STATISTICAL ANALYSIS PLAN

Based on:
Protocol Version 4.0, Dated 22-MAY-2020

SPONSOR:	Alkahest, Inc.
PROTOCOL NUMBER:	AKST4290-211
PROTOCOL TITLE:	A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of AKST4290 in Subjects with Parkinson's Disease on Stable Dopaminergic Treatment
ORIGINAL PROTOCOL DATE:	10-JAN-2019

Created by: [redacted]

SAP Approval

By signing the following, I agree to the contents in the Statistical Analysis Plan and its associated attachments. Once the SAP has been signed, the analyses and programming of the TLFs based upon this document can proceed. Any modifications to the SAP and TLFs made after signing may result in a change order.

Approved by:

Name	Title	Signature	Date
[redacted]	[redacted]	[redacted]	4/7/2021
[redacted]	[redacted]	[redacted]	07Apr2021

Abbreviations

Abbreviation	Definition
ADLB	Laboratory analysis dataset
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
ALR	Alternating logistic regression
ANCOVA	Analysis of covariance
ATC	Anatomical therapeutic chemical
AST	Aspartate transaminase
BMI	Body mass index
CI	Confidence interval
CISI-PD	Clinical Impression of Severity Index – Parkinson's Disease
CM	Concomitant medication
CSR	Clinical study report
CSF	Cerebrospinal fluid
CWS	Comfortable walking speed
DBP	Diastolic blood pressure
ECG	Electrocardiogram
EOT	End of treatment
FACS	Fluorescence-activated cell sorting
FCS MI	Fully conditional specification
FWS	Fast walking speed
GEE	Generalized Estimating Equations
ITT	Intention-to-Treat
LS	Least-square
MDS-UPDRS	Movement Disorder Society's Unified Parkinson's Disease Rating Scale
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intention-to-Treat
MMRM	Mixed-effect model for repeated measures
MoCA	Montreal Cognitive Assessment
PD	Parkinson’s disease
PDQ-39	Parkinson’s Disease Quality of Life Questionnaire-39
PP	Per-Protocol
PMM	Predictive mean matching
PT	Preferred term
QTc	Corrected QT interval
QTcF	QT interval corrected by the Fridericia formula
SAE	Serious adverse event
SAP	Statistical analysis plan
SBP	Systolic blood pressure
SE	Standard error

SE-ADL	Schwab and England Activities of Daily Living Scale
SOC	System organ class
S-STS	Sheehan-Suicidality Tracking Scale
STD	Standard deviation
TEAE	Treatment-emergent adverse events
ULN	Upper limit of normal
WHO	World Health Organization
WOCBP	Women of childbearing potential

Table of Contents

1 Introduction

This Statistical Analysis Plan (SAP) describes the planned statistical analyses to be performed for data from Protocol AKST4290-211, "A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of AKST4290 in Subjects with Parkinson's Disease (PD)on Stable Dopaminergic Treatment". This SAP was created using Clinical Protocol AKST4290-211Version 4.0 dated 22-MAY-2020.The table and listing shells will be provided in separate files as attachments to this SAP.

2 Objectives

2.1 Primary Objective

To assess the effects of AKST4290 on motor function in the practically defined off medication state, defined as ≥12 hours off levodopa, in subjects with PD.

2.2 Secondary Objectives

To assess the safety of AKST4290 in subjects with PD as well as the potential effects on clinical function and activities of daily living.

2.3 Exploratory Objectives

- To analyze pharmacokinetic parameters following twice daily dosing with AKST4290.
- To assess flow cytometry, pharmacogenomic, and biomarker evaluations conducted on blood and plasma samples.
- To assess bradykinesia, tremor, general activity, and sleep using a wearable device
- To evaluate fecal samples obtained at screening and following treatment for potential microbiome changes in consenting subjects.

3 Endpoints

3.1 Primary Endpoint

Change from Baseline (Day 1) in motor function during the practically defined off-medication state, defined as ≥12 hours off levodopa, at Week 12 (Day 84) as measured by Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS), Part 3.

3.2 Secondary Endpoints

- Incidence of treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs) identified by the Medical Dictionary for Regulatory Activities (MedDRA) preferred term (PT) and grouped by MedDRA System Organ Class (SOC).
- Incidence of abnormalities or clinically-significant changes from Baseline in laboratory test data, vital sign measurements, and electrocardiograms (ECGs).

- Change from Baseline (Day 1) in clinical function, motor function, and activities of daily living at Week 12 (Day 84) during the on-medication state as assessed by:
  - MDS-UPDRS Parts 1-4
  - Montreal Cognitive Assessment (MoCA)
  - Schwab and England Activities of Daily Living (SE-ADL) Scale
  - Clinical Impression of Severity Index – PD (CISI-PD)
  - PD Quality of Life Questionnaire-39 (PDQ-39)
  - Sheehan-Suicidality Tracking Scale (S-STS)
  - 10-meter timed walk (will also be assessed in the off-medication state)
  - Hauser 3-Day Patient Diary

3.3 Exploratory Endpoints

- Changes in concentration of AKST4290 or its major metabolites in plasma at various time points.
- Flow cytometric analyses and evaluation of pharmacogenomic characteristics and biomarkers in blood and plasma samples.
- Changes in bradykinesia, tremor, general activity, and sleep as measured by an FDA-cleared wearable sensor device.
- In consenting subjects (optional): characterization of the composition and function of the fecal gut microbiome.

