A Multicenter, Dose-Optimized, Double-Blind, Randomized, Placebo-Controlled, Parallel Efficacy Laboratory Classroom Study with KP415 in Children with Attention-Deficit/Hyperactivity Disorder

Protocol: KP415.E01

Version Number:2.0

18 JUN 2018

STATISTICAL ANALYSIS PLAN

KemPharm, Inc.
1180 Celebration Blvd, Suite 103
Celebration, FL 34747
Phone: 321-939-3416

SAP Approval Page

Study Title: A Multicenter, Dose-Optimized, Double-Blind, Randomized, Placebo-Controlled, Parallel Efficacy Laboratory Classroom Study with KP415 in Children with Attention-Deficit/Hyperactivity Disorder

Version: 2.0

Date of Issue: June 18, 2018

Study Sponsor: KemPharm, Inc.
1180 Celebration Blvd, Suite 103
Celebration, FL 34747

We, the undersigned, have read and approve of this SAP and agree on its content.

Sponsor Representative – Rene Braeckman, PhD

18 June, 2018
Date

Principal Investigator – Scott Kollins, PhD

18 JUN 2018
Date

Faculty Statistician – Sheng Luo, PhD

June 18, 2018
Date

TABLE OF CONTENTS

1.1 LIST OF ABBREVIATIONS

AE	Adverse Event
ADHD	Attention-deficit hyperactivity disorder
ADHD-RS-5	Attention-deficit hyperactivity disorder rating scale 5
AIC	Akaike's Information Criteria
ALT	Alanine transaminase
AST	Aspartate transaminase
BIC	Schwarz's Bayesian Criteria
BMI	Body Mass Index
CGI-S	Clinical Global Impressions–Severity
CGI-I	Clinical Global Impressions–Improvement
CI	Confidence Interval
C-SSRS	Columbia-Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EOS	End of Study
ET	Early Termination
FU	Follow-up
ITT	Intent-to-Treat
LLN	Lower Limit of Normal
LS	Least-square
MAR	Missing at Random
MCAR	Missing Completely at Random
MedDRA	Medical Dictionary of Regulatory Activities
MMRM	Mixed-Effect Model Repeated Measure model
MPH	Methylphenidate
PERMP	Permanent Product Measure of Performance
PERMP-A	Permanent Product Measure of Performance-Attempted
PERMP-C	Permanent Product Measure of Performance-Correct
PP	Per-Protocol
PT	Prothrombin Time
PTT	Partial Thromboplastin Time
Q1	25^{th} Percentile (1^{st} Quartile)
Q3	75^{th} Percentile (3^{rd} Quartile)
QTcF	Time between the start of the Q wave and the end of the T wave (QT interval) in the heart's electrical cycle, corrected for heart rate with Fridericias's formula
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard Error
SKAMP	Swanson, Kotkin, Agler, M-Flynn, and Pelham Rating Scale

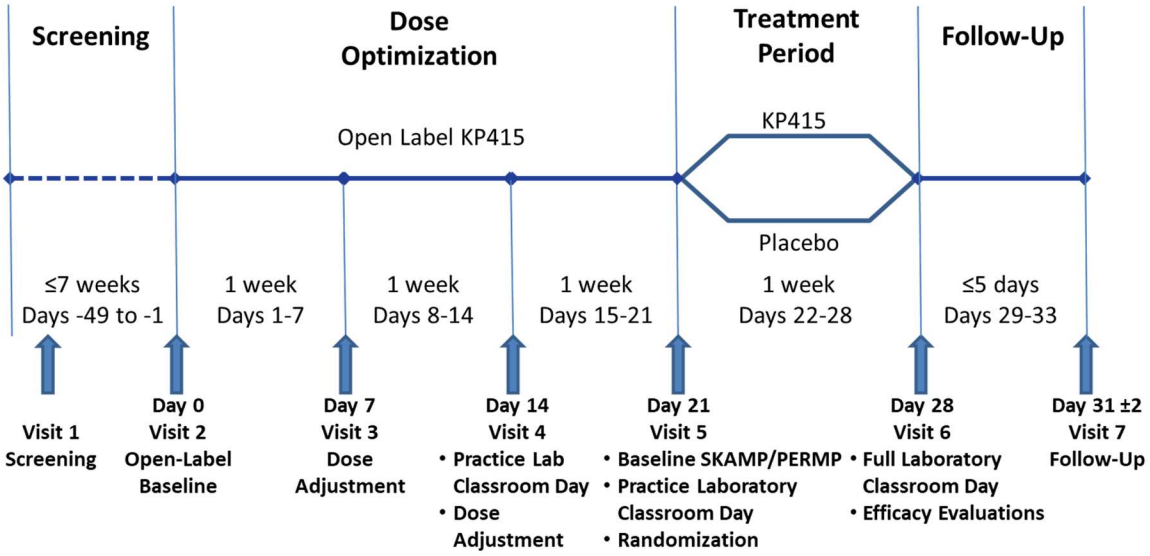
SKAMP-A	SKAMP-Attention
SKAMP-C	SKAMP-Combined
SKAMP-D	SKAMP-Deportment
TEAE	Treatment-Emergent Adverse Event
ULN	Upper Limit of Normal
WREMB-R	Weekly Rating of Evening and Morning Behavior - Revised

1.2 PROTOCOL SUMMARY

Name of Company	KemPharm, Inc.
Study Product	KP415 (a prodrug of d-methylphenidate, d-MPH)
Protocol Number	KP415.E01
Protocol Title	A Multicenter, Dose-Optimized, Double-Blind, Randomized, Placebo-Controlled, Parallel Efficacy Laboratory Classroom Study with KP415 in Children with Attention-Deficit/Hyperactivity Disorder
Study Objectives	Primary: To determine the efficacy of KP415 compared to placebo in treating children 6-12 years old with ADHD in a laboratory classroom setting. Secondary: 1. To determine the onset and duration of the clinical effect of KP415 in treating ADHD in children 6-12 years old in a laboratory classroom setting. 2. To determine the safety and tolerability of KP415 compared to placebo in treating children 6-12 years old with ADHD in a laboratory classroom setting.
Study Design	The study is a multicenter, dose-optimized, double-blind, randomized, placebo-controlled, parallel efficacy laboratory classroom study with KP415 in children with Attention-Deficit/Hyperactivity Disorder (ADHD). The study will consist of a Screening Period, an Open-Label Dose Optimization Phase, a Double-Blind Treatment Phase and a Follow-Up Visit, as follows: • Screening Period: Subjects will undergo a screening period up to 49 days prior to entering into the Open-Label Dose Optimization Phase. • Open-Label Dose Optimization Phase: During the Dose Optimization Phase, subjects will be titrated to doses of 20, 30 or 40 mg open-label KP415 based on tolerability and best individual dose-response in the opinion of the Investigator. • Double-Blind Treatment Phase: Eligible subjects will be randomized at Visit 5 (Baseline) to receive single daily doses of KP415 or Placebo for 7 days according to a randomization schedule. The dose of KP415 given in the Treatment Phase will be the same as the optimized dose of KP415 at the end of the Dose Optimization Phase. All subjects will receive their assigned treatment daily for 7 days. The dose will be the same at each day of the Treatment Period. Efficacy and safety assessments will be performed after the last dose of the Treatment Period (Visit 6). • Follow-Up Visit: 3 ±2 days after administration of the last dose of the Treatment Phase, subjects will enter a Follow-Up Visit to evaluate safety parameters.

