Statistical Analysis Plan

Sponsor:	Noven Pharmaceuticals, Inc.
Protocol No:	N25-006
[redacted]	
Version Date:	May 08, 2013
Version Number:	Amendment 4

APPROVALS

Sponsor Name:	Noven Pharmaceuticals, Inc.
Name of Sponsor Representative/Title:	[redacted]
Signature of Sponsor Representative /Date:	[redacted]
Name of Project Manager/Title:	[redacted]
Signature of Project Manager/Date:	[redacted]
Name of Biostatistician /Title:	[redacted]
Signature of Biostatistician /Date:	[redacted]

TABLE OF CONTENTS

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Noven Pharmaceuticals, Inc. Protocol: N25-006.

The format and content of this SAP are structured to provide sufficient detail to meet the requirements specified by the International Conference on Harmonization (ICH) E9: Guidance on Statistical Principles in Clinical Trials.

This SAP should be read in conjunction with the study protocol and case report form (CRF). This version of the plan has been developed using the protocol amendment 4 dated 08MAY2013 and version 2.0 of the CRF dated 05MAR2013. Any further changes to the protocol or CRF may necessitate updates to the SAP.

The SAP is to be developed in two stages. The purpose is to "finalize" a SAP so that programming can be started early in the process. Versions of the SAP up to initial sponsor approval will be known as SAP1. Changes following approval of SAP1 will be tracked in the SAP Change Log and a final version of the SAP, known as SAP2, will be issued for sponsor approval prior to database lock.

1.1 CHANGES FROM PROTOCOL

There are no changes from protocol amendment 4, dated 08 May 2013.

2. STUDY OBJECTIVES

The primary objective is to assess efficacy of d-Amphetamine Transdermal Drug Delivery System (d-ATS) compared to placebo, as measured by the Swanson, Kotkin, Agler, M-Flynn and Pelham Scale (SKAMP) total score. An additional primary objective is to assess the safety and tolerability of d-ATS.

The secondary objectives are:

- To assess onset of efficacy of d-ATS compared to placebo as measured by the SKAMP total score
- To assess duration of efficacy of d-ATS compared to placebo as measured by the SKAMP total score
- To assess additional efficacy assessments including:
    - Permanent Product Measure of Performance (PERMP)
    - Attention-Deficit/Hyperactivity Disorder Rating Scale, Version IV (ADHD-RS-IV)
    - Conners' Parent Rating Scale-Revised Short Form (CPRS-R:S)
    - Clinical Global Impression (CGI) scale.
- To assess the skin irritation, discomfort, adhesion and adhesive residue of d-ATS

3. STUDY DESIGN

This is a randomized, double-blind, cross-over, placebo controlled study to evaluate the safety and efficacy of d-ATS in children and adolescents 6 to 17 years of age with Attention Deficit Hyperactivity Disorder (ADHD).

The study will consist of a four-week screening period, a 72-hour wash-out period (if applicable), a five-week open-label, step-wise dose optimization period and two-week double blind randomized crossover treatment period with weekly classroom assessments and a safety follow-up by telephone 7 - 10 days after last dose of study drug. Approximately 90 subjects will be randomized at 3 sites.

Screening Period (Visit -1): Screening procedures may be performed within four weeks prior to the Dose Optimization Period of the study.

Washout Period: Eligible subjects will undergo a 72 hour wash-out period if applicable.

Baseline Visit (Day 0, Visit 0): Eligible subjects will return to the clinic on Day 0 for baseline assessments. Subjects who continue to meet the eligibility criteria will be entered into the open-label Dose Optimization period.

Dose-Optimization Period [Day 1 to Day 35; Visit 1 (Day 7), Visit 2 (Day 14), Visit 3 (Day 21), Visit 4 (Day 28) and Visit 5 (Day 35)]: Following Screening Period and Washout Period (if applicable), eligible subjects will enter the open-label Dose-Optimization Period. Optimal dose will be defined as the dose that will produce a reduction in ADHD RS-IV score $\geq$ 30% and CGI-Improvement (CGI-I) score of 1 or 2 and has tolerable side effects. Tolerability will be determined by the investigator. The Investigator can increase the current dose to provide additional symptom control. One dose reduction will be permitted if the subject experiences unacceptable tolerability of the current dose.

Subjects will be discontinued if they are unable to tolerate d-ATS or cannot reach their optimal dose by Week 5. If the subject has reached their optimal dose, then at the discretion of a medically qualified investigator, one dose reduction per participant may occur during the dose optimization period for safety reasons. However if the current dose is intolerable and the subject has never reached their optimal dose, the subject would be discontinued from the study. A dose increase may occur at visit 4 if no prior dose has met criteria for optimal response and the current dose is well tolerated as deemed by a medically qualified investigator. The dose used during the week 5 visit will be the dose that the subject is randomized to during the Double-Blind Treatment Period. During Visit 5, subjects will attend a half-day practice laboratory school with analog classroom sessions to become familiar with classroom schedules and procedures. Three SKAMP assessments will be performed, and 3 practice PERMP tests will be given during the practice session.