4 Study Overview

4.1 Study Design

This is a phase 2, randomized, double-blind, placebo-controlled study to assess the efficacy and safety of AKST4290 in subjects with PD.

The study will enroll approximately 120 subjects between 50 and 80 years of age with a diagnosis of PD on dopaminergic treatment from approximately 30 sites globally. Enrolled subjects will be randomized in a 1:1 ratio to active treatment in Arm 1 (approximately 60 subjects) or placebo in Arm 2 (approximately 60 subjects). Subjects will take 400 mg of AKST4290 or placebo orally, twice per day (b.i.d.) for a total daily dose of 800 mg. The study duration for the subjects will be approximately 18weeks including screening, 12 weeks of treatment, and 4weeks of follow-up. A detailed overview of the timing of study assessments and interventions is presented in the Schedule of Events section below.

4.2 Treatment Assignment and Dose Regimen

A subject number will be assigned once the subject has signed the informed consent form. Enrolled subjects that have met all eligibility criteria will be randomized to AKST4290 (Arm 1: active dosing) or placebo (Arm 2: placebo dosing). The study drug will be self-administered orally b.i.d. (2 x 200 mg per dose), approximately 12 hours

apart, for a total daily dose of 800 mg. Subjects will receive study drug or placebo for a total of 12 weeks, followed by 4weeks of follow-up.

4.3 Statistical Hypothesis

The null and alternative hypotheses for primary endpoint analysis are:

$$H0: \mu C = \mu T \text{ vs. } HA: \mu C \neq \mu T.$$

The null hypothesis is that there is no difference in mean change (μ) from baseline motor function during the practically defined off-medication state at Week 12 - as measured by the MDS-UPDRS Part 3 - between the control (C) and treated (T) groups. The alternative hypothesis is that there is a difference in mean change from baseline motor function during the practically defined off-medication state at Week 12 between the control and treated groups. P-values for testing whether the differences are equal to zero will be provided. P-values will be compared against an alpha significance level of 0.1.

4.4 Sample Size Justification

A total of approximately 120 subjects will be randomized in a 1:1 ratio to active treatment (approximately 60 subjects) or placebo (approximately 60 subjects), with the intent of obtaining ~108 evaluable subjects.

[redacted]

4.5 Randomization/Unblinding

This is a randomized and double-blind study.

4.5.1 Randomization

A randomization list, permuted blocks and stratified by sex (Male and Female) was generated by an outside vendor and is outside the scope of this SAP.

4.5.2 Unblinding

The study blind can be broken for safety reasons if the information is required for the management of SAEs, severe AEs, or pregnancies. The Investigator can obtain the treatment allocation for their subject through the web-based

randomization system (IRT). Else, the blinded code for the trial will be broken only after all subject data has been recorded and verified and the database locked.

4.6 Interim Analyses

Safety will be monitored on an ongoing basis, and a Safety Evaluation Meeting may be triggered (see protocol section 8.5). Any such ad hoc interim safety analysis performed will be outside the scope of this SAP.

4.7 Study Assessment Time Points

The study consists of 8 protocol-specified visits which will be assessed as nominal visits from an analysis perspective:

- Visit 1, Screening Visit (Day -14 to -7)
- Visit2, Baseline (Day 1): Subjects will be randomized after eligibility has been confirmed, study drug/placebo dispensed.
- Visit 3 (Day 14 ± 2)
- Visit 4 (Day 28 ± 2)
- Visit 5 (Day 56 ± 2)
- Visit 6 – End of Treatment (EOT) (Day 84 ± 2)
- Visit 7 (Day 98 ± 3 or 14 days after Early Termination)
- Safety Follow-Up (Day 114 ± 2 or 30 days after Early Termination): a documented phone call to record AEs and concomitant medications (CMs).
- Early termination: In cases of early termination, if a subject has received at least 1 dose (400 mg) of AKST4290 or placebo, the site should try to perform all assessments scheduled at the Visit 6 - EOT visit as well as the subsequent follow up visits unless the subject has withdrawn consent.

4.8 Schedule of Events

Visit Number Day Window (days)[a] Week	1 -14 to - 7	2 1	3 14 ±2 2	4 28 ±2 4	5 56 ±2 8	6/EOT 84 ±2 12	7 98 ±3 14	Phone Call 114 ±2 15
Informed Consent[1]/Optional Consent: fecal sample	X							
Demographics	X							
Medical history	X							
Inclusion/exclusion criteria	X	X						
Randomization to AKST4290 or placebo		X						
Modified Hoehn and Yahr	X							
Provide PDQ-39 to be completed prior to next visit	X	X	X	X	X	X		
Provide Hauser 3-Day Diary for next visit	X				X			
Confirmation of off-medication state		X		X		X	X	
Full physical examination	X							
Targeted physical examination		X	X	X	X	X	X	
Vital signs	X	X	X	X	X	X	X	
Laboratory tests	X	X	X	X	X	X	X	
Pregnancy testing[2]	X	X	X	X	X	X	X	
12-lead ECG	X			X		X		
MDS-UPDRS Part 3 – Off-medication		X		X		X	X	
10-meter timed walk – Off-medication		X		X		X	X	
Administration of dopaminergic medication		X		X		X	X	
MDS-UPDRS Part 1-4 – On-medication		X	X	X	X	X	X	
10-meter timed walk – On-medication		X	X	X	X	X	X	
MoCA		X				X		
SE-ADL and CISI-PD		X	X	X	X	X	X	
S-STS		X				X		

