A Phase 2, Efficacy, Safety, and Tolerability Study of ALKS 3831 in Schizophrenia with Alcohol Use Disorder

Unique Protocol ID:	ALK3831-401
NCT Number:	NCT02161718
Date of Statistical Analysis Plan:	20 Jan 2017

STATISTICAL ANALYSIS PLAN

ALK3831-401

Study Title:	A Phase 2, Efficacy, Safety, and Tolerability Study of ALKS 3831 in Schizophrenia with Alcohol Use Disorder
Document Status:	Final Version 1.0
Document Date:	20 January 2017
Based on:	Protocol Amendment 6: 01 Mar 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
ANCOVA	Analysis of Covariance
BMI	Body Mass Index
CGI-S	Clinical Global Impression–Severity
CGI-I	Clinical Global Impression–Improvement
CI	Confidence Interval
CRF	Case Report Form
C-SSRS	Columbia Suicide Severity Rating Scale
ET	Early Termination
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EEDS	Event of Exacerbation of Disease Symptom
HbA1c	Hemoglobin A1c
HDL	High Density Lipoprotein
HR	Heart Rate
ITT	Intention to Treat
LDL	Low Density Lipoprotein
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
PANSS	Positive and Negative Syndrome Scale
PBO	Placebo
PD	Pharmacodynamic
PK	Pharmacokinetics
PSP	Personal and Social Performance
PT	Preferred Term
QTcB	QTcB – Bazett's correction formula
QTcF	QTcF – Fridericia's correction formula
SAE	Serious Adverse Event

SAP	Statistical Analysis Plan
SD	Standard Deviation
SOC	System Organ Class
TLF	Table, Listing, and Figure
TLFB	Timeline Follow-Back
TEAE	Treatment-Emergent Adverse Event
WHO	World Health Organization
WHO-ATC	World Health Organization Anatomical Therapeutic Chemical

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical methods to be used for analyses and data presentation for reporting efficacy and safety results for study ALK3831-401. This document has been prepared based on Alkermes ALK3831-401 Study Protocol Amendment 6[1] (dated 01 Mar 2016).

1.1. Study Objectives

Primary:

- To evaluate the efficacy of olanzapine coadministered with samidorphan (ALKS 3831) compared with olanzapine coadministered with placebo in schizophrenia with alcohol use disorder (AUD)

Secondary:

- To evaluate the safety and tolerability of ALKS 3831 in schizophrenia with AUD

1.2. Summary of the Study Design

This is a Phase 2, multicenter, randomized, double-blind, placebo-controlled study designed to evaluate the efficacy, safety, and tolerability of ALKS 3831 in schizophrenia with AUD.

ALKS 3831 is a combination of olanzapine, an approved antipsychotic treatment for schizophrenia, and samidorphan, a new molecular entity.

Potential subjects for this study are adults with a diagnosis of schizophrenia (based on the Diagnostic and Statistical Manual of Mental Disorders 4th edition, text revision [DSM-IV-TR]) and AUD (based on DSM 5th edition [DSM-5]). Subjects must have recently (within the last 6 months) experienced an exacerbation of disease symptoms (eg, hospitalization) but cannot exceed a pre-defined level of symptom severity at the time of screening (as measured by the Positive and Negative Syndrome Scale [PANSS] and Clinical Global Impression-Severity [CGI-S]).

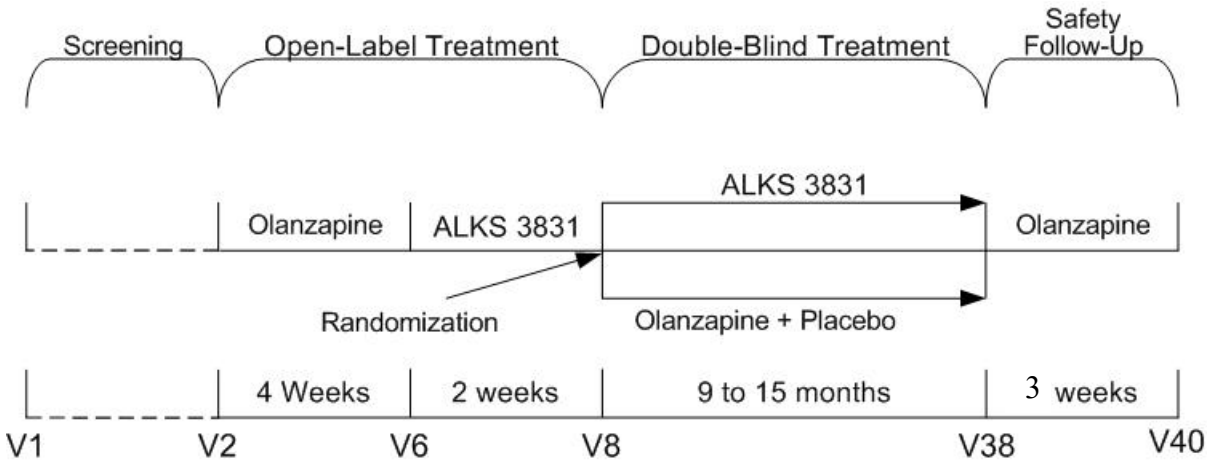
All study visits will be outpatient visits. After informed consent and screening, subjects will begin a 4-week open-label olanzapine dosing period followed by a 2-week open-label ALKS 3831 period (see Figure 1). Once-daily, open-label olanzapine dosing will begin at Visit 2 with the dose level determined by the principal investigator (PI). Once daily open-label ALKS 3831 dosing (olanzapine coadministered with 10 mg samidorphan using separate tablets) will begin at Visit 6 and last until Visit 8 (randomization). ALKS 3831 treatment will consist of a once-daily dose of olanzapine (with dose level determined by the PI) plus samidorphan 10 mg to be taken at the same time (ie, coadministered) using separate tablets.