Double-Blind Treatment Period [Day 36 to Day 49; Visit 6 (Day 42) and Visit 7 (Day 49)]: Subjects who have reached an optimal dose will enter the two-week double-blind treatment period. Subjects will be randomized to receive daily d-ATS treatment (at the optimized dose) for one week followed by daily placebo patched (identical in appearance to d-ATS) for one week, or vice versa.

During Week 6 subjects will receive study patches in accordance with the randomization schedule and during Week 7 subjects will cross-over to receive the other study treatment per the randomization schedule.

Follow-Up: A safety follow-up will be conducted by telephone 7-10 days (Visit 8, Day 56) after last dose of study drug.

The duration of the study will be approximately 56 days for each individual subjects (not including screening and washout).

3.1 SAMPLE SIZE CONSIDERATIONS

Assuming a standard deviation (SD) of [redacted] and based on an average difference in total SKAMP scores between Placebo and d-ATS of [redacted], 74 subjects would need to complete the study to detect a true difference of [redacted] in mean total SKAMP scores between placebo and d-ATS at 80% power when testing at a significance level of $\alpha = 0.05$ (2-sided). Assuming a dropout rate of 20%, 90 subjects will be randomized and it is expected that 74 subjects will complete the study. A subject will be considered to have completed the study if the subject received both treatments during the double-blind Treatment Period and have completed both 12-hour classroom sessions during the double-blind Treatment Period.

3.2 RANDOMIZATION

Randomization will occur prior to the start of the double-blind study period. Subjects will be randomized to one of the two treatment sequences (d-ATS then Placebo or Placebo then d-ATS) in a 1:1 ratio. Each eligible subject will be randomized to group assignment by a centralized Interactive Voice Response System (IVRS) /Interactive Web Response System (IWRS). Randomization will be applied centrally, without any stratification factors, across all sites to ensure even distribution of subjects across each treatment sequence.

4. STUDY VARIABLES AND COVARIATES

4.1 PRIMARY VARIABLE

The primary efficacy measure is the mean SKAMP total score. The SKAMP scale is a validated rating scale that assesses manifestations of ADHD in a classroom setting through several subscales, including deportment (behavior) and attention. Multiple SKAMP assessments will be completed at the end of individual classroom sessions across the day by observers who rated each subject on 13 items, using a 7-point impairment scale (0=normal, 6=maximal impairment).

4.2 SECONDARY VARIABLES

4.2.1 SKAMP

The SKAMP total score at each assessment time point will also be used to assess the secondary measures of onset and duration of efficacy.

4.2.2 PERMP

The PERMP, a 5-page math test consisting of 80 problems per page (total of 400 problems), will be used in this study to evaluate effortful performance in the classroom as a measure of efficacy. Subjects will be instructed to work at their seats and to complete as many problems as possible in 10 minutes. The appropriate level of difficulty for each student will be determined based on results of a math pretest administered at screening. Performance will be evaluated using two scores: PERMP-A (number of problems attempted) and PERMP-C (number of problems correct). In order to prevent the subjects from taking the same test more than once during the study, subjects will receive randomized problems in a different version of the test at each assessment.

4.2.3 ADHD-RS-IV

The ADHD-RS-IV is a clinician-rated scale that reflects current symptoms of ADHD based on Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, Text Revision (DSM-IV-TR) criteria; it is a global assessment that measures the severity of symptoms from visit to visit, but is not utilized to assess symptoms of ADHD over the course of the day. The ADHD-RS-IV consists of 18 items that are grouped into 2 subscales (hyperactivity/impulsivity and inattention). Each item is scored on a 4-point scale from 0 (no symptoms) to 3 (severe symptoms), yielding a total score of 0 to 54.

ADHD-RS-IV Hyperactivity-Impulsivity Subscale includes the following question scores:

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

ADHD-RS-IV Inattention Subscale includes the scores for the following questions:

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted].

4.2.4 CPRS-R:S

The CPRS-R:S is a standard instrument for the assessment of ADHD in children and adolescents. It evaluates problem behaviors as reported by the parent or alternative caregivers. The CPRS-R:S contains 27 items and covers a subset of the subscales and items on the long parent form. Scales include: oppositional, cognitive problems/inattention, hyperactivity, and ADHD index.