Collect PDQ-39		X	X	X	X	X	X	
Collect Hauser 3-Day Patient Diary		X				X		
Wearable sensor device[3]	X	X		X	X	X		
Fecal sample (optional)	X					X		
Dispense 5-week supply of AKST4290/placebo		X		X	X			
Study agent administration[4]		X	X	X	X	X		
Study agent accountability		X	X	X	X	X		
Pharmacokinetics blood sample[5,6]		X	X	X	X	X	X	
Biomarker plasma aliquots[5,6]		X				X	X	
FACS/CBC blood sample[5]		X				X		
Pharmacogenomics blood sample (and optional DNA banking sample)		X						
Adverse events	X	X	X	X	X	X	X	X
Concomitant medications[7]	X	X	X	X	X	X	X	X
Trial completion								X

Notes:
a. Window extensions and missed protocol assessments may be permitted to reduce the risk of COVID-19 exposure. Any deviation to the protocol to reduce the risk of COVID-19 exposure will be captured as a "Protocol Deviation related to COVID-19" to categorize the anticipated increase in protocol deviations due to the pandemic. These measures are temporary, and will be repealed as soon as the situation (e.g., governmental rules, benefit/risk assessment for the trial, etc.) allows (see Section 13.3).
1. All patients must sign an informed consent consistent with ICH-GCP guidelines prior to any trial related procedures, which includes medication washouts and restrictions.

2. Pregnancy testing may be performed using serum or urine. Test should be performed at the site and results reviewed. Positive results should be confirmed by the central lab.
3. An at home wearable watch will be provided in clinic on Visits 1, 4 and 5. The watch will be collected at Visits 2, 5, and 6.
4. Study agent (AKST4290/placebo) will be self-administered in the clinic under supervision of study personnel during every visit of the treatment period (Visits 2-6). Training on study agent administration will be conducted at Visit 2. Study agent administration will be performed after all assessments during Visit 2 and at visit start at remaining visits (after collection of initial PK sample 15 minutes prior to study agent administration, if applicable). Please Section 7.3 for timing dependencies.
5. The PK sample is drawn ~15 minutes prior to study agent (AKST4290/placebo) administration at Visits 2-6. Biomarker and FACS/CBC samples are drawn ~15 minutes prior to study agent (AKST4290/placebo) administration at Visits 2 and 6. The PK, biomarker, and FACS/CBC samples are drawn ~1 hour after study agent administration at Visit 2. Additional (optional) PK samples are drawn at ~ 2 hours after study agent administration at Visit 2. Additional (optional) PK samples are drawn at ~ 1 hour and ~ 2 hours after study agent administration at Visits 4 and 6. For additional information regarding timing, see Section 17.12.
6. Biomarker plasma aliquots are drawn from the PK blood sample.
7. [redacted]

5 Statistical Methodology

5.1 General Considerations

This section presents the statistical approaches that are anticipated for the analysis of the study data. These approaches may at times require modifications due to unanticipated features of the data. Deviations from analyses summarized in this document will be noted in the clinical study report (CSR).

All statistical analyses will be performed using SAS® Version 9.4 or higher, unless otherwise noted. All summary statistics will be descriptive unless noted otherwise. Descriptive summaries will include mean, standard deviation (STD), median, and range for continuous variables and counts, and percentages for categorical variables. Two-sided 95% confidence intervals (CIs) will be provided for the means and percentages as needed. For key outcome measures, the difference between the treatment arms and the 95% CI of the difference will be computed, as necessary. Presentations will be by treatment group, unless otherwise noted.

In general, all summary tables will be supported by a relevant subject data listing including all subjects who are randomized. The listings will include all data collected, and will be sorted by subject ID, and actual visit date, as applicable, unless otherwise noted.

Concomitant medications will be coded by the World Health Organization (WHO)Drug Dictionary and adverse events and medical history will be coded using the MedDRA. The dictionary versions will be documented in the Data Management Plan for the study.

Study day 1 is the first day of study treatment. Study day is the day relative to the study day 1. Baseline is the last non-missing valid value prior to study treatment.

5.2 Study Populations for Analysis

5.2.1 Intention-to-Treat (ITT) Population

All subjects who were randomized into the study will be included in the ITT population. Analysis will be done according to the treatment subjects were randomized to.

5.2.2 Modified Intention-to-Treat (mITT)/Evaluable Population

The mITT/evaluable population will be a subset of the ITT population that only includes subjects with non-missing baseline and Week 12 (Day 84) values for off-medication motor function as measured by the MDS-UPDRS Part 3. Analysis will be done according to the treatment subjects were randomized to.

5.2.3 Safety Evaluable Population

The safety population will consist of all subjects who received at least one dose of the study medication. All safety endpoints will be summarized for the safety population. Analysis will be done according to the actual treatment subjects received.

5.2.4 Per-Protocol (PP) Population

The PP population will be a subset of the ITT population that only includes subjects who have complete baseline and Week 12 assessments without any protocol deviation(s)which would interfere with the assessment of efficacy.

Such protocol deviations are departures from the approved protocol relating to the conduct of the study which may affect the rights, safety, and/or wellbeing of study participants or the study outcomes or data quality. A listing of study subjects with protocol deviations which would preclude inclusion in the PP population will be provided by the Sponsor prior to database lock.

