16.1.9. Documentation of Statistical Methods

The final Statistical Analysis Plan and document history are provided in this section, as described below.

Statistical Analysis Plan

Description	Date
Statistical Analysis Plan (Version 1.0)	20 Jun 2019
Statistical Analysis Plan (Version 2.0)	10 Jun 2021
Statistical Analysis Plan (Version 3.0)	06 Oct 2021

STATISTICAL ANALYSIS PLAN

ALK3831-A307

NCT03187769

Study Title:	A Study to Evaluate the Effect of ALKS 3831 Compared to Olanzapine on Body Weight in Young Adults with Schizophrenia, Schizophreniform, or Bipolar I Disorder Who are Early in Their Illness
Document Status:	Final Version 3.0
Document Date:	06 Oct 2021
Based on:	ALK3831-A307 Protocol Amendment 5.0 Date: 19 Jun 2019
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
AEDC	Adverse event leading to treatment discontinuation
ANCOVA	Analysis of covariance
Apo A1	Apolipoprotein A1
Apo B	Apolipoprotein B
ATC	Anatomical Therapeutic Chemical [classification system]
BMI	Body mass index
BPM	Beats per minute
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	Extrapyramidal symptoms
ET	Early termination
FAS	Full analysis set
HbA1C	Hemoglobin A1c
HDL	High-density lipoprotein
hs-CRP	High-sensitivity C-reactive protein
IL-6	Interleukin 6
IWQOL	Impact of Weight on Quality of Life
LDL	Low-density lipoprotein
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MMRM	Mixed model with repeated measurements
PCS	Potentially clinically significant
PK	Pharmacokinetic
QTcF	QT corrected with Fridericia formula
SAE	Serious adverse event
SAP	Statistical analysis plan

Abbreviation or Term	Explanation or Definition
TEAE	Treatment-emergent adverse event
TNFα	Tumor necrosis factor alpha
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 data presentation for reporting efficacy and safety results for study ALK3831-A307. This document has been prepared based on Alkermes ALK3831-A307 Study Protocol.

1.1. Study Objectives

Primary objective:

- To evaluate the effect of ALKS 3831, compared to olanzapine, on body weight in young adults with schizophrenia, schizophreniform, or bipolar I disorder who are early in their illness.

Secondary objective:

- To evaluate the safety and tolerability of ALKS 3831 in young adults with schizophrenia, schizophreniform, or bipolar I disorder who are early in their illness.

1.2 Summary of the Study Design

This is a Phase 3, multicenter, randomized, double-blind study to evaluate the effect of ALKS 3831 compared to olanzapine on body weight in young adults with schizophrenia, schizophreniform disorder, or bipolar I disorder who are early in their illness.

Subjects will be screened at Visit 1, up to 30 days prior to randomization. At Visit 2, eligible subjects will be randomized 1:1 to ALKS 3831 or olanzapine, and receive study drug for up to 12 weeks. Randomization will be stratified by diagnosis (schizophrenia/schizophreniform disorder vs bipolar I disorder), region (US vs non-US), and baseline BMI (<25 vs ≥25). Subjects on antipsychotic medications or mood stabilizers (excluding study medication) should be tapered off this medication by the end of week 2 (Visit 4). Subjects may be inpatient or outpatient at Screening; however, in the opinion of the Investigator, all subjects should be appropriate for eventual outpatient treatment. Subjects should be outpatient within 2 weeks post randomization. Cases that require inpatient treatment for longer than 2 weeks will require review and approval by the Medical Monitor on a case by case basis.

Starting with Week 2 (Visit 4), subjects will come for bi-weekly visits for the remaining 10 weeks. Study drug will be provided to subjects as coated bilayer tablets dispensed in blister packs at each visit to be taken at home (one tablet by mouth each day, preferably at bedtime). For the first week, at the discretion of the Investigator, subjects will receive 5, 10, 15 or 20 mg of olanzapine, or 5/10[1], 10/10, 15/10 or 20/10 mg of ALKS 3831. At the end of Week 1, for subjects initiated on 5 mg of olanzapine or 5/10 mg of ALKS 3831, the dose will be increased to 10 mg of olanzapine or 10/10 mg of ALKS 3831. For all other subjects, the dose may also be increased to either 15 or 20 mg of olanzapine, or 15/10 or 20/10 mg of ALKS 3831. Following this increase, the dose may be increased or decreased to 5, 10, 15, or 20 mg of olanzapine, or to 5/10, 10/10, 15/10, or 20/10 mg of ALKS 3831 at the Investigator's discretion. Dosing will be

[1] ALKS 3831 dose is represented as XX/XX where the first number represents the OLZ dose in mg and the second number represents the SAM dose in mg.

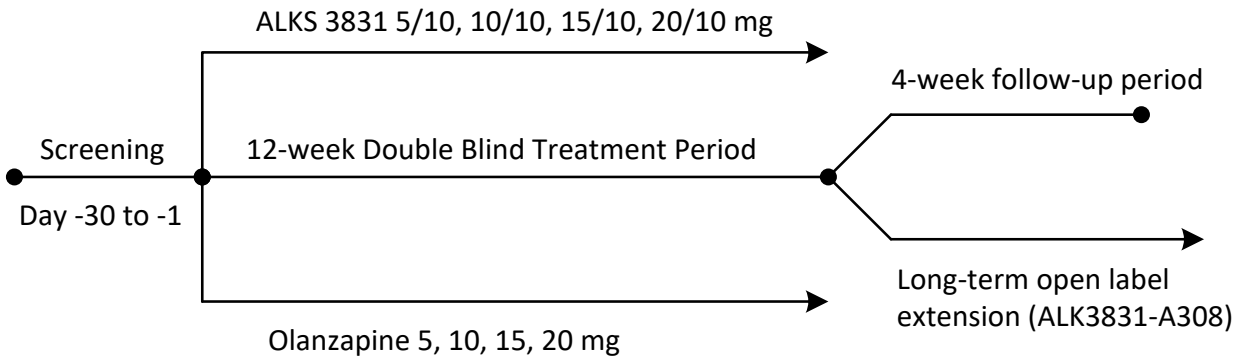
flexible throughout the study; however, frequent adjustments are discouraged. Dose adjustments can only be performed on-site at the study center. Subjects requiring dose adjustments between scheduled visits should arrange an unscheduled visit for the following procedures: study drug return and dispensation, adherence review, adverse event (AE), and concomitant medication monitoring.