4.2.5 CGI

The CGI provides a global evaluation of baseline severity and improvement over time, and, like the ADHD-RS-IV scale, measures global impressions of severity from visit to visit but not over the course of the day. At baseline, the investigator will use the CGI-Severity (CGI-S) to rate severity on a scale that ranges from 1 (normal, not at all ill) to 7 (among the most extremely ill subjects) plus a not assessed option. At each visit thereafter, the clinician will use the CGI-I to rate improvement relative to baseline on a scale ranging from 1 (very much improved) to 7 (very much worse) plus a not assessed option. For analysis, CGI-I and CGI-S scores will be dichotomized.

For CGI-S:
The **non-severe** illness is defined as subjects achieving a score of
1: Normal, not at all ill or
2: Borderline mentally ill or
3: Mildly ill.
The **severe** illness is defined as subjects achieving a score of
4: Moderately ill or
5: Markedly ill or
6: Severely ill
7: Among the most extremely ill patients.
Subjects achieving a score of 0: Not assessed will be considered as missing data. The number and percentage of subjects with severe illness and non-severe illness at each time point will be summarized.

For CGI-I
The **responders** are defined as subjects achieving a score of
1: Very much improved or
2: Much improved or
3: Minimally improved on the clinician-rated CGI global improvement item.
The **non-responders** are defined as subjects achieving a score of
4: No change or
5: Minimally worse or

6: Much worse or
7: Very much worse.

Subjects achieving a score of 0: Not assessed will be considered as missing data. The number and percentage of responder and non-responders at each time point will be summarized.

The actual values of each question of CGI-S, and CGI-I, severity (severe or non-severe), and responder (yes/no) will be listed.

4.2.6 Safety

Safety will be assessed by evaluating:

- Adverse events (AEs)
- Suicidality [assessed using Columbia Suicide Severity Rating Scale (C-SSRS)]
- Concomitant medications
- Clinical laboratory tests
- Vital signs
- Physical examinations
- Electrocardiograms (ECGs)

4.3 PREDETERMINED COVARIATES AND PROGNOSTIC FACTORS

Demographic and baseline characteristics known to be related to treatment response may be considered for inclusion as covariates in exploratory statistical analysis models.

The effect of the following factors on the primary efficacy analysis (FAS population) will be considered:

- Investigator site
- Optimized dose
- Gender
- ADHD type (inattentive, hyperactive/impulsive, combined, not otherwise specified)
- Baseline ADHD severity (defined as the pre-dose ADHD-RS-IV rating scale at baseline, Visit 0)
- Age group

To ensure summaries from the subgroups are informative, these will only be performed where subgroups include 5 or more subjects in each.Where subgroups have less than 5 subjects the small group will be combined with the next group, if there is a consecutive order to the groupings or with the next smallest group if there is no order.

5. DEFINITIONS

SKAMP Total Score

The SKAMP total score will be the sum of the 13 individual item scores at each timepoint. Missing data will not be handled at an item level, if any of the 13 items are missing for a timepoint then the SKAMP total score will be missing for that timepoint.

Mean SKAMP Total Score

The mean SKAMP total score will be the average of all SKAMP total scores collected over the course of a laboratory assessment day, with the exception of the score from the session prior to dosing.

Onset of efficacy

The onset of efficacy, based on the SKAMP total scores, will be defined as the time of the first assessment time showing statistical significance between d-ATS and placebo. If no significant difference is found at any time point for the onset of efficacy, no onset of efficacy will be deemed to have occurred, and the onset of efficacy will be defined as 'none'.

Duration of efficacy

The duration of efficacy, based on the SKAMP total scores, will be defined as the difference between the time point that indicates that the effect has ended and the onset of efficacy, as defined above. The time point that indicates that the effect has ended will be defined as the first time point at which there is a non-significant difference between the two treatment groups after a time point in which there is a significant difference between the two treatment groups. If there is no time point at which there is a non-significant difference between the two treatment groups after a time point in which there is a significant difference between the two treatment groups, then the end of effect will be defined as 12 hours post-dose.

Onset of effect (50% reduction in SKAMP total score)

Onset of effect of 50% reduction in SKAMP total score is defined for each subject as the first time point at which there is a 50% reduction in SKAMP total score from the pre-dose observation on the same classroom day.

Duration of effect (50% reduction in SKAMP total score)

Duration of effect of 50% reduction in SKAMP total score will be determined by persistence of the 50% reduction from the pre-dose observation in SKAMP total score level. The end of effect will be determined as the first timepoint after onset of effect at which 50% reduction from the pre-dose observation in SKAMP total score is not observed. The duration of effect will be calculated as the difference between the end of effect and onset of effect, in hours.