At Visit 8, subjects will be randomized in a 1:1 fashion to 1 of 2 treatment groups as follows:

- ALKS 3831 (ie, olanzapine [dose determined by the PI] plus samidorphan [10 mg])
- Olanzapine (dose determined by the PI) plus placebo

Subjects randomized to double-blind treatment will be followed for a minimum of 9 months and a maximum of 15 months. Visits during the double-blind period will be every 2 weeks. Each visit will consist of clinical assessments plus drug dispensing or drug dispensing only. After the end of the double-blind treatment period, subjects will complete a 3-week safety follow-up period.

Figure 1: ALK3831-401 Study Schematic



Abbreviation: V = Study Visit; ALKS 3831=olanzapine + samidorphan

The schedule of visits and assessments for before randomization (Visit 8) and after randomization are shown in Table 1, Table 2, and Table 3 of the protocol (Study ALK3831-401[1]).

2. SAMPLE SIZE CONSIDERATION

The sample size calculation is based on the detection of hazard ratio for the first event of exacerbation of disease symptoms (EEDS) during the double-blind treatment period.

Based on the previous clinical studies with olanzapine and clinical judgment, the 9-month cumulative proportion of subjects experiencing an event of exacerbation of disease symptoms (EEDS) will be 15% and 30% for the ALKS 3831 group and olanzapine plus placebo group, respectively. Assuming an exponential distribution of survival time, and a 2-sided test at α=0.05, a total of 70 EEDS will be required for approximately 90% power to detect a hazard ratio of 0.45 between ALKS 3831 and olanzapine plus placebo with a randomization allocation ratio of 1:1 (ALKS 3831:olanzapine plus placebo). Approximately 270 subjects (135 subjects per treatment group) will need to be randomized at Visit 8 to observe 70 total EEDS.

3. DATA ANALYSIS

3.1. General Statistical Methodology

The baseline for each efficacy and safety parameter in a period is defined as the last assessment made on or before the first dose of study drug in that period. That is, baseline for open-label Olanzapine and ALKS 3831 period is defined as the last non-missing observation on or before the first dose of study drug for the open-label period at Visit 2 and 6, respectively; baseline for double-blind period is defined as the last non-missing observation on or before the date of the first dose of study drug for double-blind period at Visit 8.

Unless specified otherwise, the efficacy parameters will be summarized for the double-blind period and the safety parameters will be summarized for the open-label, double-blind and safety follow-up periods.

Continuous variables will be summarized by number of subjects and mean, SD, median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of subjects.

3.2. Definitions of Analysis Populations

3.2.1. Open-label Olanzapine Population

The open-label olanzapine population is defined as all enrolled subjects who received at least 1 dose of olanzapine during the open-label olanzapine treatment period.

3.2.2. Open-label ALKS 3831 Population

The open-label ALKS 3831 population is defined as all enrolled subjects who received at least 1 dose of olanzapine or samidorphan during the open-label ALKS 3831 treatment period.

3.2.3. Safety Population

The safety population is defined as all randomized subjects who receive at least 1 dose of study drug (olanzapine, samidorphan or placebo) during the double-blind placebo-controlled treatment period.

Follow-up safety population is defined as subjects who entered into the follow-up period or have any newly emergent adverse events (Section 3.8.1).

3.2.4. Efficacy Population

The intent-to-treat (ITT) population is the same as the safety population.

3.3. Disposition

The number and percentage of subjects completing or prematurely discontinuing during the open-label olanzapine period and the open-label ALKS 3831 period will be summarized. For the safety population, the number and percentage of subjects completing or prematurely

discontinuing the study and reason for discontinuation will be summarized by treatment group; and overall.

3.4. Demographics and Baseline Characteristics

Demographics and baseline characteristics, including but not limited to, sex, age, race, ethnicity, region, weight, BMI, PANSS, CGI-S and drinking status will be summarized for the open-label populations and by treatment group and overall for the safety population.