Body weight, body composition (measured by Bioelectrical Impedance Analysis [BIA] using a Body Composition Analyzer), and metabolic parameters (including fasting triglycerides, cholesterol, and glycemic parameters) will be measured throughout the 12-week treatment period. Psychiatric symptoms will be evaluated using the CGI scales. Additional assessments will include Impact of Weight on Quality of Life – Lite Questionnaire (IWQOL-Lite). A daily medication adherence monitoring and reminder system (via smartphones) may be used in this study. In addition, subjects will be asked to participate in a supportive clinical care (SCC) program during the 12-week treatment period. Participation in the SCC program will be independent of study visits and based on clinician and subject preference.

In the event that a subject chooses to withdraw from the study, the Investigator should make a reasonable effort to ascertain the reason(s) for withdrawal, while fully respecting the subject's rights. Randomized subjects who terminate the treatment early (prior to Visit 9) will be asked to complete an Early Termination (ET) Visit and the Safety Follow-up Period (Visits 10 and 11). The ET Visit should be scheduled as close as possible to the subject's last dose and will mimic the assessments scheduled to be conducted at Visit 9. Subjects will then be asked to return to the study center for biweekly visits until the end of planned treatment, to collect the following information: weight, waist circumference, body composition, AEs, and concomitant medications. If the subject fails or refuses to return to the study center, an attempt must be made to contact the subject by telephone in order to assess as many safety and efficacy parameters as possible. All data collected over the telephone must be documented and kept in the subject's record.

The total study duration will be approximately 20 weeks, including an up to 4-week Screening Period, a 12-week Treatment Period, and a 4-week Follow-up Period. Subjects completing this study will be eligible to enroll in a long-term, single-arm, open-label safety study (ALK3831-A308), where it is approved by the local regulatory authority. Subjects not enrolling in the long-term safety study will enter the 4-week Safety Follow-up Period and return to their normal standard of care once their participation has ended. A schematic of the study design is provided in Figure 1.

Figure 1: Study Design



2. SAMPLE SIZE CONSIDERATION

Approximately 425 subjects will be enrolled in order to have 400 subjects (200 per treatment group) in the primary efficacy population, which is defined in Section 3.2.2. Assuming the percent weight gain at Week 12 is 5% and 8.5% for ALKS 3831 and olanzapine groups, respectively, a common standard deviation of 8%, and a cumulative dropout rate of 35%, 200 subjects per treatment group (400 subjects in total) will provide at least 90% power to demonstrate a difference of ALKS 3831 relative to olanzapine at a two-sided significance level of 0.05.

3. DATA ANALYSIS

3.1. General Statistical Methodology

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

In general, descriptive 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 for all endpoints.

The metabolic laboratory parameters will be analyzed on a linear and log-transformed scale. If log-transformed data is not normally distributed, Wilcoxon rank sum test and Hodges-Lehmann estimator will be used.

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.

All source data will be presented as subject data listings.

3.2. Analysis Sets

3.2.1. Safety Population

The Safety Population will include all randomized subjects who receive at least 1 dose of study drug (ALKS 3831 or olanzapine). The Safety Population will be used for the safety analyses.

3.2.2. Efficacy Population

The Primary Efficacy Population (ie, full analysis set [FAS]), will include all subjects in the Safety Population who have at least one postbaseline weight assessment. The FAS will be used for the efficacy analyses, unless specified otherwise.

The Retrieved Dropout population will include FAS subjects with at least one postbaseline off-treatment body weight assessment on or before Week 12. The Retrieved Dropout Population will be used for the sensitivity analysis of the primary endpoint and secondary endpoints.

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
- Subjects in the Retrieved Dropout Population

3.4. Demographics and Baseline Characteristics

Demographics and baseline characteristics such as sex, age, race, ethnicity, weight, and body mass index (BMI), psychiatric history and disease diagnosis will be summarized overall and by treatment group for the Safety Population and the FAS. If there are heterogeneities between study groups in any of the subject characteristics that are of clinical importance or could affect the treatment outcome, the impact of the imbalances will be investigated and, if necessary, appropriate adjustments made in the efficacy and safety analyses.

Medical history will be coded by System Organ Class and Preferred Term using the Medical Dictionary for Regulatory Activities (MedDRA®) Version 23.0 or higher. Medical history will be summarized for the Safety Population by treatment group and overall using the number of and percentage of subjects by Preferred Term.

3.5. Protocol Deviation

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

- Did not meet study entry 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

3.6. 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 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 WHODD B3 March 2021 or higher.

Prior and concomitant medications will be summarized by the preferred drug name for the Safety Population, separately. Concomitant medications that are taken during the Double-blind Treatment Period will be included in the summary table. Additionally, concomitant medications will be summarized by the WHO-ATC code. For subjects in the Retrieved Dropout Population, all antipsychotic medications that are taken after the last dose of study medication will be summarized separately. 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. Treatment Exposure

Duration of exposure (ALKS 3831 and olanzapine) will be 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 exposure will be summarized for the Safety Population by treatment group.