Age

Age will be calculated as (date of informed consent – date of birth + 1) / 365.25.

Baseline

Baseline is defined as data collected at Visit 0 (Day 0). If data was not collected at Visit 0, the last data available prior to Visit 0 will be used as the baseline data. However during the double

blind crossover phase change from baseline analyses will be performed taking the pre-dose data at the same classroom visit to be the baseline.

Body Mass Index

Body Mass Index (BMI) will be calculated as weight in kg / (screening height in m)^2.

Prior and Concomitant Medications

Prior medications are defined as those taken prior to the date of first dose of study medication. Concomitant medications are defined as those on or after the date of first dose of study medication through and including the day of the safety follow-up visit. Medications with a missing or incomplete start date will be assumed to be concomitant.

Concomitant medications will be assigned to either the Dose-Optimization Period or the double-blind Treatment Period of the study according to the start date, and will be summarized separately. Concomitant medications for which a missing or incomplete start date makes it impossible to determine which period of the study they started in will be counted as starting in the Dose-Optimization Period.

Treatment-emergent adverse event

A treatment-emergent adverse event (TEAE) is one which first occurs or worsens in severity after the first patch application and not later than 30 days following last patch removal. Adverse events are collected through the last study visit. Those AEs with a start date equal to the date of first patch application will be considered as TEAEs.

Adverse events for which a missing or incomplete start date makes it impossible to determine whether or not they were treatment-emergent will be counted as being treatment-emergent.

Adverse events will be assigned to either the Dose-Optimization Period or the double-blind Treatment Period of the study according to the start date, and will be summarized separately. Adverse events for which a missing or incomplete start date makes it impossible to determine which period of the study it started in will be counted as starting in the Dose-Optimization Period

6. ANALYSIS POPULATIONS

6.1 FULL ANALYSIS SET

The full analysis set (FAS) includes all consented and randomized subjects who have taken at least one dose of study medication. The FAS population will be used as the primary population for analysis of efficacy endpoints.

6.2 COMPLETERS

The completers population includes all consented and randomized subjects who

- received full prescribed dose of the double-blind study medication at both test laboratory classroom sessions
- completed the full classroom tests on both test classroom sessions
- did not miss more than four consecutive days of therapy during the double blind treatment period

- did not use prohibited concomitant medications during the double blind treatment period

The Completers population will be used for supportive analyses of efficacy endpoints.

6.3 SAFETY

The safety (SP) population includes all subjects who have taken at least one dose of the study medication and have at least one post dose safety measurement (including dermal assessments). In the unlikely event that errors may have occurred in treatment arm assignments, analyses using the Safety population will be based on treatment actually received. The SP population will be used for the analysis of dermal evaluations and safety endpoints.

7. STATISTICAL METHODS

Summaries will be presented for the SP population for all subjects combined during the Dose-Optimization Period of the study. During the double-blind Treatment Period of the study, summaries will be presented for the FAS, Completers, and SP populations by treatment (i.e, d-ATS and placebo, with all d-ATS dose levels combined), and some tables by treatment and treatment sequence. Subjects are intended to receive both treatments during the double-blind cross-over period of the study, so summaries for Visits 6 and 7 will be combined, with subjects contributing data to both treatment columns for the combined time point. As a result, the numbers of subjects in each treatment arm in the header of these summaries will reflect the total number of subjects receiving the treatment, and the number of subjects contributing data for each visit will be presented in the body of the table.

Early termination visits will be mapped to the next scheduled visit for inclusion in summary tables, where appropriate. For example, a subject who terminates after Visit 3 (e.g., has assessments at Visits 1 - 3 and early termination assessments) would have assessments from early termination mapped to the next scheduled visit for each assessment. Different assessments may be mapped to different visits; in this example, most assessments would be mapped to Visit 4, ECG would be mapped to Visit 5 as no ECG assessments are scheduled to be performed at Visit 4. This mapping is considered part of the observed values, and will be conducted before any summaries or analyses are performed on observed values. Similarly, it will be conducted before any scoring algorithm is applied. Listings will present early termination data as the mapped visit, as well as under 'ET' Visit as well.

Unless otherwise noted, categorical variables will be summarized using counts and percentages. Percentages will be rounded to one decimal place, except 0% and 100% will be displayed without any decimal places. Continuous variables will be summarized using the number of observations (n), mean, standard deviation (SD), median, minimum, and maximum. All analyses will use SAS® version 9.1 or higher.

All raw data will be presented to the original number of decimal places. The means and medians will be presented to 1 more decimal place than the raw data. The standard deviations will be presented to 2 more decimal places than the raw data.