Duration of Study Participation	Subjects will participate in the study as outpatients for up to 85 days.
Number of Subjects	An appropriate number of subjects will enter the Screening Period to enroll approximately 176 subjects in the Open-Label Dose Optimization Phase, and to randomize approximately 140 subjects in the Double-Blind Treatment Phase, with the intention to complete with approximately 126 subjects.
Number of Sites	5 sites in the United States of America
Study Population	Children 6-12 years old with Attention-Deficit/Hyperactivity Disorder who meet the inclusion/exclusion criteria.
Efficacy Endpoints	The primary efficacy evaluation is based on SKAMP and PERMP scores at pre-dose, and at 0.5. 1, 2, 4, 8, 10, 12, and 13 hours post-dose during the full laboratory classroom day at Visit 6. The baseline SKAMP score is measured at pre-dose at Visit 5. **Primary Efficacy Variable**: o Average of the change from baseline (measured at Visit 5) of the SKAMP-C scores collected post-dose across the laboratory classroom day at Visit 6. **Secondary Efficacy Variables**: o Change from baseline (measured at Visit 5) of the SKAMP-C scores measured at each time point on the laboratory classroom day at Visit 6. o The serial measures at different times post-dosing will be used to determine onset and duration of the clinical effect of KP415. o Change from baseline of the scores measured at each time point and the average of the scores collected across the laboratory classroom day at Visit 6, for the following endpoints: • SKAMP-D and SKAMP-A scores • PERMP scores • PERMP-A and PERMP-C scores • WREMB-R scores (total score, and morning and evening subscore) at Baseline (Visit 2), Visit 5 and Visit 6.

1.3 STUDY DESIGN SCHEMATIC



1.4 STUDY OBJECTIVES

Primary:

- To determine the efficacy of KP415 compared to placebo in treating children 6-12 years old with ADHD in a laboratory classroom setting.

Secondary:

- To determine the onset and duration of the clinical effect of KP415 in treating ADHD in children 6-12 years old in a laboratory classroom setting.
- To determine the safety and tolerability of KP415 compared to placebo in treating children 6-12 years old with ADHD in a laboratory classroom setting.

The primary interest of Study KP415.E01 is to assess efficacy of KP415 compared to placebo in treatment of children 6-12 years with ADHD in a laboratory classroom setting, which is an analog setting of regular school classrooms.

The study intends to employ the 13-Item SKAMP rating scale as the primary efficacy measurement for the assessment of behavioral impairment for ADHD children in the laboratory classroom setting. The SKAMP was initially developed in 1992 by Swanson, Kotkin, Agler, M-Flynn, and Pelham (Swanson, 1992), and has been used mainly in a laboratory classroom setting in which the application of a labor intensive coding system associated with the scale is possible (Greenhill et al., 1996; McBurnett et al., 1997; Swanson et al., 1998; Wigal et al., 1998; Swanson & Greenhill, 1999). The SKAMP has demonstrated sound psychometric properties, including moderate to high test-retest reliabilities, concurrent validity, and discriminant validity, especially in those situations where a non-treatment or placebo condition is involved. To date, the SKAMP has been widely and successfully utilized in several clinical trials of moderate sample sizes to assess treatment efficacy in ADHD children with respect to their behavioral impairments in the laboratory classroom setting with reference to placebo (McCracken et al., 2003; Wigal et al., 2005, Biederman et al., 2007).

In this study, nine assessment sessions (one pre-dose and eight post-dose) will occur over the laboratory classroom assessment day (i.e., Visit 6), and the SKAMP rating scale will be evaluated at different times, or in other words,

measured repeatedly. Therefore, the study intends to analyze the SKAMP rating scale using a Mixed-effects Model for Repeated Measures (MMRM) so that the issue of within-subject measurement correlation can be handled appropriately. The MMRM model is commonly used in clinical trials in which the design with repeated measures is employed. To address the primary objective of this study, the MMRM model will define the average of the change from baseline (measured before the first dose in the Treatment Phase [at pre-dose of Visit 5]) over the eight post-dose session scores (post-dose at Visit 6), as the primary efficacy comparison of KP415 vs. placebo treatments, whereas the comparison of KP415 vs. placebo at each session will be used to address the secondary efficacy objective of the study.

1.5 STATISTICAL HYPOTHESES AND ENDPOINTS

- Primary:
    - The primary hypothesis is that the test drug is superior (mean change in post-dose SKAMP-C scores from baseline are lower) when compared to placebo. The baseline measurement is collected pre-dose at Visit 5, after 2 days of no drug administration. SKAMP-C scores are then collected across the laboratory classroom day at Visit 6. We will evaluate efficacy by looking at the direction of effect. The null and alternative hypotheses are:

        - $H_0: \overline{\Delta SKAMP - C}_{control} - \overline{\Delta SKAMP - C}_{test} = 0$
          $H_A: \overline{\Delta SKAMP - C}_{control} - \overline{\Delta SKAMP - C}_{test} \neq 0$

        The hypothesis test is the main effect of the primary efficacy analysis model and will be based on the residual error of the model. The treatment-by-time interaction will be pre-specified in the primary efficacy model regardless of the statistical significance of the interaction term.

- Primary Efficacy Endpoint(s):
    - The primary efficacy endpoint will be the mean change from the pre-dose baseline measurement at Visit 5 and SKAMP-C ($\overline{\Delta SKAMP - C}$) scores collected post-dose across the laboratory classroom day at Visit 6.
        - The Swanson, Kotkin, Agler, M-Flynn, and Pelham Rating Scale (SKAMP) scale is a validated rating of subjective impairment of classroom behaviors. It is comprised of 13 items (grouped under the subcategories of attention, deportment, quality of work, and compliance) on which subjects are rated according to a 7-point scale (0 = normal to 6 = maximal impairment) by trained study personnel (Swanson 1999). The SKAMP-Combined (SKAMP-C) score is obtained by summing the rating values for the 13 items of the SKAMP scale.
        - The SKAMP scale is collected at Baseline (pre-dose at Visit 5 after 2 days of no drug administration) and at 9 time points during Visit 6: pre-dose, 0.5, 1, 2, 4, 8, 10, 12, and 13 hours post-dose.