The PP population is a supportive analysis population for the primary and secondary efficacy endpoints, and will be used in sensitivity analyses as needed. Analysis will be done according to the treatment subjects were randomized to.

5.3 Subject Disposition

All subjects who signed the informed consent will be accounted for in Subject Disposition. The following disposition information will be summarized by the number and percentages of subjects:

- Subjects who signed the informed consent, and who comprise the ITT, mITT/Evaluable, Safety, and PP populations
- Subjects who complete the study or withdraw prematurely along with the reason for discontinuation

A listing of screen failure subjects, including age, sex, race, ethnicity, education, and the inclusion/exclusion criteria not met, will be provided.

A listing of subjects excluded from each of the analysis populations, including population flags (Safety/mITT/Evaluable) and the reason for exclusion, will be provided.

5.4 Demographics and other Baseline Characteristics

5.4.1 Demographics

Sex, age (in years), race, ethnicity, education, and baseline body mass index (BMI) (kg/m^2) will be summarized by treatment group for the ITT and mITT/Evaluable populations. BMI (kg/m^2) will be calculated as: 10,000*weight (kg) / the square of height (cm). If height is collected in inches, it will be converted to cm by dividing the value by 2.54. If weight is collected in pounds, it will be converted to kg by dividing the value by 2.205.

5.4.2 Baseline Characteristics

Baseline characteristics will be summarized by treatment group for the ITT and mITT/Evaluable populations.

Baseline characteristics to be summarized are:

- Disease Duration
- Modified Hoehn and Yahr
- MDS-UPDRS Parts 3 in the off-medication state

5.4.3 Medical History

The number and percentage of subjects with any medical history will be summarized for the ITT population by treatment group, SOC, and PT using the MedDRA dictionary.

The following will be provided in listings only:

- Family history in biological parents, siblings, and offspring of PD, Alzheimer's disease (AD), or other dementias or neurological disorders
- Cause of parental death (if not living)
- History of tobacco and substance use
- Social history and exercise/activity level

5.5 Protocol Deviations

Protocol deviations will be summarized by deviation category, major vs minor and treatment group for the ITT population, with flags for the mITT/Evaluable, Safety, and PP populations.

5.6 Methods for Handling Missing Data

In order to minimize missing data, every effort will be made to obtain required data for all randomized subjects at each scheduled evaluation. However, in the event that more than 5% of data are missing, imputation methods will be used for analysis of the primary endpoint only. The imputation will be carried out in SAS version 9.4 or later using PROC MI.

5.7 Efficacy Analyses

Efficacy analyses will evaluate the effects of AKST4290 on motor function in the off-medication state, as well as the potential effects on clinical function and activities of daily living. Efficacy endpoints will be evaluated using the mITT/Evaluable and PP populations.

5.7.1 Primary Efficacy Endpoint

The primary endpoint is change from baseline at Week 12 (Day 84) in the off-medication motor function as measured by the MDS-UPDRS Part 3. The main analysis for the primary efficacy variable will be analyzed using mITT subjects.

The MDS-UPDRS (Fahn, 1987; Goetz, 2008) was developed to incorporate elements from existing scales to provide a comprehensive but efficient and flexible means to monitor PD-related disability and impairment. The MDS-UPDRS has four components (Part 1, Mentation, Behavior, and Mood; Part 2, Activities of Daily Living; Part 3, Motor; Part 4, Complications). The rating for each item is from 0 (normal) to 4 (severe). The score for each Part is obtained from the sum of the corresponding item scores.

The null and alternative hypotheses for the primary endpoint are:

$$H0: \mu C = \mu T \text{ vs. } HA: \mu C \neq \mu T.$$

The null hypothesis states that there is no difference in mean change (μ) from baseline motor function during the practically defined off-medication state at Week 12 - as measured by the MDS-UPDRS Part 3 - between the control (C) and treated (T) groups. The alternative hypothesis states that there is a difference in mean change from baseline motor function during the practically defined off-medication state at Week 12 between the control and treated groups.

The MDS-UPDRS Part 3 is administered in the off-medication state at Baseline (Study Day 1) and Weeks 4 (Day 28), 12 (Day 84), and 14 (Day 98). The sum of corresponding items for section 3 will be summarized descriptively at each scheduled visit. Additionally, the number and percentage of subjects with dyskinesia present during the evaluation, and whether that dyskinesia interfered with ratings, will be presented.

A mixed-effect model for repeated measures (MMRM) will be employed to analyze the change from baseline scores at each visit over time up to Week 12. The model will include the fixed effects of treatment (stratified, by sex), baseline of the endpoint, visit, and treatment by visit interaction. An unstructured covariance structure will be used to model the within subject error. Least-square (LS) means, standard error (SE), and 95% CI of the LS mean will be provided for each treatment arm at each visit. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated. P-value for testing whether the difference is equal to zero will be provided. The p-value at Week 12 will be compared against the alpha significance level of 0.1.