To preserve the Type I error rate in analyzing the secondary endpoints of onset and duration of efficacy, a closed test procedure will be used. The closed test principle states that the trial-wise error rate can be controlled by the ordered testing of each hypothesis in the closed family using a suitable α level significance. A condition of the Closed Test procedure is that the order of testing

of the endpoints in a family matters. Thus hypotheses in a family must be ordered and the testing of each subsequent hypothesis in the order can only proceed if the testing of the prior hypothesis achieves a significant p-value with Type I error rate of predetermined α level, which will be set at α= 0.05 for this study. Testing of the hypotheses in the family proceeds until a hypothesis fails to achieve a significant p-value. If/when a hypothesis in the family fails to achieve significance, then significant p-values cannot be claimed for all subsequent hypotheses in that family, i.e., once one hypothesis fails to achieve significance, testing should cease.

If the primary efficacy endpoint, difference between d-ATS and placebo in mean SKAMP total score for the double-blind phase, is statistically significant (i.e., $p<0.05$), the secondary variables of onset and duration of efficacy will be tested using a closed testing procedure. The closed testing procedure starts from the time-point of 1 hour post-morning dose, then 2, 3, 4.5, 6, 7, 9, 10 and 12 hours post-dose. At each timepoint the order of testing will be onset of efficacy followed by duration of efficacy, before moving to the next timepoint. Duration of efficacy will not be considered until onset of efficacy has been determined.

All data collected will be presented in data listings, and the listings will be sorted by investigator/site, subject, and time point.

7.1 EFFICACY ANALYSES

The primary population for efficacy assessments will be the FAS population. The Completers population will be used to conduct sensitivity analyses. Efficacy summaries will be presented by visit by treatment arm. Variables that are collected in both the Dose-Optimization and double-blind Treatment Periods will be combined on the same table, but no summaries will be presented in the placebo column for the Dose-Optimization Period.

7.1.1 Primary Variable

Data Summaries

Continuous summary statistics will be presented for the mean SKAMP total score for Visits 6 and 7 for the FAS. Similar summaries will be presented for the Completers population.

Primary Efficacy Analysis

A likelihood-based Mixed-effect Model Repeated Measures (MMRM) will be used to assess the primary objective.

In the model, the SKAMP total score for each timepoint will be analyzed with the fixed effects of sequence (two levels), period (two levels), treatment (two levels) and time (10 levels, one for each of 30 minutes prior and 1, 2, 3, 4.5, 6, 7, 9, 10, and 12 hours post dosing) while the repeated measures effect of timepoint will be defined for subject within sequence and "Variance Components (VC)' covariance structure. An example of the SAS procedure to be used is presented below.

```
proc glimmix data=<dataset>;
  class <treatment sequence period timepoint subject>;
  model <SKAMP> = <treatment>
                  <sequence>
                  <period>
                  <timepoint>
```

```
                <treatment * timepoint>
                /solution ddfm=kr;

 random <timepoint>  / type=vc subject=<subject * period> residual;
 lsmeans <treatment> / pdiff cl alpha=0.05;

 estimate 'd-ATS - Placebo' <treatment> -1 1 / cl;
run;
```

Note: <italics> identifies where actual variable name substitutions are required.

<sequence> will be required to be defined as <treatment * period> to ensure estimates are obtainable.The two treatment levels will be: d-ATS (5, 10, 15, 20 mg doses combined) and placebo. The treatment * time interaction is included in the model to allow the form of the relationship to be modeled. The purpose of including the interaction term is not to test it for significance and so p-value will not be presented. Raw means, least-squares (LS) means, differences in LS means, and 95% confidence interval (CI) for the difference between treatment arms, and p-values will be calculated for SKAMP total score over the treatment day for the FAS (primary), and Completers populations.

Scaled residual and random effects plots will be considered to check the assumptions behind the mixed model.

Assessment of Carryover

The sequence effect (treatment x period interaction) will be tested at the significance level of 0.10. If significant differences are detected between sequences, data will be reviewed and strength of evidence indicating a crossover effect will be considered. In this case data from Visit 6 only will be presented to aid in assessment of the treatment effect. When only one period of data is used then the efficiencies gained from conducting a crossover trial design may be lost and there is concern associated with treating data as a parallel study due to the sample size being based up on the assumed efficiencies and may no longer be large enough to provide sufficient power. In this case, permutation test will be used to validate the results. Sampling without replacement (i.e. duplicate samples are not allowed) will be performed for 5000 permutations. The average mean treatment difference will be presented along with the associated 95% confidence interval and p-value for the permutation test. This will be compared against a significance level of 0.05.