- Secondary Efficacy Endpoint(s):
    - $\overline{\Delta SKAMP - C}$ scores from baseline at each time point on the laboratory classroom day at Visit 6.
    - The serial measures of SKAMP-C scores at different times post-dosing will be used to determine onset and duration of the clinical effect of KP415.
    - The mean change from baseline at each time point collected across the laboratory classroom day at Visit 6 for the following endpoints:
        - SKAMP-D and SKAMP-A scores

        - SKAMP-Deportment (SKAMP-D) score is a measure of behavior and is comprised of 4 items.
        - SKAMP-Attention (SKAMP-A) score is a measure of attention and is comprised of 4 items.
    - PERMP scores
        - The Permanent Product Measure of Performance (PERMP) is an individually calibrated five-page mathematics worksheet consisting of 400 problems. The PERMP score is calculated by dividing the number of correct problems (PERMP-C) by the number of problems attempted (PERMP-A).
        - The PERMP score is collected at Baseline (pre-dose at Visit 5 after 2 days of no drug administration) and at 9 time points during Visit 6: pre-dose, 0.5, 1, 2, 4, 8, 10, 12, and 13 hours post-dose.
    - PERMP-A (Number of problems attempted) and PERMP-C scores (Number of problems correct)
- WREMB-R (Overall), Morning, Evening
    - The 11-item Weekly Rating of Evening and Morning Behavior - Revised (WREMB-R) questionnaire is a parent-rated questionnaire that was developed to assess behaviors for their severity during the morning hours (3 items) and evening hours (8 items) (Carlson 2007). The possible score for each item ranges from 0 (no difficulty) to 3 (a lot of difficulty).
    - The WREMB-R is collected at Visit 2, Visit 5, and Visit 6. The assessment at Visit 2 is the baseline assessment (after washout of ADHD medications, if applicable); the assessments at Visits 5 and 6 are evaluations at the end of the Dose Optimization and Treatment Phase, respectively. WREMB-R scores will be compared separately for the Dose Optimization (Visit 5 versus baseline at Visit 2) and Treatment Phase (Visit 6 versus baseline at Visit 2).

*We expect that secondary endpoint results will follow the primary endpoint in direction.

- Efficacy Baseline: For the primary and secondary SKAMP and PERMP efficacy endpoints, Baseline is considered Visit 5. For all other endpoints, Baseline is considered Visit 2 (or closest time-point prior to first dose of open-label study medication).

- Safety Endpoint(s):
    - Treatment-Emergent Adverse Events (TEAEs)
        - Adverse Events with new onset during the study between the initiation of study drug and 5 days after the last dose of study drug will be considered TEAEs.
    - Laboratory tests
        - Total Hematology as well as differential and Coagulation: red blood cell count, white blood cell count with differential (neutrophils, lymphocytes, monocytes, eosinophils, basophils), hemoglobin, hematocrit and platelets, Prothrombin Time (PT) and Partial Thromboplastin Time (PTT)
        - Serum Chemistry: aspartate aminotransferase (AST), alanine aminotransferase (ALT), albumin, alkaline phosphatase, bicarbonate, total bilirubin, blood urea nitrogen, phosphorus (inorganic) calcium, chloride creatinine phosphokinase, creatinine, gamma

glutamyl transferase, glucose, lactate dehydrogenase, potassium, sodium, total protein, Screening only thyroid stimulating hormone (TSH), and uric acid.
        - Urinalysis results: specific gravity, presence of hemoglobin, albumin, glucose, pH, and leukocyte esterase
    - Vital signs (systolic and diastolic blood pressure, pulse rate, respiratory rate, and oral temperature)
    - Height, weight, and BMI (will be derived using height and weight)
    - Frequency of body system results. Results include: Normal; Abnormal, Not Clinically Significant; or Abnormal, Clinically Significant
    - ECG parameters
        - Electrocardiograms (ECGs) will be interpreted overall as: Normal or Abnormal
        - The following ECG parameters are collected:
            - Heart rate (bpm)
            - PR interval (ms)
            - RR interval (ms)
            - QRS complex (ms)
            - QT interval (ms)
            - QTc interval (ms) [Fridericia's formula - QTcF]
        - ECGs are obtained at the Screening visit and at Early Termination (ET) or Follow-Up (FU) visit.
    - C-SSRS
        - Suicidal ideation will be assessed by the Columbia-Suicide Severity Rating Scale (C-SSRS, Pediatric Version) (Posner et al. 2010). There are 5 items asked during the subject's lifetime and during the past 6 months.
        - The C-SSRS is collected at all study visits.

- Exploratory Endpoints:
    - ADHD-RS-5 (Overall), Hyperactivity/Impulsivity subscale, Inattentiveness subscale
        - The Attention-deficit hyperactivity disorder rating scale 5 (ADHD-RS-5) is an 18-item scale based on Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) (American Psychiatric Association 2013) criteria of ADHD that rates symptoms on a 4-point scale. Each item is scored using a combination of severity and frequency ratings from a range of 0 (reflecting no symptoms or a frequency of never or rarely) to 3 (reflecting severe symptoms or a frequency of very often), so that the total ADHD-RS-5 scores range from 0 to 54. The 18 items are then divided into two 9-item subscales: One for hyperactivity/impulsivity and the other for inattentiveness.
        - The ADHD-RS-5 is collected at Visits 2 (Baseline), 3, 4, and 5.
    - CGI-S
        - The Clinical Global Impressions–Severity (CGI-S) is a clinician-rated scale that evaluates the severity of psychopathology (ADHD symptoms in the study) on a scale from 1 (not at all ill) to 7 (among the most severely ill) (Busner and Targum 2007).
        - The CGI-S is collected at Visits 2 (Baseline), 3, 4, and 5.