The overall mean and modal dose of olanzapine will be summarized by treatment group. The proportion of subjects at each olanzapine dose level at randomization and at final dose will also be summarized by treatment group.

3.8. Efficacy Analyses

3.8.1. General Considerations

Only the on-treatment weight assessments will be included in the analyses, unless specified otherwise.

3.8.2. Primary Efficacy Endpoint

The primary efficacy endpoint is:

- Percent change from baseline in body weight at Week 12

3.8.2.1. Primary Analysis

The null hypothesis is that there is no difference in percent change from baseline in body weight at Week 12 between ALKS 3831 and olanzapine. The alternative hypothesis is that there is a difference between ALKS 3831 and olanzapine.

The primary endpoint will be analyzed by analysis of covariance (ANCOVA) with the multiple imputation (MI) method for handling of missing values. The model will include the treatment group and randomization strata (diagnosis [schizophrenia/schizophreniform disorder vs bipolar I disorder], region [US vs non-US] and baseline BMI [<25 vs ≥25 kg/m^2]) as factors, and the baseline weight as the covariate. Only the on-treatment measurements will be included in the analysis. The following steps will be performed:

1. The missing data on body weight will be imputed using MI. If there is any subject whose missing data pattern is non-monotonic (defined as having missing data in between visits), the Markov Chain Monte Carlo (MCMC) method will be used to impute the data to a monotonic missing pattern (i.e., if the data is missing at the current visit, the data at all following visits are missing). A single Markov chain will be used with 200 burn-in iterations, 100 iterations between imputations and a Jeffrey's prior. Next, the missing data will be imputed sequentially by each visit using a regression method. The imputation regression model includes treatment group, all randomization strata as factors, and body weight at all previous visits (including baseline weight) as covariates. At least 100 imputations will be carried out. More imputations may be performed to improve the precision. The initial seed for both steps of the imputation will be 8704011.
2. Percent change from baseline in weight at Week 12 of each of these multiple imputed datasets will be analyzed by the ANCOVA model with treatment group, all randomization strata as factors, and the baseline weight as covariate.
3. Results from Step 2 will be combined using Rubin's method. Estimated treatment effect, standard error (SE), 95% CI, and p-value will be reported.

Descriptive statistics of body weight and percent change from baseline at each visit (including last assessment during treatment period) will be presented by treatment group based on observed and MI data.

3.8.2.2. Sensitivity Analysis

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

- **Delta-adjusted pattern mixture model**: To assess the potential impact of missing data due to missing not at random (MNAR), the delta-adjusted Pattern Mixture Model (Ratitch et al, 2013) will be conducted to assess the impact of missing data. It incorporates the clinical assumption that olanzapine subjects who discontinue at a given time point would have, on average, their unobserved weight gain decreased by some amount δ compared with the observed weight gain of subjects on the olanzapine arm who continue to the next time point. Subjects who discontinue from the ALKS 3831 arm would have the same weight gain trajectory as the ALKS 3831 subjects who stay on the study. A sequential regression-based 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 use the measurement at the current time point as the response variable and include the measurements at the previous time points as covariates and all randomization strata as factors. One hundred imputations will be carried out. For each of the 100 imputed data sets, the same ANCOVA model as in the

primary analysis will be fitted to the percent change from baseline at Week 12 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 will account for 10%, 20%, 30%, 40%, 50%, and up to 100% of the observed treatment difference between olanzapine and ALKS 3831 from the primary analysis of the percent change from baseline in body weight at Week 12 in FAS population.

- **Including on and off-treatment assessments:** To further assess the potential impact of missing data due to MNAR, the primary analysis will be repeated including both on-treatment and off-treatment weight assessments after premature discontinuation of study drug.
- **Additional covariates:** The ANCOVA and MI approach may adjust for additional covariates and/or factors, including but not limited to study sites.
- **Average treatment effect:** Average treatment difference across visits through Week 12 between ALKS 3831 and olanzapine will be analyzed using mixed model with repeated measurements (MMRM) model. The MMRM model will include treatment, visit, treatment-by-visit interaction term, all randomization strata as categorical fixed effects; baseline weight will be included as a covariate. An unstructured covariance structure will be applied. The Kenward-Roger approximation (Kenward and Roger 1997) will be used to adjust the denominator degree of freedom. The analysis will be performed on all observed post-randomization on-treatment weight measurements without imputation of missing data.

3.8.3. Secondary Endpoints

The secondary endpoints are:

- Proportion of subjects with ≥10% weight gain at Week 12
- Proportion of subjects with ≥7% weight gain at Week 12
- Change from baseline in waist circumference at Week 12
- Change from baseline in Clinical Global Impression-Severity (CGI-S) score within the ALKS 3831 Group at Week 12

3.8.3.1. Secondary Endpoint: Proportion of Subjects with ≥10% Weight Gain at Week 12 in FAS Population

3.8.3.1.1. Analysis

The analysis will be carried out using a logistic regression model. The logistic regression model will include the treatment group, all randomization strata as factors; and the baseline weight as the covariate. The proportion of subjects with ≥10% weight gain at Week 12 will be derived based on the same complete datasets obtained from the multiple imputation procedure as described in Section 3.8.2.1 for the primary endpoint. Rubin's rule will be used to combine the results from each imputed dataset. Estimated odds ratio, 95% CI and p-value will be reported. Risk difference and 95% CI will be reported as well as the number needed to treat (NNT). NNT is calculated as the absolute value of the inverse of the risk difference, rounded up to the nearest

integer. For subjects who prematurely discontinue the study drug but are willing to come back for weight assessments, only on-treatment weight assessments will be included in the analysis.