Effect Modification Analyses

The primary efficacy analysis on the FAS population will also be conducted for the following subgroups:

- Investigator site
- Optimized dose
- Gender
- ADHD type (inattentive, hyperactive/impulsive, combined, not otherwise specified)
- Baseline ADHD severity (defined as the pre-dose ADHD-RS-IV rating scale for Baseline, Visit 0)

- Age group

To ensure summaries from the subgroups are informative, these will only be performed where subgroups include 5 or more subjects in each. Where subgroups have less than 5 subjects the small group will be combined with the next group, if there is a consecutive order to the groupings or with the next smallest group if there is no order. Baseline ADHD-RS-IV total score will be categorized into the following groups: 0-18; 19-36; and 37-54. Age will be categorized into groups for 6-12 years and 13-17 years.

7.1.2 Missing Data

MMRM is the best method to control type I error rate and minimize bias in the presence of missing data, although this does not fully overcome the issues and assumptions are still required to be made regarding the missing data mechanisms. Nevertheless, in the presence of a high drop out rate performance of sensitivity analysis is critical. The purpose of sensitivity analysis is to explore whether different analyses under different set of missingness assumptions provide robust efficacy results. The extent to which efficacy results are stable across such analysis provides confidence in the statistical conclusions.

Missing data can be considered to be missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR) according to the National Academy of Sciences (2010).

To explore the effect of missing data the following sensitivity analyses will be conducted for the primary efficacy analysis for the FAS population. Under the assumption that missing data is missing at random and the analyses require the assumption of MAR to hold:

- Imputation based analysis
    - Multiple Imputation

      Subjects with missing SKAMP total scores at any timepoint during Visits 6 and 7 will have these individual SKAMP total score values imputed.

      <u>Step 1: Imputation</u>

      For each missing SKAMP total score 20 imputations will be made using SAS PROC MI on the unimputed SKAMP total score efficacy data. All imputed data will be non-negative. For intermittent missing data the Markov Chain Monte Carlo (MCMC) simulation will be used, the number of burn in iterations will be set at 200, or higher if required for convergence. The thinning option will be set to select one out of every 100 iterations to ensure sufficient separation from previous iterations.

      Once intermittent missing data has been imputed monotone missing data will be imputed using a regression model including factors for treatment and period and the SKAMP total score in the VAR statement.

      The number of multiple imputations will be 20, leading to 20 completed data sets.

Step 2: Analysis of completed data sets

The imputed data for the primary endpoint will be analyzed as per the method described in section 7.1.1. This will lead to 20 estimates for the comparison of d-ATS with placebo.

Step 3: Inference

The estimates will be averaged and the associated standard errors will be summarized based on within-imputation and between-imputation variance, as is customarily done in multiple imputation. In other words, PROC MIANALYZE will summarize the 20 model estimates yielding a final estimate with associated 95% CI and p-value.

- LOCF

Subjects who started a classroom day but did not complete all the assessments will have their last observation within the same classroom day carried forward. Data from one classroom day will not be used to impute values for another test classroom day.

The completed LOCF data will be analyzed as per the method described in section 7.1.1.

The following sensitivity analyses will be conducted for the primary efficacy analysis for the FAS population under the assumption of some data MNAR:

- Worst-case scenario analyses
  - Scenario One

Subjects with missing data which is considered due to a defined 'poor outcome' will have the worst case (largest) SKAMP total score for that subject within the same classroom day imputed. If no SKAMP total score is available for a subject in a classroom day the overall observed worst case (largest) SKAMP total score for the same treatment will be imputed. For subjects with missing data which is not considered due to the defined 'poor outcome' reasons the mean SKAMP total score per treatment will be imputed.

Poor outcome reasons are defined as missing data due to poor tolerability and withdrawal due to lack of efficacy, AE or death.

Poor outcome reasons will be reviewed on blinded data and defined prior to database lock.

- Scenario Two

Results of two analyses will be presented for comparison. The first will assign the best possible outcome to missing values in the placebo group and the second will assign the worst possible outcome to those of the d-ATS group.

Sensitivity analysis using different methods of handling missing data will provide useful information about effectiveness of the study drug and confidence in the reliability of the conclusions drawn.

7.1.3 Secondary Variables

7.1.3.1 SKAMP

The onset of efficacy, based on the SKAMP total scores, will be defined as the time of the first assessment time showing statistical significance between d-ATS and placebo. If no significant difference is found at any time point for the onset of efficacy, no onset of efficacy will be deemed to have occurred, and the onset of efficacy will be defined as 'none'.