Assignment of the stage of readiness to change will be based on the quick method[3]. The scale score is the sum of the item scores. The highest scale score represents the stage of change designation. The score of the precontemplation scale is reversed to compare with the contemplation and action scales. If two scale scores are equal, the scale farther along the continuum of change (ie precontemplation - contemplation - action) represents the subject's stage of change designation.

Medical history will be summarized for the open-label olanzapine population (overall) and for the safety population (by treatment group and overall).

3.5. Prior and Concomitant Medication

Prior medications are defined as medications taken prior to the first dose of open-label olanzapine. Concomitant medications are defined as medications taken on or after the first dose of open-label olanzapine. All medications as documented by the investigator will be coded using the World Health Organization Anatomical Therapeutic Chemical (WHO-ATC) drug dictionary Version March 2015.

Prior and concomitant medications will be summarized by the preferred term and treatment group for the safety population. Concomitant medications that are taken during the double-blind treatment period will be included in the summary table. All reported medications (including those initiated after the last dose of double-blind study medication) will be included in the listing. 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.

3.6. Treatment Adherence Rate and Extent of Exposure

3.6.1. Treatment Adherence Rate

Treatment adherence to the daily dosing schedule of samidorphan and placebo during the double-blind period will be summarized by treatment group for the safety population. Treatment adherence will be calculated as follows:

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

3.6.2. Extent of Exposure

Exposure (days) on study drug 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, for the relevant period. Exposure to

study drug will be summarized for each population; overall for the open-label periods and by treatment group for the double-blind period.

The olanzapine dose level at the start of the double-blind treatment period (V8) will be summarized by treatment group and overall.

3.7. Efficacy Analyses

3.7.1. 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, unless specified otherwise.

3.7.2. Primary Efficacy Endpoint

The primary efficacy endpoint is defined as the time from randomization to the first EEDS.

An EEDS is defined as any of the following occurring during the double-blind period and is related to worsening of disease symptoms, as confirmed by the IAC (protocol section 11.1[1]):

1. Subject requires hospitalization due to worsening of disease symptoms including a psychiatric hospitalization or a hospitalization due to or for acute treatment of alcohol intoxication or withdrawal
2. Subject has a ≥25% or ≥15-point increase from randomization in PANSS total score that is confirmed at a second, ad hoc visit shortly after the first assessment
3. Subject has a PANSS item score of P1, P2, P3, P6, P7, or G8 that meets one of the following criteria:
   - For subjects with a score ≤3 at randomization, a score ≥5 that is confirmed at a second, ad hoc visit shortly after the first assessment
   - For subjects with a score of 4 at randomization, a score ≥6 that is confirmed at a second, ad hoc visit shortly after the first assessment
4. Subject commits deliberate self-injury, engages in aggressive behavior, or displays signs of suicidal or homicidal ideation that is clinically significant as judged by the PI
5. Subject requires initiation of rescue medication, including an increase in prescribed olanzapine dose level, due to worsening of disease symptoms
   - For this study, the increase in olanzapine dose must be above the randomization dose level even if the dose level is previously decreased at some point after randomization.
6. Subject requires an emergency room visit due to worsening of disease symptoms
7. Subject withdraws or PI discontinues the subject from the study for one of the following reasons:
   a. Lack of efficacy
   b. Lost to follow-up
   c. Withdrawal by subject

8. Subject is involved in an incident leading to arrest or incarceration during the study period that is related to the subject's underlying disease

For criterion requiring a consecutive assessment, the second assessment should be scheduled within 3-7 days of the first occurrence and conducted as per Section 8.5 of the protocol[1]. All EEDS cases will be reviewed by the IAC in a blinded manner. Only events confirmed by the IAC will be used for efficacy analyses. Further details on the functioning of the IAC are included in the IAC charter[2].

Subjects who completed or discontinued the double-blind period without an EEDS will be censored at the last EEDS assessment date. For all other subjects with EEDS, the date of first EEDS will be counted as event date.

3.7.2.1. Primary Analysis of the Primary Efficacy Endpoint

For the primary analysis of time to the first EEDS, comparison of the 2 treatment groups will be carried out using a log rank test. The Cox proportional hazards model will be used to obtain an estimate of the hazard ratio for the ALKS 3831 group to the olanzapine group adjusting for the baseline PANSS total score and drinking status. A 95% confidence interval will be computed for the hazard ratio.

Kaplan-Meier estimates of EEDS rate will be generated and the Kaplan-Meier plot of time to first EEDS will be presented by the treatment group.

3.7.2.2. Sensitivity Analysis of the Primary Efficacy Endpoint

As sensitivity analysis, if the first adjudicated EEDS only satisfies the discontinuation criterion, (either satisfies 7b or 7c criteria in EEDS definition, or if IAC doesn't specify a criterion), it will be censored at the last EEDS assessment date. Analysis will be the same as the primary analysis of EEDS.