- CGI-I
  - The Clinical Global Impressions–Improvement (CGI-I) is scored from 1 (very much improved) to 7 (very much worse).
  - The CGI-I is collected at Visits 2 (Baseline), 3, 4, and 5.
- Conners 3-P
  - The Conners 3-P (short form) is a 43-item parent/guardian/caregiver report that provides evaluation of 6 assessment subscales: Inattention (IN), Hyperactivity/Impulsivity (HY), Learning Problems (LP), Executive Functioning (EF), Aggression (AG) and Peer Relations (PR). In addition, there are 2 validity/response style subscales: Positive Impression (PI) and Negative Impression (NI). Each item is scored on a 4-point scale ranging from Not true at all (Never, Seldom) to Very much true (Very often, Very frequently). The additional questions and scores beyond item 43 on the forms will not be used.
  - For each item, sites will use a software package (www.mhs.com/MHS-Assessment?prodname=conners3) to calculate raw total scores. The score for each of the items contributes uniquely to one of the eight sub-scores, 6 assessment subscales and 2 validity/response style subscales.
  - Using the software, the raw total scores of the 6 assessment subscales will then be converted to T-scores specific for the age and gender of the child being evaluated. The DSM-5 scoring option will be used. T-scores for each of the 6 assessment subscales will be assessed. The 2 validity/response style subscales (raw total scores) will also be captured in the database.
  - The Conners 3-P is collected at Visits 2 (Baseline), 3, 4, 5, and 6.

1.6 SAMPLE SIZE DETERMINATION

Drug efficacy studies are sized to be able to detect at least the Minimum Clinically Important Difference (MCID). MCID has not been established for the SKAMP-C scores. However, Rai et al. (Rai et al. 2015) noted “0.5 standard deviations has been suggested to correspond to the MCID in a number of studies“. Biederman et al. (Biederman et al. 2007) also noted, in a similar study on lisdexamfetamine dimesylate, “Previous studies of amphetamine products using the SKAMP-D to assess efficacy in children with ADHD have disclosed an effect size of greater than .50.”

Assuming a mean difference between test and control for ΔSKAMP-C of 0.5 units and a standard deviation (SD) of 1.0, 126 subjects would need to complete the study, assuming 80% power when testing a significance level of $\alpha = 0.05$ (2-sided two-sample z-test).

Assuming an approximate 20% dropout during the Open-Label Dose Optimization Phase, 176 subjects are planned to be enrolled in the Open-Label Dose Optimization Phase. We assume that there will be an approximate 10% dropout rate during the Double-Blind Treatment Phase. Therefore, approximately 140 subjects will be randomized into the Double-Blind Treatment Phase with the intention to complete with approximately 126 subjects.

Subjects will be randomized (1:1) into two groups: KP415 (N~63) or Placebo (N~63). Randomization will be stratified by site.

1.7 DATA SOURCES

At each study site, data will be entered on the electronic Case Report Forms (eCRFs) stored on IBM Clinical Development (run by Merge Healthcare, an IBM company). Laboratory evaluations will be received from a central laboratory and will be integrated with the clinical database. A 12-lead ECG will be read by a centralized ECG laboratory (ERT) and results will be transferred to the DCRI and integrated with the clinical database. One master list of randomized trial assignments will be uploaded to IBM Clinical Development from DCRI Statistics. Prior to database lock, programmed computer edit checks will be run against the database to identify discrepancies and verify reasonableness of the data. Queries to resolve discrepancies will be generated and resolved by the sites. Periodically, DCRI Statistics will receive or download from DCRI Data Export Services the eCRF database as Clinical Data Interchange Standards Consortium Study Data Tabulation Model (CDISC SDTM version 3.1.3) datasets. CDISC Analysis Data Model (ADaM version 2.1) datasets will be created by DCRI Statistics for production of tables, figures, and listings. All planned reporting will be based off of CDISC datasets, but in the case of emergent safety data, some reporting may occur from the raw eCRF data. All programs written to create analysis datasets and perform analyses will be validated according to SOPs established by the DCRI Statistical Programming group.

1.8 DOCUMENTATION CONVENTION

The statistical analyses described in this SAP, as well as production of tables, listings, and figures will be performed using SAS®, version 9.4 or higher (SAS Institute, Cary, NC). Additional statistical software may be used as needed.

1.9 VERIFICATION OF RESULTS

All tables, listings, and graphs will be verified and reviewed before considered final. The verification process will ensure that the numbers are produced by a statistically valid method and that the execution of the computations is correct. Qualified statisticians or statistical programmers employed by the DCRI who have not been previously involved in the production of the original programming will perform the verification procedures. Methods of verification include independent programming, prior to issuance of the draft statistical report, of all analysis datasets/ADaM and comparison to data listings. Tables, listings, and graphs will be reviewed for accuracy, consistency with this analysis plan, consistency within tables/listings/graphs, and consistency with corresponding output. Once verification is complete, all documentation of the verification process will be filed in a statistical programming documentation notebook as required by the Statistical Standard Operations Procedures of the DCRI.

1.10 SUBJECT DISPOSITION

The disposition of subjects (number randomized, number who received any amount of the randomly assigned treatment, number completing study drug administration, number who withdrew consent or discontinued from study drug early, and number lost to follow-up, and number who completed the trial) will be summarized overall and by treatment arm. The number of subjects screened for inclusion and a breakdown of reasons for exclusion will be summarized. The timing and reasons for early discontinuation of study drug and/or withdrawal from the study will be summarized by treatment group. A listing of all patients discontinued from the study after enrollment, broken down by site and treatment group will be provided. The listing will include: reason for discontinuation, treatment group, duration of treatment, and whether or not the blind was broken. Also, for patients who discontinued from the study after enrollment, a listing of adverse events will be provided. Lastly, the number of subjects with major protocol violations will be summarized and listed.

1.11 POPULATIONS FOR ANALYSES

- **Intent-to-Treat (ITT) Population**: All randomized subjects who receive at least one dose of double-blind study medication and have at least one of the SKAMP-C assessments at Visit 6. Subjects will be grouped according to their randomized allocation (KP415 or placebo), regardless of whether the allocated therapy was administered or switched.
- **Per-Protocol (PP) Population**: ITT subjects who received the morning dose of double-blind study medication at the laboratory test session, who have all SKAMP-C assessments at Visit 6, who did not miss more than 2 days of therapy during the double-blind Treatment Phase, and did not use prohibited medications during the double-blind treatment period. Subjects included in the PP population will have received the study drug assigned to them by their randomization in a manner consistent with the intention of the protocol; subjects who prematurely discontinue study drug because of non-compliance or unwillingness to comply with the procedures required by the protocol will be excluded.
- **Overall Safety Population**: All subjects that entered the Open-Label Dose Optimization phase and received at least one dose of open-label study medication and had at least one post-dose safety assessment. This will include subjects who were randomized as well as those subjects who were never randomized. This population will be used for safety analyses.