Onset of efficacy as measured by the SKAMP total score (defined in Section 5) will be analyzed using the same MMRM model as for the primary analysis. In the model, the SKAMP total scores for each timepoint will be analyzed with the fixed effects of sequence (two levels), period (two levels), treatment (two levels) and time (10 levels, one for each of 30 minutes prior and 1, 2, 3, 4.5, 6, 7, 9, 10, and 12 hours post dosing) while the repeated effect will be defined for subject-within-sequence. The interaction term for treatment * time will be included in the model, but not tested for significance. The example SAS code shown in section 7.1.1 may be used for this analysis, substituting the estimate statement given (mean estimates across timepoints) for the following set of statements (mean estimates at each timepint):

```
estimate 'd-ATS – Placebo 1hr post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 -1 0 0 0 0 0 0 0 0   0 1 0 0 0 0 0 0 0 0
      / cl;

estimate 'd-ATS – Placebo 2hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 -1 0 0 0 0 0 0 0   0 0 1 0 0 0 0 0 0 0
      / cl;

estimate 'd-ATS – Placebo 3hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 0 -1 0 0 0 0 0 0   0 0 0 1 0 0 0 0 0 0
      / cl;

estimate 'd-ATS – Placebo 4.5hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 0 0 -1 0 0 0 0 0   0 0 0 0 1 0 0 0 0 0
      / cl;

estimate 'd-ATS – Placebo 6hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 0 0 0 -1 0 0 0 0   0 0 0 0 0 1 0 0 0 0
      / cl;

estimate 'd-ATS – Placebo 7hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 0 0 0 0 -1 0 0 0   0 0 0 0 0 0 1 0 0 0
      / cl;
```

```
estimate 'd-ATS – Placebo 9hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 0 0 0 0 0 -1 0 0   0 0 0 0 0 0 0 1 0 0
      / cl;

estimate 'd-ATS – Placebo 10hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 0 0 0 0 0 0 -1 0   0 0 0 0 0 0 0 0 1 0
      / cl;

estimate 'd-ATS – Placebo 12hrs post dose'
      <treatment> -1 1
      <treatment*timepoint> 0 0 0 0 0 0 0 0 0 -1   0 0 0 0 0 0 0 0 0 1
      / cl;
```

The two treatment levels will be: d-ATS (5, 10, 15, 20 mg doses combined) and placebo. Continuous summary statistics will be presented for the SKAMP total score at each time point for Visits 6 and 7 for the FAS population. Raw means, LS means, differences in LS means, and 95% CI for the difference between treatment arms, and p-values will be calculated for each post-dose time point and for mean score over the treatment day for the FAS population.

The SKAMP total scores will also be plotted by time point for the FAS population.

Scaled residual and random effects plots will be considered to check the assumptions behind the mixed model.

The assessment of treatment differences will follow the closed testing procedure described in Section 7. The closed testing procedure starts from the time-point of 1 hour post-morning dose, then 2, 3, 4.5, 6, 7, 9, 10 and 12 hours post-dose. At each timepoint the order of testing will be onset of efficacy followed by duration of efficacy, before moving to the next timepoint. Duration of efficacy will not be considered until onset of efficacy has been determined.

- The onset time of efficacy action will be determined as the first post-dose time where the difference between the two treatments is statistically significant (i.e., $p \leq 0.05$). .
- If the difference between the two treatments is statistically significant (i.e., $p \leq 0.05$) at the 1 hours post-dose time point, the duration of efficacy will be claimed as the difference in hours between the end of efficacy; the last consecutive time point at which the difference is still statistically significant (i.e., $p \leq 0.05$) and onset of efficacy. If there is no time point at which there is a non-significant difference between the two treatment groups after a time point in which there is a significant difference between the two treatment groups, then the end of efficacy will be defined as 12 hours post-dose.

For example, if a statistically significant difference in mean SKAMP total scores for d-ATS vs. placebo was determined at 1 hour post-dose and statistical significance was measured at all time-points up to and including 10 hours post-dose, but statistical significance was not reached at 12 hours post-dose, onset of clinical efficacy and end of clinical efficacy would be defined as 1 hours post-dose and 10 hours post-dose, respectively, with duration of efficacy being 9 hours.

Supportive analysis of onset of efficacy and duration of efficacy will be conducted using the same model for the Completers population.

Onset and duration of effect as measured by the SKAMP total score (defined in Section 5) will be analyzed using two methods, firstly an MMRM model approach will be considered for the analysis of Onset and Duration of efficacy. Secondly a summary measures and survival analysis approach for onset and duration of effect for 50% reduction in SKAMP total score will be conducted.

Onset of effect for 50% reduction in SKAMP total score is defined for each subject as the first time point at which there is a 50% reduction in SKAMP total score from the pre-dose observation on the same classroom day.

Continuous summary statistics will be presented for the SKAMP total score at each time point for Visits 6 and 7, for the time of effect onset and for the duration of effect (hours) for the FAS.