Additional sensitivity analysis will be performed to assess the potential impact of informative censoring. 10%, 20% and 30% of subjects who discontinued from the study without any EEDS will be randomly assigned as having EEDS one day after the discontinuation. Twenty imputations will be carried out. The generated datasets will be analyzed similarly as the primary analysis using the cox proportional hazard model. Rubin's rule will be used to combine the treatment effect estimates and standard errors across imputations.

3.7.2.3. Additional Analysis of the Primary Efficacy Endpoint

Exploratory analysis will be performed to analyze the relationship of the primary endpoint to the drinking profiles during the double-blind treatment period. Percent of heavy drinking days during the double-blind treatment period will be subset into quantiles and used as the factor in the Cox proportional hazards model. The model will also adjust for the baseline PANSS total score. Kaplan-Meier estimates of EEDS rate will be generated.

3.7.2.4. Examination of Subgroups

The following subgroup category will be examined for the primary efficacy endpoint:

- Male vs. Female

- Black vs. non-Black
- US vs. Non-US
- Prior olanzapine exposure (subjects on olanzapine at Visit 1 vs. subjects not on olanzapine at Visit 1)
- Baseline PANSS (≥70 vs. <70)
- Baseline alcohol consumption risk category (medium, high or very high risk vs. abstinence or low risk)
- Baseline AUD severity (severe vs. mild or moderate)
- Baseline number of prior hospitalizations or ER visits for the underlying disease in the last 12 months (≥1 vs. 0)

For each subgroup category, the Cox proportional hazards model will be used to obtain an estimate of the hazard ratio of the ALKS 3831 group to the olanzapine group. The model will include treatment group, subgroup, and treatment by subgroup interaction as factors and adjust for the relevant baseline covariates/factors when appropriate.

3.7.3. Secondary Efficacy Endpoints

Secondary efficacy endpoints include:

- The rate and number of EEDS
- Proportion of subjects with at least 1 level decrease in World Health Organization (WHO) drinking risk level from baseline to week 24 of the double-blind treatment (study week 27)

3.7.3.1. Rate and Number of EEDS

3.7.3.1.1. Primary Analysis on the Rate and Number of EEDS

Considering EEDS as recurrent events, an Andersen-Gill mean/rate intensity model will be employed to analyze the time-to-recurrent EEDS. This model will allow for an arbitrary dependence structure of the recurrent events within the same subject by using a robust "sandwich" estimate of the covariance matrix. The model will also adjust for the baseline PANSS total score and drinking status. The estimated hazard ratio (and 95% confidence interval) of the ALKS 3831 group to the olanzapine group will be calculated under the Anderson-Gill mean/rate intensity model.

3.7.3.1.2. Sensitivity Analysis on the Rate and Number of EEDS

Annualized EEDS rate will be analyzed by a negative binomial regressionmodel. The logarithmic transformation of the time from the first dose of the double-blind study drug to the last EEDS assessment visit will be included in themodel as the "offset" parameter. The model will adjust the baseline PANSS total score and drinking status. If the data are underdispersed, or if the negative binomial regression model does not converge, a Poisson regression model will be used instead of the negative binomial regression model. Dispersion will be evaluated from the

Pearson Chi-Square statistic. If the ratio of the Pearson Chi-Square statistic to the degrees of freedom is ≤ 1 which indicates no overdispersion, then a Poisson regression model will be used.

The descriptive statistics for the annualized rate of EEDS for each treatment group will be calculated as the total number of EEDS experienced in the group divided by the total number of subject-years on study. The descriptive statistics for the annualized rate of EEDS rate for an individual subject will be calculated as the number of EEDS for that subject divided by the number of subject-years and presented by the treatment group.

3.7.3.1.3. Additional Analysis on the Rate and Number of EEDS

Since the EEDS consists of a series of individual criteria, number and percent of subjects who met each individual criterion will be provided. The treatment groups will be compared by the logistic regression model adjusting for the baseline PANSS total score and drinking status.

In addition, the duration of IAC confirmed hospitalizations related to the underlying disease will be summarized by treatment group.

3.7.3.2. Proportion of Subjects with at least 1 Level Decrease in World Health Organization (WHO) Drinking Risk Level from Baseline to Week 24 of Double-blind Treatment (Study Week 27)

The alcohol consumption per day at each visit will calculated based on the number of drinks per day as follows.

$$\text{Alcohol consumption per day} = \frac{\text{total number of drinks} \times 14 \text{ gram}}{\text{total number of days}}$$

The risk of the alcohol consumption will be categorized based on the alcohol consumption per day using the criteria in Table 1. Proportion of subjects with at least 1 level decrease in World Health Organization (WHO) drinking risk level[3] from baseline to week 24 of double-blind treatment (study week 27) will be compared between two treatment groups using the logistic regression model based on LOCF (last observation carried forward) imputation for the missing data, that is, the last observed non-missing post-baseline value will be carried forward for the missing post baseline assessments. The model will include the treatment group as factor and adjust for the baseline PANSS total score and drinking status.