5.7.1.1 Sensitivity Analyses

The main analysis of the primary endpoint using the above MMRM will use all available data. To check the robustness of the results, the following sensitivity analyses will be performed:

- Analysis of covariance (ANCOVA) model on the observed change from baseline value at Week 12. The model includes treatment as the fixed effect and baseline of the endpoint as a covariate.
- The same MMRM model on data from the PP subjects.
- The same MMRM model on data from the ITT subjects.
- Depending on the status of missing data among the mITT population, multiple imputations may be employed to handle missing data and apply the MMRM on the imputed datasets. See Methods for Handling Missing Data section (section 5.6) for details on imputation methods.

5.7.2 Secondary Efficacy Endpoints

Secondary efficacy endpoints will be analyzed using the mITT/Evaluable, ITT, and PP populations. Please see Safety Analysis section (Section5.8) for information regarding analysis of safety endpoints.

Secondary efficacy endpoints include the following:

Change from Baseline (Day 1) in clinical function, motor function, and activities of daily living at Week 12 (Day 84) during the on-medication state as assessed by:

- MDS-UPDRS Parts 1-4
- Montreal Cognitive Assessment
- Schwab and England Activities of Daily Living Scale
- Clinical Impression of Severity Index – PD
- PD Quality of Life Questionnaire-39
- Sheehan-Suicidality Tracking Scale
- 10-meter timed walk (will also be assessed in the off-medication state)
- Hauser 3-Day Patient Diary

5.7.2.1 MDS-UPDRS Parts 1-4

The MDS-UPDRS Parts 1-4 is administered in the on-medication state at Baseline and Weeks 2, 4, 8, 12, and 14. The scores for each section (i.e. sum of corresponding items for each section) will be summarized descriptively at each scheduled visit. Additionally, the difference in the MDS-UPDRS Part 3 between the on-medication and off-medication state will be reported.

To evaluate change from baseline, the same MMRM model used to evaluate the primary endpoint will be employed to produce LS means, SE, and 95% CIs of the LS means for each treatment arm at each scheduled visit. The unadjusted mean change from baseline, and the difference in LS

means between treatment arms and 95% CI of the difference will also be calculated.

5.7.2.2 Montreal Cognitive Assessment

The MoCA (Nasreddine, 2005) is a screening test used to assess attention and concentration, executive function, memory, language, visuo-constructional skills, conceptual thinking, calculations, and orientation. The total possible score is 30 points with a score of 26 or more considered normal.

The MoCA is administered at Baseline and Week 12 in the on-medication state. The total score will be summarized at each scheduled time point. To evaluate change from baseline, an analysis of covariance (ANCOVA) model will be used to produce the LS mean, SE, and 95% CI of the LS means for each treatment arm. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated. A summary of each cognitive domain will be provided as well.

5.7.2.3 Schwab and England Activities of Daily Living Scale

The SE-ADL evaluates patients' perceptions of global functional capacity and dependence (Schwab, 1968). Scoring is expressed in terms of percentage, in 10 steps from 100 to 0 (100%, normal status; 0%, bedridden with impaired vegetative functions), so that the lower the score, the worse the functional status.

Scores are collected at Baseline, and Weeks 2, 4, 8, 12, and 14. Scores will be summarized descriptively at each scheduled time point (i.e. n, %) by treatment group. A Generalized Estimating Equations (GEE) for alternating logistic regression (ALR) with an exchangeable working correlation structure will be employed to analyze change from baseline. The model will include the fixed effects of treatment, baseline of the endpoint, visit, and treatment by visit interaction. The common odds ratio and 95% CI for odds of higher versus lower SE-ADL scores between treatment groups will be presented. Mean scores at each timepoint, and change from baseline, will also be presented.

5.7.2.4 Clinical Impression of Severity Index – PD

The CIS-IPD is a severity index formed by four items (motor signs, disability, motor complications, and cognitive status), rated 0 (not at all) to 6 (very severe or completely disabled) (Martinez-Martin, 2006). A total score is calculated by summing the item scores.

The CISI-PD is administered in the on-medication state at Baseline, and Weeks 2, 4, 8, 12, and 14. The total score will be summarized

descriptively at each scheduled visit. To evaluate change from baseline in the total score, the same MMRM model used to evaluate the primary endpoint will be employed to produce LS means, SE, and 95% CIs of the LS means for each treatment arm at each scheduled visit. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated. A summary of each parameter will be presented as well.

5.7.2.5 PD Quality of Life Questionnaire-39

The PDQ-39 is a self-administered questionnaire of 39 questions relating to 8 key areas of health and daily activities, including both motor and non-motor symptoms (Peto, 1998). The eight dimensions include: mobility, activities of daily living, emotional well-being, stigma, social support, cognition, communication, and bodily discomfort. It is scored on a scale of 0-100 with lower scores indicating better health and high scores indicating more severe symptoms.

The PDQ-39 is collected at Baseline and Weeks 2, 4, 8, 12, and 14. The total scores for each of the eight dimensions and the overall total score will be summarized descriptively at each scheduled visit. To evaluate change from baseline, the same MMRM model used to evaluate the primary endpoint will be employed to produce LS means, SE, and 95% CIs of the LS means for each treatment arm at each scheduled visit. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated.

5.7.2.6 Sheehan-Suicidality Tracking Scale

The S-STS was developed to provide a brief but efficient instrument for use in assessing change in suicidal ideation and behavior while providing a comprehensive description of suicidal ideation and behavior (Sheehan, 2014). The standard version of the S-STS is a 16-item scale that assesses the seriousness of suicidality phenomena on a Likert-type scale (0-4) ranging from “not at all” (0) to “extremely” (4). It also assesses the frequency of key phenomena and the overall time spent in suicidality.

