will be summarized by treatment group and preferred term. The selection of AEs per subset will be based on the preferred term and Standardized MedDRA queries (SMQs).

3.9.2. Clinical Laboratory Parameters

Laboratory parameters will be presented in conventional (ie, US) units. Only scheduled laboratory parameters will be included in the summaries. All laboratory data, including those collected at unscheduled visits, will be included in the listings.

Laboratory results (baseline and change from baseline) for the Safety population for chemistry and hematology parameters will be summarized by treatment group and by visit.

Number (percentage) of subjects with potentially clinically significant (PCS) values at any post-baseline visit, will be summarized by treatment group. PCS criteria are presented in Table 1. The denominator is all subjects with non-PCS baseline and at least one post-baseline assessment in the Safety Population and the numerator is the number of subjects with non-PCS baseline and PCS at post-baseline. All PCS values including baseline PCS values will be included in supportive listings.

Shift tables for selected metabolic parameters (glucose, total cholesterol, LDL, HDL, tryglyceride, and HbA1c) and liver function tests will be presented. The criteria are summarized in Table 2, Table 3 and Table 4.

Table 1: Criteria of Potentially Clinically Significant (PCS) Abnormality for Selected Analytes

Parameters	Criteria
Chemistry	
Albumin	<2.5 g/dL
Alkaline Phosphatase (U/L)	≥3 × ULN
Alanine Aminotransferase (U/L)	≥3 × ULN
Aspartate Aminotransferase (U/L)	≥3 × ULN
Bilirubin, Total	≥2.0 mg/dL
Blood Urea Nitrogen	>30 mg/dL
Cholesterol, Random	>300 mg/dL
Cholesterol, Fasting	≥240 mg/dL
Cholesterol, HDL Fasting	≤30 mg/dL
Cholesterol, LDL Fasting	≥160 mg/dL
Creatine Kinase (U/L)	≥3 × ULN
Creatinine	≥2.0 mg/dL
Glucose, Random	<50 mg/dL or ≥200 mg/dL
Glucose, Fasting	<50 mg/dL or ≥126 mg/dL
Potassium	<3 mmol/L or >5.5 mmol/L
Lactate Dehydrogenase (U/L)	>3 × ULN
Prolactin (Female)	>30 ng/mL
Prolactin (Male)	>20 ng/mL
Sodium	<130 mmol/L or >150 mmol/L
Triglycerides, Fasting (Female)	≥120 mg/dL
Triglycerides, Fasting (Male)	≥160 mg/dL

Table 1: Criteria of Potentially Clinically Significant (PCS) Abnormality for Selected Analytes (Continued)

Parameters	Criteria
Hematology	
Eosinophils	$>1.0 \times 10^3/\mu L$
Hematocrit (Female)	≤32%
Hematocrit (Male)	≤37%
Neutrophils, Absolute	$<1.5 \times 10^3/\mu L$
Platelets	$<75.0 \times 10^3$ cells/μL or $\geq 700.0 \times 10^3$ cells/μL
Leukocytes	$\leq 2.8 \times 10^3/\mu L$ or $\geq 16.0 \times 10^3/\mu L$

Table 2: Shifts Category from Baseline to Any Post-Baseline for Selected Lipid Parameters

Total cholesterol (fasting) mg/dL
Normal (<200) to high (≥240)
Borderline (≥200 and < 240) to high (≥240)
Normal/Borderline (<240) to high (≥240)
Normal (<200) to borderline/high (≥200)
Increase ≥40 mg/dL
LDL Cholesterol (fasting) mg/dL
Normal (<100) to high (≥160)
Borderline (≥100 and <160) to high (≥160)
Normal/borderline (<160) to high (≥160)
Normal (<100) to borderline/high (≥100)
Increase ≥30 mg/dL
HDL Cholesterol (fasting) mg/dL
Normal (≥40) to low (<40)
Decrease ≥20 mg/dL
Triglycerides (fasting) mg/dL
Normal (<150) to high (≥200)
Normal (<150) to very high (≥500)
Borderline (≥150 and <200) to high (≥200)

Borderline (≥150 and <200) to very high (≥500)
Normal/borderline (<200) to high (≥200)
Normal/borderline (<200) to very high (≥500)
Normal (<150) to borderline/high/very high (≥150)
Increase ≥50 mg/dL

Table 3: Shift Category from Baseline to Any Post-Baseline in Glucose and Hba1c

Serum glucose (fasting) mg/dL
Normal (<100) to high (≥126)
Impaired (≥100 and <126) to high (≥126)
Normal/Impaired (<126) to high (≥126)
Increase ≥10 mg/dL
HbA1c %
Shift from baseline (<5.7%) to post-baseline ≥5.7%
Shift from baseline (<5.7%) to post-baseline ≥5.7% and <6.5%
Shift from baseline (<5.7%) to post-baseline ≥6.5%

Table 4: Shift Category from Baseline to Any Post-Baseline in Liver Function Test

Alanine Aminotransferase (ALT) (U/L)
Shift from Normal to ≥3 x ULN
Shift from Normal to ≥5 x ULN
Shift from Normal to ≥10 x ULN
Aspartate Aminotransferase (AST) (U/L)
Shift from Normal to ≥3 x ULN
Shift from Normal to ≥5 x ULN
Shift from Normal to ≥10 x ULN
Bilirubin, Total (mg/dL)
Shift from Normal to >1 xULN
Shift from Normal to ≥2 xULN

3.9.3. Vital Signs, Body Weight, and ECG

3.9.3.1. Vital Signs

Descriptive statistics for vital signs and changes from baseline values at each scheduled time point will be presented by treatment group.