  Baseline analyses, including demographics and other subject characteristics, will be performed using all populations. Further details are provided in section 1.12.17. All efficacy analyses will be performed using both the ITT and PP populations. All safety analyses will be performed using the Overall Safety population.

  In each analysis population, the distribution of dose levels will be reported as follows:
  - For the Dose Optimization Phase, the number of subjects in each week by dose level and the average* dose level between study visit intervals.
  - For the Treatment Phase, the number of subjects in each treatment group by dose level.

*Calculation of the average dose level: For the intervals between Visit 2 and Visit 3 and between Visit 3 and Visit 4, we will first get the number of days in the interval (the denominator). The formula will be as follows: date dose ended – date dose started +1. To calculate the sum of the doses during that interval (the numerator), we will first derive the number of days in the interval that study drug was taken. Using the number of days in the interval, we will subtract the number of days that the subject reported as having a missed dose. We will then take the number of days remaining and multiply that by the dose administered to get the sum. If there are any overdoses reported, we will add the additional dose to the numerator. We will then use the formula: sum of all doses in the interval/number of days in the interval = average dose. For the interval between Visit 4 and Visit 5, we will subtract the number of washout days from both the numerator and denominator to derive the average using the same formula.

1.12 STATISTICAL ANALYSES

1.12.1 GENERAL APPROACH

- Statistical significance: Statistical comparisons will be performed using two-sided significance tests. An alpha level of 0.05 will determine significance (unless otherwise noted).

- Descriptive statistics:
  - Continuous variables will be presented as n, mean, standard deviation, coefficient of variation (CV%), standard error (SE), median, Q1, Q3, and minimum and maximum. For comparisons of treatment

groups, if data is normally distributed we will use the t-test with unequal variances. If data is not normally distributed, we will use the non-parametric Wilcoxon/Mann-Whitney rank-sum test.
- Categorical variables will be presented as percentage (number). Group comparisons will use the conventional chi-square test or Fisher's Exact Test.
- Binary variables will be presented as percentage (number). Group comparisons will use the conventional chi-square test or Fisher's Exact Test.

- Handling of Missing Data:

  Missing data will be assessed in the primary and secondary efficacy analyses. Every effort will be made to minimize the occurrence of missing data. However, if data which are not efficacy related are missing, the general approach will be to use only data which are complete with regard to the variable(s) and time point(s) being analyzed for descriptive summaries, or the evaluation of safety endpoints.

  The Mixed-Effect Model Repeated Measure (MMRM) approach will be used in the primary and secondary efficacy analyses. Despite careful planning and study conduct, the occurrence of incomplete data cannot be completely eliminated. MMRM is based on the assumption of data that is missing at random (MAR). The MMRM uses all available data. If a score is missing, it has no effect on other scores from that same subject. Any missing data will be assumed to be MAR. As such, the MMRM approach provides valid inference and no additional steps need to be taken to handle the missing data.

- Handling of Missing Items or Scales:

  In the context of health outcomes in the format of rating scales, either the entire scale can be missing (unit non-response) or several items from the scale can be missing (item non-response). If <20% of items are missing from a scale, the items will be imputed using the mean of the remainder of the items. If >= 20% of items are missing from a scale, the entire scale will be considered missing and will be excluded from the analysis.

1.12.2 ANALYSIS OF THE PRIMARY EFFICACY ENDPOINT(S)

The primary efficacy endpoint treatment effect is the change from baseline in all post-dose SKAMP-C scores over the day during Visit 6. Though the SKAMP questionnaire uses the Likert scale, the SKAMP-C, which aggregates the scores from all the domains, it will be treated as a continuous measure.

To estimate the difference between test (KP415) and control (placebo), a repeated measures analysis using MMRM will be performed. The MMRM model will include post-dose time (session), treatment, the interaction of time and treatment, and site as fixed effects, and subject as random effect; and, the baseline SKAMP-C scores will be pre-specified in the model as a covariate. We will present the model-adjusted average change from baseline of all post-dose SKAMP-C scores for each treatment group and treatment group differences (KP415-placebo) with standard errors (SEs) and 95% confidence intervals (CIs). SAS procedure PROC MIXED will be utilized to conduct the primary efficacy analysis. If the model specified is not estimable, the compound symmetry type of the covariance matrix will be used. The SAS code for the MMRM model is as follows:

```
proc mixed data=adef method=reml;
Class trt02p usubjid atptn siteid;
```

```
Model chg= trt02p atptn trt02p*atptn base siteid/ ddfm=residual solution chisq;
repeated atptn /type= toep  group=trt02p local subject=usubjid rcorr;
parms / ols;
Lsmeans trt02p/diff cl pdiff alpha=0.05;
run;
```

trt02p=treatment group; usubjid=subject ID number; atptn= assessment time point; siteid=site; chg=change in SKAMP-C score from Baseline, base = SKAMP-C Baseline (pre-dose at Visit 5) score

Below is a sample table of results:

	KP415 (N=)	Placebo (N=)	Treatment Difference: KP415-Placebo
Pre-Dose (Baseline Visit 5) Mean (SD)	xx.x (xx.x)	xx.x(xx.x)	
Pre-Dose (Visit 6) Mean (SD)	xx.x (xx.x)	xx.x(xx.x)	
Average Post-Dose Change from Baseline			
LS Mean (SE)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)
95% Confidence Interval	xx.x – xx.x	xx.x – xx.x	xx.x – xx.x
p-value			x.xxx

To check if the assumptions of the MMRM model are met, residuals will be examined through histograms, normal plots, Shapiro-Wilk's test, and plots of the residuals versus fitted values. If there is strong evidence that the assumptions are not satisfied (defined as both the p-value<0.05 and W statistics <0.85), the non-parametric Wilcoxon Rank-Sum test will be utilized to perform between-treatment comparisons for the average score and at each post-dose time point.

1.12.3 ANALYSIS OF THE SECONDARY EFFICACY ENDPOINT(S)

The secondary endpoint analyses described in this section will be performed if the primary endpoint hypothesis test is significant. If the primary endpoint hypothesis test is not significant, only descriptive statistics will be generated at each time point during Visit 6 and for the average of the scores collected during Visit 6. The following will be reported: n, mean, standard deviation, coefficient of variation (CV%), standard error (SE), median, Q1, Q3, and minimum and maximum.

If the primary endpoint hypothesis test is significant, the following additional analyses will be performed:

- **The mean difference in SKAMP-C scores from baseline measured at each time point at Visit 6**. Similar to the primary endpoint analysis, MMRM will be used to estimate the least square means, SEs and their corresponding 95% confidence intervals (CIs) between treatment groups for each time point at Visit 6.