Table 1: Criteria for risk of alcohol consumption on a single drinking day

Risk Category	Males	Females
Abstinence	0 g	0 g
Low Risk	1 to 40 g	1 to 20 g
Medium Risk	41 to 60 g	21 to 40 g
High Risk	61 to 100 g	41 to 60 g
Very High Risk	≥101 g	≥61 g

3.7.4. Other Efficacy Endpoints

Other efficacy endpoints for the double-blind treatment period are:

- The proportion of subjects with psychosocial events indicative of exacerbation of symptoms during the double-blind treatment period
- The rate and number of psychosocial events indicative of exacerbation of symptoms during the double-blind treatment period
- Change from baseline in PANSS total Score, Positive Score, Negative Score, and General Psychopathology Score during the double-blind treatment period
- Change from baseline in Clinical Global Impression–Severity (CGI-S) score during the double-blind treatment period
- Clinical Global Impression–Improvement (CGI-I) score during the double-blind treatment period
- Complete response profile of drinking during the double-blind treatment period
- Change from baseline in visual analogue score (VAS) for perception of desire for alcohol during the double-blind treatment period

3.7.4.1. Proportion of Subjects with Psychosocial Events Indicative of Exacerbation of Symptoms during the Double-blind Treatment period

The proportion of subjects with psychosocial events indicative of exacerbation of symptoms (defined in study protocol section 11.2[1]) will be analyzed similarly as the primary analysis of the primary efficacy endpoint (Section 3.7.2.1). The incidence rate will be estimated by Kaplan-Meier estimate. The treatment comparison will be performed using the log-rank test and hazard ratio will be estimated by the Cox proportional hazards model.

3.7.4.2. Rate and Number of Psychosocial Events Indicative of Exacerbation of Symptoms during the Double-blind Treatment Period

The annualized rate and number of psychosocial events indicative of exacerbation of symptoms will be analyzed in the same way as the rate and number of EEDS (Section 3.7.3.1).

3.7.4.3. Change from Baseline in PANSS Total Score, Positive Score, Negative Score, and General Psychopathology Score during the Double-blind Treatment Period

Change from baseline of double-blind period in PANSS total score and subscale scores will be analyzed using ANCOVA (analysis of covariance) model with treatment group as factor and corresponding baseline value as the covariates for the ITT population. The ANCOVA model will be based on the LOCF imputation.

In addition, PANSS total score will be summarized by visit for the open-label olanzapine population. Plot of mean PANSS total score over time from the open-label periods (as overall) and through the double-blind period (by treatment group) will be provided for the open-label olanzapine population.

3.7.4.4. Change from Baseline in CGI-S Score during the Double-blind Treatment Period

The analysis of change from baseline in CGI-S score during the double-blind treatment period will be ANCOVA and LOCF method for the ITT population, similar to Section 3.7.4.3.

3.7.4.5. Clinical Global Impression–Improvement (CGI-I) Score during the double-blind Treatment Period

CGI-I scores will be analyzed based on the subject's by-visit responder status (≤2 versus >2) using the logistic regression model based on LOCF imputation for missing data. The model will include the treatment group as factor and the baseline PANSS as the covariate.

3.7.4.6. Complete Response Profile of Drinking during the Double-blind Treatment Period

Complete response profile of drinking during the double-blind treatment period will be analyzed by the following endpoints:

- Proportion of subjects with at least 1 level decrease in World Health Organization (WHO) drinking risk level from baseline by visit
- Proportion of subjects with at least 2 levels decrease in World Health Organization (WHO) drinking risk level from baseline by visit
- Proportion of subjects abstinent from heavy drinking days from randomization through week 24 of the double-blind treatment period (study week 27)
- Proportion of subjects abstinent from heavy drinking days by visit
- Proportion of subjects abstinent from drinking days by visit
- Number and rate of heavy drinking days
- Number and rate of drinking days
- Percent of heavy drinking days cumulatively from baseline through weeks 12, 24, 36, 48 and 60 of the double-blind treatment period
- Percent of drinking days cumulatively from baseline through weeks 12, 24, 36, 48 and 60 of the double-blind treatment period
- Percent of heavy drinking days by visit during the double-blind treatment period
- Percent of drinking days by visit during the double-blind treatment period
- Number of drinks per day by visit during the double-blind treatment period
- Number of drinks per drinking day by visit during the double-blind treatment period

In general for the TLFB assessment, by-visit result is defined as the drinking assessments between previous and current visits. It will be 2 weeks of drinking assessments for Visit 9 Week 5 and Visit 10 Week 7, and monthly thereafter.