In addition, to assess the impact of different cutoffs for weight gain (ie, 5%, 7%, 10%, etc), a plot of cumulative responder distribution at Week 12 for each treatment group will be presented. The horizontal axis will present percent weight gain, and the vertical axis will present the proportion of subjects with at least that amount of weight gain. The analysis will be based on MI for missing data.

3.8.3.1.2. Sensitivity Analysis

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

- **Delta-adjusted pattern mixture model**: To assess the potential impact of missing data due to MNAR, the delta-adjusted Pattern Mixture Model will be conducted to assess the impact of missing data. Proportion of subjects with ≥10% weight gain at Week 12 will be derived based on the same complete datasets obtained from the multiple imputation procedure of the sensitivity analysis of the primary endpoint (Section 3.8.2.2). The logistic regression model will include the treatment group, all randomization strata as factors; and the baseline weight as the covariate. It will be fitted 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.
- **Including on and off-treatment assessments:** To further assess the potential impact of missing data due to MNAR, the approach described in Section 3.8.3.1.1 will be repeated including both on-treatment and off-treatment weight assessments after premature discontinuation of study drug.
- **Additional covariates:** The logistic regression model may also adjust for additional covariates and/or factors, including but not limited to study sites.
- **Average treatment effect:** Average treatment difference across visits through Week 12 between ALKS 3831 and olanzapine will be analyzed.

3.8.3.2. Secondary Endpoint: Proportion of Subjects with ≥7% Weight Gain at Week 12 in FAS Population

The analyses will be the same as proportion of subjects with ≥10% weight gain at Week 12, including relevant sensitivity analyses (Section 3.8.3.1).

3.8.3.3. Secondary Endpoint: Change from Baseline in Waist Circumference at Week 12 in FAS Population

For the change from baseline in waist circumference at Week 12, the analysis will be the same as the primary endpoint, including relevant sensitivity analyses (Section 3.8.2.1 and Section 3.8.2.2).

3.8.3.4. Secondary Endpoint: Change from Baseline in Clinical Global Impression-Severity (CGI-S) score within the ALKS 3831 Group at Week 12

To assess the change from baseline in CGI-S at Week 12 within the ALKS 3831 Group, the MMRM will be used. The model will include treatment, visit, treatment-by-visit interaction and all randomization strata as factors. An unstructured variance-covariance matrix will be used. The least squares (LS) mean, 95% CI and p-value of change from baseline at Week 12 within the ALKS 3831 Group will be reported.

3.8.4. Multiple Comparison/Multiplicity

The study will be claimed positive when the primary analysis of the primary efficacy endpoint is statistically significant. The primary endpoint will be tested at a full α level of 0.05. To control the overall Type I error rate of the primary and secondary endpoints, hierarchical testing will be performed in the following order: (1) percent change from baseline in body weight at Week 12, (2) proportion of subjects with ≥10% weight gain from baseline at Week 12, (3) proportion of subjects with ≥7% weight gain from baseline at Week 12, (4) change from baseline in waist circumference at Week 12, (5) change from baseline in CGI-S score within the ALKS 3831 Group at Week 12.

3.8.5. Examination of Subgroups

Subgroup analyses of the primary and secondary endpoints will be performed for each of the following categories:

- Diagnosis (schizophrenia/schizophreniform disorder, bipolar I disorder)
- Sex (male, female)
- Age (<18 years, ≥18 years)
- Age (<30 years, ≥30 years)
- Race (Black or African American, Non-Black or Non-African American)
- Baseline BMI (<25 kg/m^2, ≥25 kg/m^2)
- Region (US, Non-US)

A forest plot of the estimated treatment effect along with 95% CIs will be provided by the subgroup factors listed above. Only summary statistics for the age <18 years subgroup will be provided due to its limited enrolment expected.

3.8.6. Other Endpoints

Other endpoints include:

- Absolute change from baseline in body weight by visit
- Absolute change from baseline in BMI by visit
- Proportion of subjects in each BMI category (normal weight: <25, overweight: ≥25 to <30, obese: ≥30 kg/m^2) by visit and proportion of subjects with shifts in BMI category from baseline by visit

- Proportion of subjects with ≥5 cm increase from baseline in waist circumference at Week 12
- Body composition endpoints, as assessed by bioimpedance analysis, including:
  - Change from baseline in fat mass by visit
  - Change from baseline in percent body fat by visit
  - Change from baseline in fat mass index by visit
  - Change from baseline in visceral adipose tissue by visit
  - Change from baseline in lean mass by visit
  - Change from baseline in fat free mass index by visit
  - Change from baseline in total skeletal muscle mass by visit
- Percent and absolute change from baseline in fasting lipids (fasting triglycerides, low-density lipoprotein [LDL], high-density lipoprotein [HDL], total cholesterol), fasting glucose, hemoglobin A1c (HbA1c), fasting insulin, and homeostatic model assessment of insulin resistance (HOMA-IR) by visit
- Percent and absolute change from baseline in additional laboratory parameters including apolipoprotein B (Apo B), apolipoprotein A1 (Apo A1), Apo B/Apo A1, high-sensitivity C-reactive protein (hs-CRP), interleukin-6 (IL-6) and tumor necrosis factor α (TNFa) by visit
- Change from baseline in blood pressure (supine systolic and diastolic) by visit
- Proportion of subjects in each blood pressure category (normal, elevated, stage 1 hypertension, stage 2 hypertension) and proportion of subjects shifting from normal or elevated blood pressure at baseline to Stage 1 or 2 hypertension by visit
- Proportion of subjects meeting criteria for metabolic syndrome by visit and proportion of subjects shifting from no metabolic syndrome at baseline to having metabolic syndrome by visit
- Proportion of subjects meeting individual metabolic syndrome component criterion by visit
- Change from baseline in CGI-S score (between and within treatment groups) by visit
- Change from baseline in CGI-S score (between and within treatment groups) by visit and by diagnosis
- CGI-I score by visit
- Change from baseline in IWQOL-Lite score by visit

For all the other endpoints, presentation of analysis results will be limited to estimated treatment effect, 95% CI and/or descriptive statistics.