  Below is a sample table of results:

Time point	Change from Baseline (Visit 5 pre-dose) in SKAMP-C Score				
Visit 6	LS Mean (SE)		Difference in LS Mean (SE)	95 % Confidence Interval	Unadjusted p-value
	KP415 (N=)	Placebo (N=)	KP415-Placebo	KP415-Placebo	
Pre-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
0.5 hours post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
1 hour post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
2 hours post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
4 hours post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
8 hours post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
10 hours post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
12 hours post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
13 hours post-dose	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x – xx.x	x.xxx
Mean difference in change from baseline across all post-dose time-points			xx.x(xx.x)	xx.x – xx.x	x.xxx

- **Onset and duration of the clinical effect of KP415**. The onset of effect is defined as the first post-dose assessment time showing statistical significance ($p<0.05$) between KP415 and placebo as measured by the SKAMP-C Score. The duration of treatment effect is defined as the length of the time interval, such that statistical significance was reached at each time point of this interval. Onset and duration results will be presented in a table as well as graphically displayed.

- **All other secondary efficacy endpoints**. Scores and subscores of SKAMP and PERMP will be treated as continuous data. Individual time point estimates will be determined using the same model analysis methods and output described for the SKAMP-C endpoint.

- Raw means ± SE of changes from Baseline of secondary efficacy endpoints (scores and subscores of SKAMP and PERMP) will be plotted versus time by treatment. Baseline (Visit 5) and Visit 6 LS Means ± SE of change from Baseline of secondary efficacy endpoints (scores and subscores of SKAMP and PERMP) will be plotted versus time by treatment. The statistical significance of treatment differences at each time point will be indicated.

- WREMB-R scores will be analyzed as continuous measures. Scores will be examined separately for the Open-Label Dose Optimization Phase and the Double-Blind Treatment Phase. For the Open-Label Dose Optimization Phase, change in mean scores between Baseline (Visit 2) and Visit 5 will be analyzed using a paired t-test. Mean scores and 95% confidence intervals will be plotted by visit. For the Double-Blind Treatment Phase, changes in mean scores from baseline between treatment groups will be analyzed using a two-sample t-test. Mean scores and 95% confidence intervals will be plotted by visit and treatment. The statistical significance of treatment differences at Visit 6 between treatment groups will be indicated.

1.12.4 SAFETY ANALYSES

All analyses of safety will be conducted using the Overall Safety Population. The frequencies of adverse events (AEs), the results of laboratory assessments, physical examinations, vital signs, ECG results, and the frequency of suicidal ideation or behavior (assessed using the C-SSRS) will be summarized descriptively by treatment arm and overall. All AE-

related data will be summarized separately for the Dose Optimization Phase and Treatment Phase Frequencies of AEs in each phase will be expressed in percentage relative to the respective number of subjects in each phase (a denominator for the Dose Optimization Phase and another denominator for the Treatment Phase). Safety analyses performed during the Treatment Phase will be presented by treatment group (KP415 vs. Placebo) as well as overall.

AEs will be coded in accordance with the Medical Dictionary for Regulatory Activities (MedDRA). AEs will be graded with regard to severity according to criteria defined in the Common Terminology Criteria for Adverse Events v4.0 (CTCAE). Each AE will be counted once only for a given subject. A listing of AEs will be provided including: treatment arm, sex, age, preferred term (PT), study phase, onset date/time, time to onset, stop date/time, duration, action taken with study treatment, severity, outcome, relationship to study treatment, and whether it was a serious adverse event (SAE).

AEs with new onset during the study between the initiation of study drug and 5 days after the last dose of study drug will be considered treatment-emergent (TEAEs). This will include any AE with onset prior to initiation of study drug and increased severity after the treatment initiation. TEAEs will be summarized by system organ class and preferred term, and by treatment arm. The most common (occurring in >2% of subjects) TEAEs will be summarized by preferred term and by Dose Optimization and Treatment phases. In the Treatment phase, the most common TEAEs occurring more frequently in subjects in the KP415 treatment group than in subjects on placebo will be presented. Overall incidence rates (regardless of severity and relationship to study drug) and incidence rates for moderate or severe AEs will be summarized by treatment arm. A summary of SAEs and AEs leading to early discontinuation from the study will be presented by treatment arm. SAE rate will be calculated overall and by treatment arm. To compare the SAE rate by treatment arm, we will perform a Fisher's Exact test.

AEs will also be flagged for potential for abuse based on the following MedDRA PTs: Euphoria-related terms (Euphoric mood; Elevated mood; Feeling abnormal; Feeling drunk; Feeling of relaxation; Dizziness; Thinking abnormal; Hallucination; Inappropriate affect), Terms indicative of impaired attention, cognition, and mood (Somnolence; Mood disorders and disturbances), Dissociative/psychotic terms (Psychosis; Aggression; Confusion and disorientation), and Related terms not captured elsewhere (Drug tolerance; Habituation; Drug withdrawal syndrome; Substance-related disorders). Abuse-related AEs will be categorized by PT and tabulated by treatment group and dose. A listing of abuse-related AEs will be provided including: treatment group, time of onset, duration of event, dose of drug taken, severity, action taken, and outcome.

Post-treatment changes from baseline for all safety laboratory tests, vital signs, and ECGs will be summarized using descriptive statistics (n, mean, standard deviation, median, Q1, Q3, minimum and maximum). The eight statistics will be given at each time point for current reading and for change from baseline. For physical exam, vital signs (weight, height, BMI), ECG, and lab/urinalysis results, the last observation (ET/FU) will be compared to the baseline result. If any laboratory tests, vital signs, or ECGs are collected at an unscheduled visit post-baseline, results will be included in listings. Those subjects with significant laboratory abnormalities will be identified in data listings.

For the purposes of safety analyses, Baseline is considered the closest time point prior to first dose of open-label study medication. For the safety measurements of height, weight, BMI, chemistry and hematology panels, urinalysis, and ECGs, baseline is measured at the Screening visit. For all other tests, baseline is measured at Visit 2. If tests are scheduled to be measured at both Screening and Visit 2 and Visit 2 tests are missing, screening measurements will be used as the baseline.

Additional safety analyses include descriptive statistical summaries of abnormal laboratory values, frequency distributions and shift tables summarizing abnormal laboratory values, vital signs, concomitant medications, prior medications, medical history, and pregnancy test results.