The S-STS is collected at Baseline and Week 12. Total score will be summarized descriptively at each scheduled time point. To evaluate change from baseline, an analysis of covariance (ANCOVA) model will be used to produce the LS mean, SE, and 95% CI of the LS means for each treatment arm. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated.

5.7.2.7 10-Meter Timed Walk

The 10-meter walk test is a commonly used tool for assessing gait speed in individuals with gait limitations (Lang, 2016). Gait speed is positively correlated with the amount of community ambulation and quality of life, and it is an important measure of mobility in individuals with PD.

The 10-meter timed walk is conducted in the on-medication state at Baseline and Weeks 2, 4, 8, 12, and 14, as well as the off-medication state at Baseline and Weeks 4, 12, and 14. Based on two consecutive measurements, average comfortable walking speed (CWS) (m/s) and fast walking speed (FWS)(m/s) will be calculated for each subject at each scheduled visit. CWS and FWS will then be summarized descriptively at each scheduled visit by treatment arm. The difference in both the CWS and FWS between the on-medication and off-medication state will also be reported.
To evaluate change from baseline, the same MMRM model used to evaluate the primary endpoint will be employed to produce LS means, SE, and 95% CIs of the LS means for each treatment arm at each scheduled visit. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated.

5.7.2.8 Hauser 3-Day Patient Diary

The Hauser Patient Diary was developed to assess functional status over a period of time in patients with motor fluctuations and dyskinesia (Hauser, 2000). It is a self-completed reference diary designed to separate dyskinesia that had a negative impact on patient-defined functional status from dyskinesia that did not. With this diary, the effect of an intervention can be expressed as the change in off-medication time and the change in on-medication time with troublesome dyskinesia (bad time). The sum can be used as an outcome variable and compared to baseline or across groups.

The Hauser 3-day patient diary is collected at Baseline and Week 12. Mean “bad time” (in hours) over the three days of collection will be summarized descriptively at both time points. Based on the research conducted by Hauser (2000, 2004), bad time is defined as the sum of off time and on time with troublesome dyskinesia. Mean “good time”, defined as on time without dyskinesia plus on time with non-troublesome dyskinesia, will also be summarized descriptively. Additionally, mean % day spent in “bad time” and “good time” will be summarized descriptively. Mean % day spent in “bad time” or “good time” is the mean number of hours spent in bad time or good time, divided by the mean number of waking hours over the three-day period.

To evaluate change from baseline in bad time, good time, mean % day spent in bad time, and mean % day spent in good time, an analysis of covariance (ANCOVA) model will be used to produce the LS mean, SE, and 95% CI of the LS means for each treatment arm. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated.

5.7.3 Exploratory Endpoints

Exploratory endpoints will be analyzed using the mITT/Evaluable populations. Only descriptive statistics will be provided, treatment comparisons will not be performed.

Exploratory endpoints included in this SAP are:

- Changes in concentration of AKST4290 and its major metabolites in plasma at various time points.
- Changes in bradykinesia, tremor, general activity, and sleep as measured by an FDA-cleared wearable sensor device.

Analysis of other exploratory endpoints may be performed, and are outside the scope of this SAP.

5.7.3.1 Pharmacokinetics

Plasma concentration measurements of AKST4290 and the major metabolites M227 and M373 will be collected to assess systemic exposure to the study agent. Pharmacokinetic samples will be collected immediately prior (-15 min) to study drug administration at Baseline and Weeks 2, 4, 8, 12, and 14, as well as 1 hour and 2 hours after administration at Baseline and Weeks 4 and 12. A final PK sample will be collected at Follow-Up (Week 14). Concentration of AKST4290 in plasma will be summarized descriptively at each visit timepoint.

5.7.3.2 Wearable Sensor Device

A wearable sensor device allowing for the continuous measurement of health-related symptoms, including the motor and nonmotor features of bradykinesia, tremor, general activity, and sleep, will be provided to applicable subjects at Screening, Week 4, and Week 8. Data will then be collected over a 6-day period and returned at the following visit (i.e. Baseline, Week 8, and End of Treatment). Bradykinesia, tremor, general activity, and sleep parameters will be summarized descriptively at each applicable timepoint.

To evaluate change from baseline, the same MMRM model used to evaluate the primary endpoint will be employed to produce LS means, SE, and 95% CIs of the LS means for each treatment arm at each scheduled

visit. The unadjusted mean change from baseline, and the difference in LS means between treatment arms and 95% CI of the difference will also be calculated.

5.8 Safety Analyses

Safety analysis includes study drug accountability, incidence and severity of treatment emergent adverse events, laboratory parameters, vital signs, 12-lead ECG, physical examination, pregnancy tests, and prior and concomitant medications. Unless otherwise specified, safety parameters will be summarized by treatment group using the safety population.

5.8.1 Study Drug Accountability

The number of subjects receiving daily doses of 800 mg of study medication will be summarized over the entire course of study treatment and between baseline and Week 4, Week 4 and Week 8, and Week 8 and End of Treatment (Week 12) by treatment group, for the safety population. The number of subjects with missed or additional doses will also be summarized.

5.8.2 Adverse Events

All TEAEs will be reported. A TEAE is defined as an AE that occurs on or after the date of the first study treatment.