3.7.4.6.1. Categorical endpoints for the Complete Response Profile of Drinking

The following categorical endpoints for the complete response profile of drinking during the double-blind treatment period will be analyzed using a logistic regression model.

- Proportion of subjects with at least 1 level decrease in World Health Organization (WHO) drinking risk level from baseline by visit
- Proportion of subjects with at least 2 levels decrease in World Health Organization (WHO) drinking risk level from baseline by visit
- Proportion of subjects abstinent from heavy drinking days from randomization through week 24 of the double-blind treatment period (study week 27)
- Proportion of subjects abstinent from heavy drinking days by visit
- Proportion of subjects abstinent from drinking days by visit

The model will include treatment as factor and adjust for baseline PANSS and drinking status. LOCF imputation will be used for the by-visit summary.

3.7.4.6.2. Number and Rate of Heavy Drinking Days or Drinking Days

Considering heavy alcohol drinking days or drinking days as recurrent events, an Anderson-Gill model will be used to analyze the time-to-recurrent events, similar to Section 3.7.3.1.1.

Weekly alcohol drinking and heavy drinking rate will be summarized by treatment groups and compared by the negative binomial model, similar to Section 3.7.3.1.2.

3.7.4.6.3. Percent of Heavy Drinking Days or Drinking Days Cumulatively through Weeks 12, 24, 36, 48 and 60 of the Double-blind Treatment Period

Percent of heavy drinking days or drinking days cumulatively through weeks 12, 24, 36, 48 and 60 of the double-blind treatment period (corresponding to study weeks 15, 27, 39, 51 and 63) will be compared between two treatment groups using the Wilcoxon rank sum test.

Cumulative frequency distributions of percent of heavy drinking days or drinking days through week 60 of the double-blind treatment period will be presented by treatment group with horizontal axis as percent of heavy drinking days and vertical axis as corresponding proportion of subjects in each treatment group.

Two sensitivity analyses will be conducted. First analysis will exclude any drinking records after the discontinuation of the double-blind study drug or after any forced abstinence to control for carryover effects. Forced abstinence is defined as greater than 4 days of hospitalization or any incarceration. Additional sensitivity analysis will exclude any drinking records after discontinuation of the double-blind study drug or incarceration or 1 week of drinking records after greater than 4 days of hospitalization.

3.7.4.6.4. Percent of Heavy Drinking Days or Drinking Days by Visit during the Double-blind Treatment Period

Percent of alcohol heavy drinking days or drinking days will be presented by visit and by treatment groups. In addition, plot of mean percent of heavy drinking days or drinking days by

visit from the open-label periods (as overall) and through the double-blind period (by treatment group) will be provided for the open-label olanzapine population.

3.7.4.6.5. Number of Drinks Per Day and Per Drinking Day by Visit during the Double-blind Treatment Period

Number of drinks per day and per drinking day during the double-blind treatment period will be summarized by the treatment group and by visit.

3.7.4.7. Change from Baseline in Visual Analogue Score (VAS) for Perception of Desire for Alcohol during the Double-blind Treatment Period

The analysis of change from baseline in VAS score for perception of desire for alcohol during the double-blind treatment period will be ANCOVA and LOCF method, similar to Section 3.7.4.3.

3.7.5. Exploratory Endpoints

Exploratory endpoints during the double-blind treatment period are:

- Change from baseline in cigarette use questionnaire during the double-blind treatment period
- Change from baseline in Personal and Social Performance (PSP) score during the double-blind treatment period
- Time from randomization to the discontinuation of the study drug during the double-blind treatment period

Cigarette use questionnaire and PSP score (results and change from baseline) will be summarized using descriptive statistics at each visit in the double-blind treatment period.

Time from randomization to the discontinuation of the study drug during the double-blind treatment period will be compared between ALKS 3831 and olanzapine for the safety population, similar to Section 3.7.2.1.

3.8. Safety Analysis

3.8.1. Adverse Events

Adverse events will be coded by System Organ Class and Preferred Term using the Medical Dictionary for Regulatory Activities Version 18.0 (MedDRA®). 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 first dose of study drug.

A TEAE will be summarized for the open-label olanzapine population in the open-label olanzapine period, and for the ALKS 3831 populations in the ALKS 3831 period.

TEAE will be summarized by treatment group for the safety population in the double-blind period. A TEAE will be summarized in the double-blind period if it starts between the date of the

first dose of study drug (inclusive) and the date of the last dose of samidorphan or placebo in the double-blind period.

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

- system organ class and preferred term
- preferred terms in decreasing frequency
- system organ class, preferred term, and severity
- 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 for the same period, 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 the study drug.

The incidence of most frequently reported (≥ 5% of subjects in any treatment group) TEAEs for the double-blind period will be summarized by system organ class, preferred term, and treatment group and will be sorted by decreasing frequency for ALKS 3831 group. It will also be summarized by the olanzapine dose level at randomization.