1.12.5 LABORATORY DATA

Descriptive statistics for baseline and end of study (EOS) will be generated for abnormal laboratory data (hematology and chemistry values). Additionally, shift tables will be presented. The shift tables tabulate the number of lab values determined to be “Normal” and “Abnormal” at baseline and post-baseline time points. The classifications on shift tables will categorize lab values as “Normal” when they are in acceptable range, using a combination of clinician judgment specific to each lab test, and the reference ranges. Laboratory values outside the normal range will be graded on the Labs Listing as 1x, 2x, 3x, 5x, or 10x the upper limit (or .5x, 1x the lower limit) of normal value based on appropriate increases or decreases. Laboratory data will also be summarized graphically to show the magnitude and changes in individual subject values over time relative to normal ranges. Details of any abnormalities will be listed.

1.12.6 VITAL SIGNS

Descriptive statistics for observed values and changes from baseline will be tabulated by treatment group at each visit. Vital sign data plots will be produced to graphically display data summaries.

1.12.7 PHYSICAL EXAM

Results (Normal; Abnormal, Not Clinically Significant; or Abnormal, Clinically Significant) for each body system (general appearance, skin, lymph nodes, head and neck, lungs, cardiovascular system, abdomen, musculoskeletal/extremities, mental status, neurological system, and thyroid) will be presented by baseline and EOS. Results will be tabulated by treatment group and overall.

1.12.8 12-LEAD ECG

Overall interpretation (Normal, Abnormal NCS, and Abnormal CS) will be summarized by treatment group and overall (where applicable) at the Screening and EOS Visit (ET or FU) along with corresponding shifts between Screening and EOS assessments. In addition, descriptive statistics for observed values and changes from Baseline in ECG measurements will be tabulated by treatment group at the EOS Visit. If an ECG measurement was repeated, only the repeated measurement will be presented. Details of any abnormalities will be listed.

1.12.9 PRIOR AND COMCOMITANT MEDICATIONS

Prior and concomitant medication use will be tabulated by generic drug name.

1.12.10 PREGNANCY TESTS

Pregnancy test results at baseline, Visit 5, and EOS will be listed for female subjects of child-bearing potential.

1.12.11 URINE ALCOHOL/DRUGS OF ABUSE

Urine alcohol/drugs of abuse screen results at Screening will be listed.

1.12.12 URINE MPH SCREEN

Urine MPH screen results (positive or negative) at Screening, Visit 2 and Visit 5 will be listed.

1.12.13 DRUG ACCOUNTABILITY AND COMPLIANCE

Compliance will be tabulated by visit (weekly from Visit 3 through Visit 6), study phase (Dose Optimization and Treatment Phase), and overall. Percent compliance by visit will be calculated by dividing the number of missed doses reported at that visit by the number of days between that visit and the previous visit. For the interval between Visit 4 and Visit 5, 2 days will be subtracted for the wash-out period.

Number of missed doses and reason for missed dose will be listed by subject at Visits 3, 4, 5, and 6. Drug accountability for abuse potential will be tabulated by subject and will include any instances of overdose (more capsules taken than indicated), lost/stolen medication, and unused drug not returned by the end of the trial.

1.12.14 C-SSRS

The number of instances where the response is “yes” to items 1 and 2 will be tabulated as well as the number of subjects excluded due to suicidality.

1.12.15 ABUSE POTENTIAL SAFETY ANALYSES

In accordance with the 2017 FDA Guidance for Industry, Assessment of Abuse Potential of Drugs, the following analyses will be performed in the Overall Safety Population:

- Systematic categorization and tabulation of AEs that are abuse potential-related, using the MedDRA Preferred Terms, including euphoria-related terms, terms of altered attention, cognition and mood, and dissociative/psychotic terms (FDA Guidance 2017, Sellers and Romach 2017). An analysis by dose, age, and gender; and by case, in order to understand the incident that led to the AEs, establish the time at which AEs appear following drug administration, the duration of the AEs, and which AEs overlap temporally.

- Analysis of subject’s study drug accountability assessments that may provide information about the incidence of signals suggestive of abuse, such as substance use disorders, overdose, drug diversion or drug loss.

1.12.16 EXPLORATORY ANALYSES

- Changes in ADHD severity will be compared from week to week in the ITT and PP population based on ADHD-RS-5, CGI-S, CGI-I, and Conners 3-P.
    - For the ADHD-RS-5, CGI-S, and CGI-I endpoints, we will present the projection trajectories as a function of time (i.e. by study visit), without treatment group comparison. Change in mean scores between Baseline (Visit 2) and subsequent study visits will be analyzed using a paired t-test. Means and 95% confidence intervals of endpoints will be plotted by visit.

- To assess changes in ADHD severity between treatment arms, the Conners 3-P T-scores will be analyzed as continuous measures. The treatment effects will be examined at all assessment time points. Scores will be examined separately for the Open-Label Dose Optimization Phase and the Double-Blind Treatment Phase. For the Open-Label Dose Optimization Phase, change in mean scores between Baseline (Visit 2) and subsequent study visits will be analyzed using a paired t-test. Mean scores and 95% confidence intervals will be plotted by visit. For the Double-Blind Treatment Phase, changes in mean scores from baseline between treatment groups will be analyzed using the previously described repeated measures model. LS Means ± SE will be plotted versus visit by treatment. The statistical significance of treatment differences at Visit 6 will be indicated.
- Note: For the Conners 3-P, the statistical analysis will be performed with the 6 assessment subscale T scores. For the 2 total raw validity/response style subscale scores, descriptive statistics only will be calculated, by treatment group.

- No imputation or adjustments for missing data will be performed for the exploratory endpoints.

1.12.17 BASELINE DESCRIPTIVE STATISTICS

Key baseline clinical characteristics will be summarized by treatment group within all study populations (ITT, PP, and Overall Safety). Characteristics include: age, gender, ADHD subtype, ethnicity, race, weight, height, ADHD rating scale, and CGIS. Baseline is considered the closest time-point prior to first dose of open-label study medication.

Counts and percentages will be presented for categorical variables. Continuous variables will be presented as medians (Q1, Q3). For continuous variables, data will be assessed as to whether they are normally distributed using the SW test. If there is no evidence of lack of normality (SW $p>0.05$), then the two treatment arms will be compared statistically using the t-test with unequal variances. If the SW indicates that there is evidence that the variable is not normally distributed (SW $p<0.05$), then the non-parametric Wilcoxon rank sum test will be used to compare treatment groups. To compare categorical variables between treatment groups, chi-square or Fisher's exact tests will be used as appropriate.

1.12.18 PLANNED INTERIM ANALYSES

There are no planned interim analyses.