Absolute change in body weight, BMI, systolic and diastolic blood pressure and all body composition endpoints will be analyzed similarly by ANCOVA model and MI for missing data

as noted in Section 3.8.2.1. The ANCOVA model will include treatment and all randomization strata as factors and baseline value as a covariate. The proportion of subjects with ≥5 cm increase of waist circumference will be analyzed by the logistic regression model and MI method in a similar fashion as the proportion of subjects with ≥10% weight gain, using the same imputed dataset from the primary analysis of change from baseline in waist circumference in Section 3.8.3.3.

The proportion of subjects in each BMI category will be analyzed based on each subject's by-visit assessment status using a logistic regression model including treatment group and randomization strata as factors, and baseline BMI as covariate. The model is based on the MI for missing postbaseline assessments. Rubin's rule is used to combine results from applying the logistic regression model on imputed datasets.

Proportion of subjects with shifts in BMI category by visit will also be analyzed based on each subject's by-visit assessment status using the logistic regression model mentioned above. Shifts are defined as follows:

- Proportion of subjects shifting from BMI <25 kg/m^2 at baseline to BMI ≥25 kg/m^2 by visit
- Proportion of subjects shifting from BMI <30 kg/m^2 at baseline to BMI ≥30 kg/m^2 by visit

The proportion of subjects in each blood pressure category will be analyzed based on the subject's by-visit assessment status using a logistic regression model including treatment group and randomization strata as factors, and baseline systolic and diastolic blood pressure as covariates. The model is based on the MI for missing postbaseline assessments. Rubin's rule is used to combine results from applying the logistic regression model on imputed datasets. Number of responders and proportion for each blood pressure category are the mean number of responders and mean proportion from the imputed datasets, respectively.

Proportion of subjects with shifts in blood pressure category by visit will also be analyzed based on each subject's by-visit assessment status using the logistic regression model described above. Shifts are defined as follows:

- Proportion of subjects shifting from Normal blood pressure at baseline to Stage 1 or Stage 2 Hypertension
- Proportion of subjects shifting from Normal or Elevated blood pressure at baseline to Stage 1 or Stage 2 Hypertension

The blood pressure categories will be defined as follows based on American Heart Association criteria:

- Normal: systolic and diastolic blood pressure of <120 and <80 mmHg, respectively
- Elevated: systolic and diastolic blood pressure of ≥120 to ≤129 and <80 mmHg, respectively
- Stage 1: systolic blood pressure of ≥130 to ≤139 mmHg or diastolic blood pressure of ≥80 to ≤89 mmHg

- Stage 2: systolic blood pressure of ≥140 mmHg or diastolic blood pressure of ≥90 mmHg.

The following laboratory parameters will be analyzed by MMRM model as described in Section 3.8.3.4 . The model will include treatment, visit, treatment-by-visit interaction, all randomization strata as categorical fixed effects; the baseline value will be included as the covariate. An unstructured variance-covariance matrix will be used.

- Percent and absolute change from baseline in fasting lipids (fasting triglycerides, LDL, HDL, total cholesterol), fasting glucose, HbA1c, fasting insulin, and HOMA-IR by visit
- Percent and absolute change from baseline in additional laboratory parameters including Apo B, Apo A1, Apo B/Apo A1, hs-CRP, IL-6 and TNFa by visit.

The proportion of subjects with metabolic syndrome at each visit will be analyzed by logistic regression model including treatment group and randomization strata as factors and baseline metabolic syndrome as covariate.

The proportion of subjects meeting each individual component criterion for metabolic syndrome (listed below) will be analyzed by visit using the same logistic regression model described above with the respective baseline value for each component as the covariate.

Additionally, a similar logistic regression model including treatment and randomization strata as factors will be used to analyze the proportion of subjects shifting from no metabolic syndrome at baseline to developing metabolic syndrome postbaseline. Data are based on observed values with no imputation for missing values.

Metabolic syndrome is defined as meeting at least 3 of the following individual component criteria:

- Waist circumference, cm>88 (women), >102 (men)
- Triglycerides, mg/dL ≥150
- HDL cholesterol, mg/dL <50 (women), <40 (men)
- Glucose, fasting mg/dL ≥100
- Blood pressure, mm Hg ≥130/85

The following endpoints will be analyzed by MMRM model as described in Section 3.8.3.4, as applicable.

- Change from baseline in CGI-S score (between and within treatment groups) by visit
- Change from baseline in CGI-S score (between and within treatment groups) by visit and by diagnosis
- Change from baseline in IWQOL-Lite score by visit

IWQOL-Lite scale is a 31-item self-report measure of obesity-specific quality of life. The IWQOL-Lite provides an overall total score as well as scores on five domains: (1) physical function, (2) self-esteem, (3) sexual life, (4) public distress, and (5) work. The raw IWQOL scores will be transformed as follows (Tessier et al, 2012) before analyzed by the MMRM

method. The transformed scores range from 0 to 100, with 100 representing the best and 0 the most impaired quality of life.

$$\text{Transformed Score} = \frac{\text{maximum theoretical score} - \text{actual score}}{\text{test score range}} \times 100$$

CGI-I score and CGI-I responder status (≤2 versus >2) at each visit will be summarized by treatment group. Descriptive statistics will be presented.