Newly-emergent AEs (NEAEs), defined as AEs that started or worsened after the date of the last dose of study drug during the treatment period will also be summarized for the follow-up safety population.

The number and percentage of subjects who have serious adverse events (SAE) will be summarized by preferred term and overall for open-label periods. It will be summarized by preferred term and treatment group for the double-blind period and will be sorted by decreasing frequency for ALKS 3831 group. A listing of subjects with SAEs for entire study will be provided by period. The number and percentage of subjects who have AEs leading to premature discontinuation from the study will be summarized in the same way as SAEs.

3.8.2. Other significant adverse events

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

3.8.3. Clinical Laboratory Parameters

Laboratory parameters will be presented in conventional 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 chemistry and hematology parameters will be summarized as overall for open-label periods and by treatment group for the double-blind period.

Clinical laboratory test values, scheduled or unscheduled, will be considered potentially clinically significant (PCS) if they meet PCS criteria listed in Table 2. The number and percentage of subjects who have PCS post baseline clinical laboratory values will be tabulated by treatment group for the double-blind period. The percentages will be calculated relative to the number of subjects with available non-PCS baseline values and at least 1 post baseline assessment for the double-blind treatment period. The numerator will be the total number of subjects with available non-PCS baseline values and at least 1 PCS post baseline value for the double-blind treatment period. A supportive tabular display of subjects with PCS post baseline values will be provided, including the subject ID number, study center number, baseline, and all post baseline (including non-PCS) values.

Table 2: Criteria of Potentially Clinically Significant (PCS) Abnormality for Selected Analytes

Parameters	**Criteria**
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	≥126 mg/dL
Potassium	<3 mmol/L or >5.5 mmol/L
Lactate dehydrogenase (U/L)	>3 × ULN
Phosphate	<2 mg/dl or >5 mg/dl
Prolactin (female)	>30 ng/mL
Prolactin (male)	>20 ng/mL
Sodium	<130 mmol/L or >150 mmol/L

Table 2: Criteria of Potentially Clinically Significant (PCS) Abnormality for Selected Analytes (Continued)

Parameters	Criteria
Triglycerides, fasting (female)	≥120 mg/dL
Triglycerides, fasting (male)	≥160 mg/dL

Table 3: Potentially Clinically Significant Results Criteria for Hematology

Parameters	Criteria
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 $\geq 16.0 \times 10^3/\mu L$

Shift tables for metabolic parameters (glucose, total cholesterol, LDL, HDL, triglycerides, and HbA1c will be presented.

3.8.4. Weight, Body Mass Index, Waist Circumferences and Vital Signs

3.8.4.1. Weight, Body Mass Index, Waist Circumferences

Weight (kg), BMI (kg/m^2) and waist circumferences (baseline and change from baseline) for each visit will be summarized as overall for the open-label periods and by treatment group for the double-blind periods. Height (cm) will be measured at screening visit and this value will be used for the calculation of the BMI. All weight and BMI 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 for the double-blind period. Criteria for PCS are presented below. The percentages will be calculated relative to the number of subjects in the safety population. A supportive listing will present all body weight for a subject and parameter for cases where at least one value was potentially significant.

Table 4: Criteria for Potentially Clinically Significant (PCS) Changes from Baseline in Body Weight

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

3.8.4.2. Vital Signs

Descriptive statistics for vital signs and changes from baseline values at each scheduled time point will be presented as overall for the open-label periods and by treatment group for the double-blind period.

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 for the double-blind period. The percentages will be calculated relative to the number of subjects with available baseline values and at least 1 post baseline assessment for the double-blind period. The numerator will be the total number of subjects with available baseline values and at least 1 PCS post baseline value for double-blind period. A supportive tabular display of subjects with PCS post baseline values will be provided, including the subject ID number, study center number, and 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 mmHg 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

3.8.5. Electrocardiograms

Descriptive statistics for ECG parameters (heart rate, PR interval, QRS interval, QT interval, and QTc interval) at baseline and change from baseline values at each assessment time point will be presented as overall for the open-label periods and by treatment group for the double-blind period. 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 6. The number and percentage of subjects with PCS post baseline ECG values will be tabulated by treatment group for the double-blind treatment period. The percentages will be calculated relative to the number of subjects with available non-PCS baseline values and at least 1 post baseline assessment for the double-blind treatment period. The numerator is the total number of subjects with available non-PCS baseline values and at least 1 PCS post baseline value for the double-blind treatment period. 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 6: Criteria for Potentially Clinically Significant (PCS) QTcB and QTcF

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

3.8.6. Abnormal Movement Scales

Extra pyramidal symptoms (EPS) will be evaluated as AEs and also as 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 post-baseline 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.