Vital sign values will be considered PCS if they meet both the observed-value criteria and the change-from-baseline criteria listed in Table 5. The number and percentage of subjects with PCS post-baseline values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with available non-PCS baseline values and at least 1 post-baseline assessment. The numerator will be the total number of subjects with available non-PCS baseline values and at least 1 PCS post-baseline value. A supportive tabular display of subjects with PCS post-baseline values will be provided, including the subject ID number, baseline, and all post-baseline (including non-PCS) values.

All vital signs will be presented in the subject data listing.

Orthostatic hypotension (20/10) is defined as a fall in systolic blood pressure of at least 20 mmHg and a fall in the diastolic blood pressure of at least 10 mmHg upon standing from supine. Orthostatic hypotension (30) is defined as a fall in systolic blood pressure of at least 30 mm Hg upon standing from supine.

Orthostatic tachycardia is defined as a heart rate increase of 30 beats per minute (bpm) or more upon standing from supine, or over 120 bpm upon standing.

The number and percentage of subjects with orthostatic hypotension or orthostatic tachycardia occurring at any post-baseline visit will also be summarized by treatment group for the double-blind period.

Table 5: Criteria for Potentially Clinically Significant (PCS) Blood Pressure or Pulse Rate

Parameter	Criteria
Supine Systolic Blood Pressure	≤90 and decrease ≥20 mm Hg ≥180 and increase ≥20 mm Hg
Supine Diastolic Blood Pressure	≤50 and decrease ≥15 mm Hg ≥105 and increase ≥15 mm Hg
Supine Heart Rate	≤50 and decrease ≥15 bpm ≥120 and increase ≥15 bpm

3.9.3.2. Weight and Body Mass Index

Weight (kg), BMI (kg/m^2) and waist circumference (baseline and change from baseline) will be summarized by treatment group. Height (cm) will be measured at the screening visit and this value will be used for the calculation of the BMI. All weight, BMI and waist circumference data will be listed.

Number and percentage of subjects with weight change values considered as PCS occurring at any post-baseline visit will be summarized by treatment group. Criteria for PCS are presented below. The percentages will be calculated relative to the number of subjects in the Safety Population with at least one post-baseline value. A supportive listing will be provided for subjects with PCS values.

Table 6: Criteria for Potentially Clinically Significant (PCS) Changes from Baseline in Body Weight

Parameter	Criteria
Body Weight	Decrease from Baseline ≥7% Increase from Baseline ≥7%

3.9.3.3. Electrocardiograms

Descriptive statistics for ECG parameters (heart rate, PR interval, RR interval, QRS interval, QT interval, and QTc interval) at baseline and change from baseline values at each assessment timepoint and at the end of the double-blind treatment period will be presented by treatment group. QTc interval will be calculated using both Bazett ($QTcB = QT/(RR)^{1/2}$) and Fridericia ($QTcF = QT/(RR)^{1/3}$) corrections; if RR is not available, it will be replaced with 60/HR in the correction formula.

Electrocardiogram parameter values are considered PCS if they meet or exceed the higher-limit PCS criteria listed in Table 7. The number and percentage of subjects with PCS post-baseline ECG values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with normal baseline values and at least 1 post-baseline assessment. The numerator is the total number of subjects with normal baseline values and at least 1 PCS post-baseline value. A supportive tabular display of subjects with PCS post-baseline values will be provided, including the subject ID number, study center number, baseline, all post-baseline (including non-PCS) values, and change from baseline.

Table 7: Criteria for Potentially Clinically Significant (PCS) QTcB and QTcF

Parameter	Criteria
QTcB and QTcF	>450 to ≤480 msec
	>480 to ≤500 msec
	>500 msec
	Change from baseline >30 to ≤60 msec
	Change from baseline >60 msec

3.9.4. Abnormal Movement Scales

Extra pyramidal symptoms (EPS) will be evaluated as AEs and also assessed by abnormal movement scales. Abnormal movement scales will include the following: Abnormal Involuntary Movement Scale (AIMS), Barnes Akathisia Rating Scale (BARS), and Simpson-Angus Scale (SAS).

For all abnormal movement scales, total scores and subscale scores will be summarized by treatment group at each visit for the absolute value and for changes from baseline.

Number and percentage of subjects meeting the criteria for treatment emergent Parkinsonism (SAS total score>3), for treatment emergent akathisia (BARS global clinical assessment of akathisia score ≥2), for treatment emergent dyskinesia (AIMS score ≥3 on any of the first 7

items, or a score ≥2 on two or more of any of the first 7 items) at any postbaseline visit will be summarized by treatment group.