A summary of all TEAEs by the number and percentages of subjects who experienced any of the following will be provided:

- Any TEAE, including severity (Mild, Moderate, and Severe), and relationship to study treatment (Unrelated, Possibly Related, Definitely Related), and AEs leading to discontinuation of study participation.
- Any serious TEAE, including severity (Mild, Moderate, and Severe), and relationship to study treatment (Unrelated, Possibly Related, Definitely Related), and AEs leading to discontinuation of study participation.
- Any event of Special Interest (AESI).
- Any fatal TEAEs, including relationship to study treatment (Unrelated, Possibly Related, Definitely Related).

5.8.2.1 Summary for All TEAEs

The number and percentage of subjects who experience TEAEs will be tabulated by SOC and PT using the MedDRA dictionary. Adverse events will be counted only once for a subject within each PT and SOC; thus, since a subject may have more than one PT within a SOC, percentages of PT may not sum to the percentages in the SOC.

For all TEAEs the following will be summarized:

- Overall summary of any TEAEs, presented by SOC/PT

- Overall summary of any TEAEs by worst severity, presented by SOC/PT
- Overall summary of any TEAEs by closest relationship to study treatment, presented by SOC/PT

A listing of all AEs will be provided. This listing will include a flag identifying if an AE is a TEAE.

5.8.2.2 Summary for Serious TEAEs

For all serious TEAEs the following will be summarized:

- Overall summary of any serious TEAEs by worst severity, presented by SOC/PT
- Overall summary of any serious TEAEs by closest relationship to study treatment, presented by SOC/PT

A listing of all serious AEs will be provided. This listing will include a flag identifying if an AE is a TEAE.

[redacted]

5.8.2.5 Deaths

Deaths, if any, will be provided in a listing only.

5.8.3 Laboratory Data

Biological samples (e.g. whole blood, serum/plasma, urine) will be collected at all in-person visits (i.e. Visits 1-7) for laboratory evaluations. Laboratory test will include hematology, chemistry, coagulation, infectious disease serology, qualitative urinalysis, and pregnancy testing in women of childbearing potential (WOCBP). Laboratory data summaries will consist of the following:

- Observed values at screening, baseline and Weeks 2, 4, 8, 12, and 14 along with change from baseline at each scheduled post-baseline visit.
- Shift tables by treatment group and by out of range flag at Weeks 2, 4, 8, 12, and 14, compared to baseline.
- Number and percentage of subjects with abnormal laboratory tests at all visits.

The tests will be presented in the original units collected. If the original value contains the "<" sign, AVAL in the laboratory analysis dataset (ADLB) dataset will be derived by extracting the numeric value from LBSTRESC and subtracting 0.1(N+1) to that numeric value, where N is the number of decimals in the numeric value. For example, AVAL will be set to 0.29 if LBSTRESC ="< 0.3". If the LBSTRESC value contains the ">" sign, AVAL in the ADLB dataset will be derived by extracting the numeric value from LBSTRESC and adding 0.1(N+1) to that numeric value, where N is the number of decimals in the numeric value. For example, if LBSTRESC =">1000", AVAL will be set to "1000.1". Subject listings will present all data collected along with toxicity grading.

Results for chemistry, hematology, urinalysis, coagulation, infectious serology, and pregnancy tests in WOCBP will be provided in separate listings. A listing of abnormal laboratory tests will also be provided.

5.8.4 Vital Signs

Vital signs, including blood pressure (mmHg), heart rate (beats/min), respiratory rate (breaths/min), and body temperature (C), will be collected at all in-person visits (i.e. Visits 1-7). Observed values will be tabulated at screening, baseline, and at each scheduled post-baseline time point. Change from baseline at each scheduled post-baseline time point will also be summarized by treatment group.

Vital signs at each scheduled time point will also be provided in subject data listings.

5.8.5 12-lead ECG

A 12-lead ECG will be performed to obtain Heart rate (beats/min), QT interval (msec), and QT interval corrected by the Fridericia formula (QTcF). ECGs will be collected at Screening and at Weeks 4 and 12. Observed values, along with the change from baseline at each scheduled post-baseline visit, will be summarized by treatment group. Additionally, a frequency table detailing the overall interpretation of ECG findings will be provided.

A listing of abnormal 12-lead ECG will also be generated.

5.8.6 Physical Examination

Full or targeted physical examinations will be conducted at every in-person study visit (i.e. Visits 1-7). The number and percentage of subjects with abnormal clinically significant and not clinically significant physical examination results will be summarized by the body systems at each scheduled visit.

5.8.7 Pregnancy tests

Pregnancy test results will be provided for women of childbearing potential in a listing only.

5.8.8 Elevated Liver Enzyme Investigations

Adverse events of abnormal liver enzyme values meeting the following criteria may require additional investigation:

- [redacted]
- [redacted]
- Drug-Induced Liver Injury

Elevated liver enzyme investigations may include: imaging, repeat lab testing, and review of symptoms, medical history, and exposure to risk factors (e.g. concomitant drug and alcohol use, etc.). All elevated liver enzyme investigations will be provided in listings. The number and percentage of subjects that met the criteria for the above liver enzyme categories will be summarized from the laboratory data.

5.8.9 Prior and Concomitant Medications

The number and percentage of subjects taking any prior and concomitant medications will be summarized by anatomical therapeutic chemical (ATC) level 3 terms and preferred medication names.