1.12.19 SUB-GROUP ANALYSES

The following subgroups have been pre-specified:

- Gender
- Site
- Dose
- Age (ages 6-9 and ages 10-12)

The estimated treatment effect of KP415 within each of the subgroup sets listed above will be examined. The subgroup analyses will be conducted using the same analysis model outlined for the primary efficacy endpoint. The subgroup analyses will performed for the ITT population

1.13 REFERENCES

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington, VA, American Psychiatric Association, 2013.

Biederman J, Boellner SW, Childress A, Lopez FA, Krishnan S, Zhang Y. Lisdexamfetamine dimesylate and mixed amphetamine salts extended-release in children with ADHD: a double-blind, placebo-controlled, crossover analog classroom study. Biol Psychiatry. 2007; 62(9): 970-6.

Busner J, Targum SD. The Clinical Global Impressions Scale: Applying a research tool in clinical practice. Psychiatry 4:28–37, 2007.

Carlson GA, Dunn D, Kelsey D, Ruff D, Ball S, Ahrbecker L, Allen AJ. A pilot study for augmenting atomoxetine with methylphenidate: safety of concomitant therapy in children with attention-deficit/hyperactivity disorder (2007). Child Adolesc Psychiatry Ment Health. 1(1):10.

Food and Drug Administration, Assessment of Abuse Potential of Drugs. (2017, January). Guidance for Industry, U.S. Department of Health and Human Services, Center for Drug Evaluation and Research, Washington, DC.

Rai SK, Yazdany J, Fortin PR and Aviña-Zubieta JA. Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. Arthritis Research & Therapy. 2015; 17(1): 143.

Posner K, Brent D, Lucas C, Gould M, Stanley B, Brown G, Fisher P, Zelazny J, Burke A, Oquendo M, Mann J. Columbia-Suicide Severity Rating Scale (C-SSRS) Pediatric Baseline Version. New York, NY: The Research Foundation for Mental Hygiene, Inc., 2010.

Sellers EM, Romach MK. Categorization of Abuse Potential-Related Adverse Events. Clin Pharmacol Drug Dev. 2017 Oct 11. doi: 10.1002/cpdd.394. [Epub ahead of print].

Swanson JM, Agler D, Fineberg E, Wigal S, Flynn S, Fineberg K, Quintana Y and Talebi H, The University of California, Irvine, Laboratory School Protocol for Pharmacokinetic and Pharmacodynamic Studies. in: L Greenhill, B Osman (Eds.) Ritalin. 2nd ed. Mary Ann Liebert, Larchmont, NY; 1999:405–430.

1.14 APPENDICES

1.14.1 APPENDIX 1. TABLE OF CONTENTS FOR STATISTICAL TABLES, FIGURES, AND LISTINGS

Tables

Physical Exam Body System Results by Treatment Group

Prior Medications by Generic Name by Treatment Group

Concomitant Medications by Generic Name by Treatment Group

Summary of C-SSRS by Treatment Group

Descriptive Summary of SKAMP-C Endpoint

Descriptive Summary of SKAMP-D and SKAMP-A Endpoints

Descriptive Summary of PERMP Endpoint

Descriptive Summary of PERMP-A and PERMP-C Endpoints

Descriptive Summary of WREMB-R Endpoints

Descriptive Summary of ADHD-RS-5 Endpoints

Descriptive Summary of CGI-I and CGI-S Endpoints

Descriptive Summary of Conners 3-P Endpoints

MMRM Analysis of SKAMP-C Endpoint – Primary Analysis

Mean SKAMP-C Scores by Time Point

Analysis of the Onset and Duration of the Clinical Effect of KP415

Mean SKAMP-D and SKAMP-A Endpoints by Time Point

Mean PERMP Endpoint by Time Point

Mean PERMP-A and PERMP-C Endpoints by Time Point

Mean WREMB-R Scores – Open-Label Dose Optimization Phase

Mean WREMB-R Scores – Treatment Phase

Mean ADHD-RS-5 Scores – Open-Label Dose Optimization Phase

Mean ADHD-RS-5 Scores – Treatment Phase

Mean CGI Scores – Open-Label Dose Optimization Phase

Mean CGI Scores – Treatment Phase

Mean Conners 3-P T Scores – Open-Label Dose Optimization Phase

Mean Conners 3-P T Scores – Treatment Phase

Listings

Subject Disposition, Discontinued Subjects – Overall Safety Population

All Adverse Events – Enrolled Subjects Population

Adverse Events – Overall Safety Population

Adverse Events, Discontinued Subjects – Enrolled Subjects Population

Treatment Emergent Adverse Events – Overall Safety Population

Abuse-Related Adverse Events – Overall Safety Population

Study Drug Compliance for Abuse Potential –Overall Safety Population

Subjects Excluded from the Efficacy Analysis –Overall Safety Population

Protocol Violations – Overall Safety Population

Pregnancy Test – Enrolled Subjects Population

Urine Alcohol/Drugs of Abuse Screen – Screened Subjects Population

Urine MPH Screen – Enrolled Subjects Population

Significant Laboratory Abnormalities –Overall Safety Population

Significant ECG Abnormalities – Overall Safety Population

Figures

Bar Plot of Dose Level Changes during Dose Optimization Phase

Worst Post-Baseline vs. Baseline Laboratory (Scaled by Normal Values): Hematology Tests

Worst Post-Baseline vs. Baseline Laboratory (Scaled by Normal Values): Hematology Tests by Treatment Group

Worst Post-Baseline vs. Baseline Laboratory (Scaled by Normal Values): Chemistry Tests

Worst Post-Baseline vs. Baseline Laboratory (Scaled by Normal Values): Chemistry Tests by Treatment Group

Worst Post-Baseline vs. Baseline Laboratory (Scaled by Normal Values): Urinalysis Tests

Worst Post-Baseline vs. Baseline Laboratory (Scaled by Normal Values): Urinalysis Tests by Treatment Group

Worst Post-Baseline vs. Baseline ECG Results

Worst Post-Baseline vs. Baseline ECG Results by Treatment Group

Summary Plot of Vital Signs

Summary Plot of Vital Signs by Treatment Group

WREMB-R Scores Over Time

ADHD-RS-5 Scores Over Time

CGI Scores Over Time

Conners 3-P Scores- Open-Label Dose Optimization Phase Over Time

Treatment Difference Over Time in SKAMP-C

Treatment Difference Over Time in SKAMP-D

Treatment Difference Over Time in PERMP

Treatment Difference Over Time in PERMP-A

Treatment Difference Over Time in PERMP-C

Treatment Difference Over Time in WREMB-R Scores

Treatment Difference Over Time in Conners 3-P Scores Treatment Phase