3.9. Safety Analysis

3.9.1. Adverse Events

AEs will be coded by System Organ Class and Preferred Term using the Medical Dictionary for Regulatory Activities (MedDRA®) Version 23.0 or higher. The verbatim term will be included in the AE listings.

An AE (classified by preferred term) will be considered as a treatment-emergent AE (TEAE) if the event is newly occurring or worsening on or after the date of first dose of study drug and within 7 days after the date of last dose of the study drug.

An overview table, including number and percentage of subjects with TEAEs, study-drug-related TEAEs, severe TEAEs, AEs leading to treatment discontinuation (AEDCs), serious AEs (SAEs), study drug-related SAEs, and deaths will be provided.

The number and percentage of subjects reporting TEAEs during the Double-blind Treatment 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 olanzapine 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
  - Experienced by ≥2% of subjects in ALKS 3831 group by preferred term

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 study drug.

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

For the subjects who prematurely discontinue the study drug and return for the bi-weekly visits, all AEs collected after discontinuation of the study drug will be summarized by system organ class, preferred term, and treatment group.

3.9.1.1. Deaths, Serious and Other Significant Adverse Events

The number and percentage of subjects who have SAEs and AEDCs will be summarized by preferred term and treatment group, and will be sorted by decreasing frequency for the ALKS 3831 group.

3.9.1.1.1. AEs of Special Interest

The incidence of a selected subset of relevant AEs in this class of drugs (eg, Movement Disorders (Including Extrapyramidal Sydrome [EPS] and Tardive Dyskinesia), and suicidal ideation and behavior) will be summarized by treatment group and preferred term. The selection of AEs per subset will be based on the preferred term, Standardized MedDRA queries (SMQs) or customized MedDRA queries (CMQs) in Appendix 2 and Appendix 3.

3.9.2. Clinical Laboratory Parameters

Laboratory parameters will be presented in conventional (ie, US) units. Unless specified otherwise, 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 during Double-blind Treatment Period for chemistry and hematology parameters will be summarized by treatment group and by visit.

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

For selected metabolic parameters (fasting total cholesterol, fasting HDL, fasting LDL, fasting triglycerides, fasting glucose, and HbA1c), an analysis of sustained PCS values will also be conducted. The percentage of subjects with sustained PCS values through the treatment period will be summarized. The denominator is all subjects with non-PCS at baseline and at least two postbaseline assessments, and the numerator is the number of subjects who met PCS criteria at the last 2 assessments within the treatment period.

Shift tables for metabolic parameters (fasting total, fasting HDL, fasting LDL cholesterol, fasting triglycerides, fasting glucose, and HbA1c) and liver function tests will be presented. The criteria are summarized in Table 2, Table 3, Table 4 and Table 5.

A listing of subjects who met Hy's Law will be provided. Subjects who met Hy's Law is defined as subjects who meet all of the following three criteria:

1. Alanine aminotransferase (ALT) or aspartate aminotransferase (AST) ≥3× upper limit of normal (ULN)
2. Total bilirubin ≥2× ULN
3. A non-elevated alkaline phosphatase (ALP) <2× ULN

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

Parameter	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	≤40 mg/dL
Cholesterol, HDL Fasting	Decrease ≥20 mg/dL
Cholesterol, LDL Fasting	≥160 mg/dL
Cholesterol, LDL Fasting	Increase ≥30 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
Glucose, Fasting	Increase ≥10 mg/dL
HbA1c	≥ 5.7%
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
Prolactin (Female)	>3 x ULN
Prolactin (Male)	>3 x ULN
Sodium	<130 mmol/L or >150 mmol/L
Triglycerides, Fasting	≥200 mg/dL
Triglycerides, Fasting	Increase ≥50 mg/dL

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

Parameter	Criteria
Hematology	
Hematocrit (Female)	≤32%
Hematocrit (Male)	≤37%
Platelets	<75.0 × 10^3 cells/μL or ≥ 700.0 × 10^3 cells/μL
Leukocytes	≤ 2.8 × 10^3/μL or ≥ 16.0 ×10^3/μL
Neutrophils, Absolute	<1.5 x 10^3/μL <1.0 x 10^3/μL

Abbreviations: HDL= high-density lipoprotein; LDL=low-density lipoprotein; ULN=upper limit of normal.

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

Parameter	Shift Criterion
Total cholesterol (fasting) mg/dL	Normal (<200) to borderline (≥200 and <240)
	Borderline (≥200 and <240) to high (≥240)
	Normal (<200) to high (≥240)
LDL cholesterol (fasting) mg/dL	Normal (<100) to borderline (≥100 and <160)
	Borderline (≥100 and <160) to high (≥160)
	Normal (<100) to high (≥160)
HDL cholesterol (fasting) mg/dL	Normal (≥40) to low (<40)
Triglycerides (fasting) mg/dL	Normal (<150) to borderline (≥150 and <200)
	Borderline (≥150 and <200) to high (≥200)
	Normal (<150) to high (≥200)

Abbreviations: HDL=high-density lipoproteins; LDL=low-density lipoproteins.

Table 3: Shift Category from Baseline to Any Postbaseline in Glucose and HbA1c

Parameter	Shift Criterion
Glucose (fasting) mg/dL	Normal (<100) to impaired (≥100 and <126)
	Impaired (≥100 and <126) to high (≥126)
	Normal (<100) to high (≥126)
HbA1c %[a]	Normal (<5.7%) to borderline (≥5.7% and <6.5%)
	Borderline (≥5.7% and <6.5%) to high (≥6.5%)
	Normal (<5.7%) to high (≥6.5%)

[a] Subjects with HbA1c ≥6.0% were excluded from enrolling in the studies.