6 Document Version Control

Revision History:

REVISION	RELEASE DATE	AUTHOR	SUMMARY OF CHANGES
A	23Dec2020	M. Arthur	Original release
B	08Apr2021	M. Arthur/M. Barstow	-Modify specification of MMRM model to ANCOVA model, when only baseline and Week 12 data. -Per Sponsor request add ITT population for efficacy tables; add breakdown of each cognitive domain for MoCA, include mean scores for SE-ADL; add breakdown of each parameter for CISI-PD; include overall summaries for abnormal Chemistry lab summary table. -Other minor edits

Appendix A - Programming Specifications for Tables and Listings

The following specifications will be used in the production of tables and listings.

1. Page Setup
   Unless otherwise noted, tables and listings will use landscape orientation. Margins will be at least 3/4 of an inch on the left side of page, at least 3/4 of an inch at the top, and 3/8 of an inch on the other sides.
   Upper left: Sponsor name and protocol number
   - Center: CONFIDENTIAL; Database Download Date: ddmmmyy
   - Upper right: Page number shown as Page n of N. Page numbers should be sequential within a table or listing.

   The footer should include:
   - Left: the name of the SAS program used to generate the output
   - Center: run date/time and the words "by CTDS".
   - Right: output file name.

2. Footnotes
   Unless otherwise specified, footnotes should appear on all pages within the table.

3. Font
   Font will be 9-point Arial, or smaller if needed for space constraints. If possible, small tables should appear on one page. If tables continue on to multiple pages, there should be a page break after an assessment so that all the statistics for an assessment appear on the same page.

4. Tables
   Table titles should reflect the content of the table. Under the main title, in parentheses, the name of the analysis population being summarized should appear.

   Summary Statistics - Continuous Data

   Unless otherwise noted, the mean and median and confidence interval (CI) of a set of values should be printed out to one decimal place more than the original value. The standard deviation and standard error should be printed out to 2 decimal places more than the original value. The number of subjects on whom the parameter is assessed should appear. Minimum and maximum should be consistent with the original value.

   Summary Statistics - Categorical Data
   Numbers of subjects are reported as whole numbers. Null counts are represented as 0. Table percentages should be reported to one decimal unless otherwise noted. Null percentages should be reported as 0.0. For all categories, the total number of subjects with data will be presented as N.

5. Subjects Included in Listings
   In general, subject data listings should include all subjects who are randomized. The population flag (Safety/mITT/Evaluable/PP) should be included in all listings as a column to indicate which population(s) a subject belongs to. If a listing includes a subset of subjects who meet a certain condition (e.g., subjects with SAEs) then this should be clear from the title of the listing. If there is no record for a listing, then a statement, such as There is no serious adverse events in any of the treatment groups, will be presented.

References

Athauda D, Maclagan K, Skene SS, Bajwa-Joseph M, Letchford D, Chowdury K, et al. Exenatide once weeklyversus placebo in Parkinson's disease: a randomised, double-blind, placebo-controlled trial. Lancet.2017;390(10103):1664-1675.

Fahn S, Elton RL. Unified Parkinson's disease rating scale. In: Fahn S, Marsden CD, Calne D, Goldstein M, eds. Recent Developments in Parkinson's Disease. Florham Park: Macmillan Health Care Information. 1987:153-163.

Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder SocietysponsoredrevisionoftheUnitedParkinson'sDiseaseRatingScale(MDS-UPDRS):ScalePresentationandClinimetricTestingResults.MovDisorders.2008;23(15):2129-2170.

Hauser RA, Deckers F, Lehert P. Parkinson's disease home diary: further validation and implications for clinical trials. Mov Disorders. 2004;19:1409-1413.

Hauser RA, Friedlander J, Zesiewicz TA, Adler CH, Seeberger LC, O'Brien CF, et al. A home diary to assessfunctional status in patients with Parkinson's disease with motor fluctuations and dyskinesia. ClinNeuropharmacol. 2000;23:75-81.

Lang JT, Kassan TO, Devaney LL, Colon-Semenza C, Joseph ME. Test-retest reliability and minimal detectable change for the 10-meter walk test in older adults with Parkinson's disease. J GeriatrPhyTher. 2016;39:165-170.

Martinez-Martin P, Forjaz MJ, Cubo E, Frades B, de Pedro Cuesta J, ELEP Project Members. Global versusfactor-related impression of severity in Parkinson's disease: a new clinimetric index (CISI-PD). Mov Disord.2006;21:208-214.

Nasreddine ZS, Phillips NA, BedirianV,Charbonneau S, Whitehead V, Collin I, et al. The Montreal CognitiveAssessment, MoCA: a brief screening tool for mild cognitive impairment. Am Geriatr Soc. 2005;53:695-699.

Peto V, Jenkinson C, Fitzpatrick R. PDQ-30: a review of the development, validation and application of aParkinson's disease quality of life questionnaire and its associated measures. J Neurol. 1998;245[Suppl1]:S10S14.

Schwab RS, England AC. Parkinson syndrome due to various and specific causes. In: Vinkin PJ, Bruyn, eds.Handbook of Clinical Neurology. Amsterdam: Elsevier North Holland. 1968: 227-247.

Sheehan DV, Giddens JM, Sheehan IS. Sheehan-Suicidality Tracking Scale (S-STS) 2014. Innov Clin Neurosci.2014;11(9-10):93-140.