3.8.7. 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 at lifetime history, open-label periods and double-blind treatment period and safety follow-up period will be summarized by overall or 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 7: 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

3.9. Pharmacokinetic/Pharmacodynamic Data Analysis

Subject listings for PK sampling time, prior drug dosing time, and plasma concentrations of olanzapine, samidorphan, and RDC-9986 (metabolite of samidorphan) will be provided. Plasma concentration data obtained from this study may be included in a subsequent population PK analysis or other post-hoc analyses.

4. INTERIM ANALYSES

Not applicable.

5. CHANGES IN CONDUCT OR PLANNED ANALYSES FROM THE PROTOCOL

Two endpoints, which have been included in the study protocol as the other endpoints, are elevated to the secondary endpoints in the SAP (Section 3.7.3).

Three exploratory endpoints, which have not been specified in the protocol, are included in the SAP (Section 3.7.5).

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 upon scheduled time points as specified in the protocol (Table 1-3 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.

The V38/ET visit can be mapped to a scheduled visit and be used for analysis and summaries only if there is no valid value already at the mapped visit.

A mapped V38/ET visit will be treated as a scheduled visit to be included in table summaries. Otherwise, it will be treated as an unscheduled visit.

Four rules will be used to map an ET visit.

- First, a 7-day (inclusive) window (≤ last dose date + 7) will be used to determine if the assessment at the ET visit is considered to be an on-treatment result. If an ET visit is within 7 days after the last dose of samidorphan/placebo, it will be considered as a valid on-treatment visit.
- Second, visit windows (inpatient stay is excluded) will be applied to determine to which visit a V38/ET valid on-treatment assessment can be mapped to. Visit windows are defined in Table 8 below.

Table 8: Visit Window Definition

Analysis Visit to be Mapped to	Target Study Day	Visit Day	Visit Window
Visit 2	Study day -28	Day -28	[-28, -25]
Visit 3	Study day -21	Day -21	[-24,-18]

Table 8: Visit Window Definition (Continued)

Analysis Visit to be Mapped to	Target Study Day	Visit Day	Visit Window
Visit 4	Study day -14	Day -14	[-17,-11]
Visit 5	Study day -7	Day -7	[-10, 0]
Visit 6	Study day 1	Day 1	1
Visit 7	Study day 8	Day 8	[2, 11]
Visit 8	Week 3	Day 15	[12, 21]
Visit 9	Week 5	Day 29	[22, 35]
Visit 10	Week 7	Day 43	[36, 57]
Visit 12	Week 11	Day 71	[58, 85]
Visit 14	Week 15	Day 99	[86, 113]
Visit 16	Week 19	Day 127	[114, 141]
Visit 18	Week 23	Day 155	[142, 169]
Visit 20	Week 27	Day 183	[170, 197]
Visit 22	Week 31	Day 211	[198, 225]
Visit 24	Week 35	Day 239	[226, 253]
Visit 26	Week 39	Day 267	[254, 281]
Visit 28	Week 43	Day 295	[282, 309]
Visit 30	Week 47	Day 323	[310, 337]
Visit 32	Week 51	Day 351	[338, 365]
Visit 34	Week 54	Day 379	[366, 393]
Visit 36	Week 59	Day 407	[394, 421]
Visit 38	Week 63	Day 435	[422, Day of Safety follow-up -1]

Visit Day is calculated as visit date – date of the first dose of ALKS 3831 (scheduled on Visit 6) + 1.

- Third, V38/ET visit can only be mapped to the same treatment period as the last scheduled visit. Otherwise, ET visit will be considered as unscheduled visit. Additionally, V38/ET visit can only be mapped to a scheduled visit that a particular assessment is taken based on the protocol.
- Four, for subjects who discontinued before visit 6, an on-treatment V38/ET results as defined above, will be mapped to the next scheduled visit, if within the same treatment period as the last scheduled visit.

Last post baseline values are defined as the last valid post baseline values collected for each subject during the treatment period or a specified reporting period.

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, or the derivation of the last post-baseline value during treatment, but will be included in the analyses for the PCS post-baseline 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 body mass index (BMI), one decimal place will be used for summary statistics.

8. PROGRAMMING SPECIFICATIONS

Programming specifications will be provided in a separate document.

9. MOCK TABLE LISTING AND FIGURES (TLF)

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

10. REFERENCES

1. Alkermes ALK3831-401 Study Protocol Amendment 6 (dated 01 March 2016)
2. Independent Adjudication Committee Charter Version 0.6 (dated 01 March 2016)
3. Heather N and Rollnick S. Readiness to change questionnaire: user's manual (revised version) 1993.
4. Witkiewitz K, Hallgren KA et al. Clinical validation of reduced alcohol consumption after treatment for alcohol dependence using the World Health Organization risk drinking level. *Alcoholism: Clinical and Experimental Research* 2017; 41: 179-186