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

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

Abbreviations: ULN=upper limit of normal.

Table 5: Prolactin – Shift Categories from Baseline to Any Postbaseline

Prolactin (ng/mL)
Shift from (Low or Normal) to (≥3 × ULN)
Shift from (≥3 × ULN) to (Low or Normal)

Abbreviations: ULN=upper limit of normal.

3.9.3. Vital signs and Electrocardiograms

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 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 6. The number and percentage of subjects with PCS postbaseline 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 postbaseline assessment for the double-blind period. The numerator will be the total number of subjects with available baseline values and at least 1 PCS postbaseline value for double-blind period. A supportive tabular display of subjects with PCS postbaseline values will be provided, including the subject ID number, and baseline and all postbaseline (including non-PCS) values.

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

Table 6: 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. 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 the end of the Double-blind Treatment Period will be presented by treatment group. QTc interval will be calculated using 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 postbaseline 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 non-PCS baseline values and at least 1 postbaseline assessment for the Double-blind Treatment Period. The numerator is the total number of subjects with non-PCS baseline values and at least 1 PCS postbaseline value for the Double-blind Treatment Period. A supportive tabular display of subjects with PCS postbaseline values will be provided, including the subject ID number, study center number, the baseline value, all postbaseline (including non-PCS) values, and change from baseline.

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

Parameter	Criteria
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. 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 (Table 8). 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[a]	Preparatory acts or behavior Aborted attempt Interrupted attempt Actual attempt Complete suicide
Suicidal ideation[a]	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
Non-suicidal self-injurious behavior	Non-suicidal self-injurious behavior

[a] Derived based on responses to individual items listed within the category

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.

3.11. Other Analysis

Analyses to evaluate the potential impact of the covid-19 pandemic to efficacy and safety results will be conducted as outlined in Appendix 1.

4. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE

Interim analysis is not planned.

5. CHANGES IN CONDUCT OR PLANNED ANALYSES FROM THE PROTOCOL

If, after the study has begun, but prior to any unblinding, changes are made to primary and/or key secondary hypotheses, or the statistical methods related to those hypotheses, then the protocol will be amended (consistent with ICH Guideline E-9). Changes to exploratory or other non-confirmatory analyses made after the protocol has been finalized, along with an explanation as to when and why they occurred, will be listed in the Clinical Study Report (CSR) for the study. Post hoc exploratory analyses will be clearly identified in the CSR.

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 3 Schedule of Visits and Assessments).

Scheduled analysis visits during the study period will be the same as the nominal visits collected in the 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 visit data as collected in eCRFs will be included in listings.

Last postbaseline values are defined as the last valid postbaseline values collected for each subject during the Double-blind Treatment Period.

An early termination (ET) visit during the on-treatment period can be mapped to a scheduled visit, if there is no valid value already at that visit. Visit windows are defined for CGI, weight or body composition related endpoints (Table 9), laboratory assessments (Table 10), IWQOL-Lite (Table 11), vital sign (Table 12) and ECG (Table 13). An ET visit that is not mapped will not be summarized in the tables or figures, but will be included in the listing.

Table 9: Visit Window for Efficacy and Safety Assessments for CGI, Body Weight, Waist Circumference and Body Composition

Analysis Visit to be Mapped to	Target Study Week	Target Visit Day[a]	Visit Window
Visit 3	Week 1	Day 8	[2, 11]
Visit 4	Week 2	Day 15	[12, 21]
Visit 5	Week 4	Day 29	[22, 35]
Visit 6	Week 6	Day 43	[36, 49]
Visit 7	Week 8	Day 57	[50, 63]
Visit 8	Week 10	Day 71	[64, 77]
Visit 9	Week 12	Day 85	[78, 91]

[a] Visit Day is calculated as date of visit – date of the first dose of study drug + 1 day.

Table 10: Visit Window for Laboratory Assessments

Analysis Visit to be Mapped to	Target Study Week	Target Visit Day	Visit Window
Visit 4	Week 2	Day 15	[2, 50]
Visit 9	Week 12	Day 85	[51, 91]

Table 11: Visit Window for IWQOL-Lite

Analysis Visit to be Mapped to	Target Study Week	Target Visit Day	Visit Window
Visit 5	Week 4	Day 29	[2, 57]
Visit 9	Week 12	Day 85	[58, 91]

Table 12: Visit Window for Vital Sign

Analysis Visit to be Mapped to	Target Study Week	Target Visit Day	Visit Window
Visit 4	Week 2	Day 15	[2, 28]
Visit 6	Week 6	Day 43	[29, 63]
Visit 9	Week 12	Day 85	[64, 91]

Table 13: Visit Window for ECG

Analysis Visit to be Mapped to	Target Study Week	Target Visit Day	Visit Window
Visit 1	Screening	Day [-30,-1]	[-50, 0]
Visit 9	Week 12	Day 85	[64, 91]

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 a completely missing stop date, medication will be assumed to be ongoing.

6.3. Handling of Safety Data

All efforts should be made to obtain any 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, 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 14: 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 GRAPHS (TLGS)

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

10. REFERENCES

Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics.* 1997;53(3):983-997.

Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharm Stat.* 2013;12(6):337-347. doi: 10.1002/pst.1549.

Tessier A, Zavorsky GS, Kim dJ, Carli F, Christou N, Mayo NE. Understanding the Determinants of Weight-Related Quality of Life among Bariatric Surgery Candidates. *J Obes.* 2012;2012:713426. doi: 10.1155/2012/713426.