A listing will be provided for every abnormal movement scale. Listing for treatment emergent EPS will be provided.

3.9.5. Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is a questionnaire used to measure the presence and intensity of suicidal ideation and behavior.

Suicidal behavior and suicidal ideation will be summarized descriptively. The number of subjects with suicidal ideation and suicidal behavior during the double-blind treatment period will be summarized by treatment group when applicable.

Supportive tabular display of subjects with all values will be provided, including subject ID number, treatment group, visit number, intensity of suicidal ideation, suicidal behavior type, and lethality of suicidal behavior.

Table 8: C-SSRS Categories for Analysis

Category	C-SSRS Item response is "YES"
Suicidal behavior	Preparatory acts or behavior Aborted attempt Interrupted attempt Actual attempt Complete suicide
Suicidal ideation	Wish to be dead Non-specific active suicidal thoughts Active suicidal ideation with any methods (not plan) without intent to act Active suicidal ideation with some intent to act, without specific plan Active suicidal ideation with specific plan and intent

3.10. Pharmacokinetic/ Pharmacodynamic Data Analysis

Subject listings for the PK sampling time and concentrations of olanzapine, samidorphan, and metabolites of interest will be provided. PK data obtained from plasma samples collected in this study may be included in a subsequent population PK analysis or other post-hoc analyses conducted outside the scope of this SAP.

4. INTERIM ANALYSIS

No interim analysis is planned for this study.

5. CHANGES IN CONDUCT OR PLANNED ANALYSES FROM THE PROTOCOL

Changes have being made to the analyses specified in the Protocol[1] Section 14.4 to include the region (US vs. non-US) as a factor in the analysis model for all the relevant efficacy endpoints as described above.

6. DEFINITIONS AND CONVENTIONS FOR HANDLING OF THE DATA

Dataset specifications will be provided in a separate document.

6.1. Analysis Visit Windows

Scheduled analysis visits are visits at scheduled timepoints as specified in the protocol (Table 1 Schedule of Visits and Assessments).

Scheduled analysis visits during the study period will be the same as the nominal visits collected in eCRF. There will be one valid value of assessment kept for each scheduled analysis visit in summary/analysis statistics.

Unscheduled visits are visits with data not collected on the scheduled time point. Unscheduled visits will not be used for by-visit summary/analysis statistics unless specified otherwise.

All unscheduled visits will be included as collected in eCRF in listings.

Visit Day is calculated as (visit date) – (date of the first dose of study drug) +1.

Last postbaseline values are defined as the last valid postbaseline values collected for each subject during the double-blind treatment period.

Any post-baseline values are defined as any valid values collected from scheduled visits if applicable.

6.2. Handling of Partial Dates of Concomitant Medication

Partial start dates of prior and concomitant medications will be assumed to be the earliest possible date consistent with the partial date. Partial stop dates of prior and concomitant medications will be assumed to be the latest possible date consistent with the partial date. In the case of completely missing stop date, medication will be assumed to be ongoing.

6.3. Handling of Safety Data

All efforts should be made to obtain the missing information from the investigator. For C-SSRS, vital signs, laboratory testing (chemistry, hematology, urinalysis), and 12-lead ECGs, only observed data will be used for analyses, and missing data will not be imputed.

7. GENERAL STATISTICAL METHODOLOGY

In general, summary statistics (n, mean, standard deviation [SD], median, minimum, and maximum for continuous variables, and number and percentage of subjects in each category for categorical variables) will be provided by treatment group. All summary tables will be based on observed data, and missing values will not be imputed unless otherwise indicated. Measurements collected from unscheduled visits or repeated assessments will not be included in the by-visit summary tables or figures, but will be included in the analyses for the PCS postbaseline values, and subject listings. Source data for the summary tables and statistical analyses will be presented as subject data listings.

7.1. Reporting Precision

Summary statistics will be presented to the following degree of precision, unless otherwise specified:

Table 9: Degree of Precision

Statistics	Degree of Precision
Mean, Median, Confidence limit boundaries	One more than the raw data, up to 3 decimal places.
Standard deviation, Standard error	One more than the mean, up to 3 decimal places.
Minimum, Maximum	The same as the raw data, up to 2 decimal places.
p-value	Rounded to 3 decimal places and therefore presented as 0.xxx; p-values smaller than 0.001 as '<0.001'; p-values greater than 0.999 as '>0.999'.
Percentage	One decimal place. A percentage of 100% will be reported as 100%. Percentages of zero will be reported as 0.

Fractional numeric values will be presented with a zero to the left of the decimal point (for example, 0.12-0.30).

For weight, height, and BMI, one decimal place will be used for summary statistics.

8. PROGRAMMING SPECIFICATIONS

Programming specifications will be provided in a separate document.

9. MOCK TABLES, LISTINGS AND FIGURES (TLFS)

Mock-up tables and listings will be provided in a separate document.

10. REFERENCES

1. Alkermes ALK3831-A305 Study Protocol Amendment 1.0 Date: 22 September 2016

2. Kenward MG and Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* 1997; 53: 983–997
3. Ratitch B, O'Kelly M, and Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharmaceutical statistics* 2013; 12: 337-347