APPENDIX 1. COVID-19 ANALYSIS

Pandemic has resulted in certain disruptions in the study conduct since early 2020, including site closure, delayed study drug supply, subjects discontinuation and missing assessments. These impacts were limited in scope upon review of the protocol deviations, and also confounded by the initiation of new regions. Sites in Ukraine and Russia were first initiated in March 2020 and July 2020. Therefore, only descriptive statistics will be provided to explore the potential impact of the pandemic on the primary and secondary objectives of the study. The following subgroup analysis (before vs during pandemic) will be conducted:

- Disposition
- Demographics and Baseline Characteristics
- Protocol Deviation
- Primary Efficacy Endpoint
- Key Secondary Endpoints
- TEAE
- SAE

Subjects enrolled before March 1, 2020 will be grouped as before pandemic, and subjects enrolled after this date will be grouped as during pandemic. While the pandemic has started in various regions a few weeks apart, a single cutoff date will be applied, due to the relatively slow enrollment at the time (about 4 subjects/month) and short treatment duration.

APPENDIX 2. ADVERSE EVENTS ASSOCIATED WITH MOVEMENT DISORDERS (INCLUDING EXTRAPYRAMIDAL SYNDROME AND TARDIVE DYSKINESIA)

Preferred Term (SMQ [#20000095] of Extrapyramidal Syndrome)	Preferred Term Code
Abnormal involuntary movement scale	10075002
Action tremor	10072413
Akathisia	10001540
Akinesia	10001541
Athetosis	10003620
Ballismus	10058504
Blepharospasm	10005159
Bradykinesia	10006100
Bradyphrenia	10050012
Buccoglossal syndrome	10006532
Chorea	10008748
Choreoathetosis	10008754
Chronic tic disorder	10076661
Cogwheel rigidity	10009848
Complex tic	10076663
Dopamine dysregulation syndrome	10067468
Drooling	10013642
Dyskinesia	10013916
Dyskinesia neonatal	10013922
Dyskinesia oesophageal	10013924
Dysphonia	10013952
Dystonia	10013983
Dystonic tremor	10073210
Emprosthotonus	10014566
Extrapyramidal disorder	10015832
Facial spasm	10063006
Fine motor skill dysfunction	10076288
Freezing phenomenon	10060904
Gait disturbance	10017577
Gait inability	10017581
Grimacing	10061991

Preferred Term (SMQ [#20000095] of Extrapyramidal Syndrome)	Preferred Term Code
Hyperkinesia	10020651
Hyperkinesia neonatal	10020652
Hypertonia	10020852
Hypertonia neonatal	10048615
Hypokinesia	10021021
Hypokinesia neonatal	10021022
Hypokinetic dysarthria	10082243
Laryngeal tremor	10078751
Laryngospasm	10023891
Meige's syndrome	10027138
Micrographia	10057333
Mobility decreased	10048334
Motor dysfunction	10061296
Movement disorder	10028035
Muscle contractions involuntary	10028293
Muscle rigidity	10028330
Muscle spasms	10028334
Muscle spasticity	10028335
Muscle tightness	10049816
Muscle tone disorder	10072889
Muscle twitching	10028347
Musculoskeletal stiffness	10052904
Oculogyric crisis	10030071
Oesophageal spasm	10030184
On and off phenomenon	10030312
Opisthotonus	10030899
Oromandibular dystonia	10067954
Oropharyngeal spasm	10031111
Parkinsonian crisis	10048868

Preferred Term (SMQ [#20000095] of Extrapyramidal Syndrome)	Preferred Term Code
Parkinsonian gait	10056242
Parkinsonian rest tremor	10056437
Parkinsonism	10034010
Parkinsonism hyperpyrexia syndrome	10071243
Parkinson's disease	10061536
Parkinson's disease psychosis	10074835
Pharyngeal dyskinesia	10070912
Pharyngeal dystonia	10081226
Pleurothotonus	10035628
Postural reflex impairment	10067206
Postural tremor	10073211
Posture abnormal	10036436
Posturing	10036437
Propulsive gait	10082328
Protrusion tongue	10037076
Provisional tic disorder	10076694
Psychomotor hyperactivity	10037211
Rabbit syndrome	10068395
Reduced facial expression	10078576
Respiratory dyskinesia	10057570
Resting tremor	10071390
Restlessness	10038743
Risus sardonicus	10039198
Secondary tic	10076702
Spasmodic dysphonia	10067672
Tardive dyskinesia	10043118
Tic	10043833
Tongue spasm	10043981
Torticollis	10044074
Torticollis psychogenic	10044076
Tremor	10044565
Tremor neonatal	10044575
Trismus	10044684
Uvular spasm	10050908

Preferred Term (SMQ [#20000095] of Extrapyramidal Syndrome)	Preferred Term Code
Walking disability	10053204
Writer's cramp	10072249

Note: The list is based on MedDRA version 24.0 and will be updated based on the dictionary or relevant SMQ updates at the time of the analysis.

APPENDIX 3. ADVERSE EVENTS ASSOCIATED WITH SUICIDAL IDEATION AND BEHAVIOR

Preferred Term	Preferred Term Code
Assisted suicide	10079105
Columbia suicide severity rating scale abnormal	10075616
Completed suicide	10010144
Depression suicidal	10012397
Intentional overdose	10022523
Intentional self-injury	10022524
Overdose	10033295
Poisoning deliberate	10036000
Self-injurious ideation	10051154
Suicidal behaviour	10065604
Suicidal ideation	10042458
Suicide attempt	10042464
Suicide threat	10077417
Suspected suicide	10082458
Suspected suicide attempt	10081704

Note: The list is based on MedDRA version 24.0 and will be updated based on the dictionary or relevant SMQ updates at the time of the analysis.