Time of onset of effect for 50% reduction in SKAMP total score will be summarized descriptively and displayed graphically using the Kaplan-Meier approach for the FAS. The upper and lower quartiles (and CI) as well as the 50th percentile (and CI) will be used to estimate median time of onset of effect in each treatment arm. The difference between treatment arms, d-ATS and placebo, will be assessed using the Log-Rank test. Subjects that do not record a 50% reduction in SKAMP total score from the baseline value will be deemed to be censored at the last timepoint a SKAMP total score is recorded, for each classroom day.

Duration of effect for 50% reduction in SKAMP total score will be determined by persistence of the reduction of 50% from the pre-dose observation in SKAMP total score. The end of effect will be determined as the first timepoint after onset of effect at which the 50% reduction in SKAMP total score is not observed. The duration of effect will be calculated as the difference between the end of effect and onset of effect, in hours. Duration of effect will be analyzed using the same methodology as the onset of effect. Subjects that record a 50% reduction in the SKAMP total score from the baseline value such that duration of effect continues at their final assessment time for that classroom day will be censored at that time. Subjects without an onset of effect will be excluded from the analysis of duration of effect. A sensitivity analysis will be performed where subjects without an onset of effect will be imputed to have duration of effect of 0 hours.

Supportive analysis of onset of efficacy and duration of efficacy will be conducted using the same approach for the Completers population.

Once onset and duration of efficacy are estimated (using the MMRM approach analysis for the FAS, as described above), the sensitivity analysis will be conducted for primary variable as specified in section 7.1.1 by excluding data from time points before onset and after the last time point included into the duration of efficacy.

All data, including derived values, will be listed.

7.1.3.2 PERMP

Continuous summary statistics will be presented for PERMP-A and PERMP-C scores on the PERMP math test at screening and at each protocol-specified time point during the double blind period for the FAS. The PERMP-A and PERMP-C scores will also be plotted by time point for the FAS.

The number of items attempted and the number of items correct at each time point after dosing for Visits 6 and 7 will be analyzed in the same way as for the SKAMP primary efficacy analysis.

Raw means, LS means, differences in LS means, and 95% CI for the difference between treatment arms, p-values, and model results will be presented for each variable at each time point after dosing.

Change from pre-dose for PERMP-A and PERMP-C will be examined. Each post-dose PERMP-A and PERMP-C score will be analyzed by t-test on the change from pre-dose within group.

7.1.3.3 ADHD-RS-IV

Continuous summary statistics will be presented for the ADHD-RS-IV Total, Inattention, and Hyperactivity-Impulsivity scores at each visit from baseline through Visit 7 and for the change from baseline to each subsequent visit for the FAS. The ADHD-RS-IV Total and subscale scores will also be plotted by visit for the FAS.

The ADHD-RS-IV Total and subscale scores at Visits 6 and 7 will be analyzed in a similar way to the primary efficacy analysis, with the exclusion of timepoint from the model. Raw means, LS means, differences in LS means, and 95% CI for the difference between treatment groups, p-values, and model results will be presented for each variable.

7.1.3.4 CPRS-R:S

Continuous summary statistics will be presented for the CPRS-R:S Total, Oppositional, Cognitive Problems/Inattention, Hyperactivity, and ADHD Index scores at each visit from Visit 1 through Visit 7 for the FAS. The CPRS-R:S Total and subscale scores will also be plotted by visit for the FAS.

The CPRS-R:S Total, Oppositional, Cognitive Problems/Inattention, Hyperactivity, and ADHD Index scores at Visits 6 and 7 will be analyzed in a similar way to the primary efficacy analysis, with the exclusion of timepoint from the model. Raw means, LS means, differences in LS means, and 95% CI for the difference between treatment arms, p-values, and model results will be presented for each variable.

7.1.3.5 CGI

Categorical summary statistics will be presented for the CGI-S scores at baseline for the FAS. Categorical summary statistics will also be presented for the CGI-I scores at each visit after baseline for the FAS. The categorical CGI-I summaries will be presented both for all categories of the scale and for the dichotomized version of the scale defined in Section 4.2.5. The percentages of subjects who improved according to the dichotomized version of the scale will also be plotted by visit for the FAS.

Differences between the treatment arms in the dichotomized versions of the CGI-I scores at Visits 6 and 7 will be analyzed by using McNemar's test. In order to adjust for a potential period effect, the Mainland-Gart test, which is a test of sequence versus period preference, will also be performed. Finally, Prescott's test, which extends the Mainland-Gart test by including subjects who show no preference between the two periods, will also be conducted. Test statistics and p-values for the difference between the treatment arms will be presented for each test.

7.2 SUBJECT DISPOSITION

There will be a clear accounting of all subjects who are enrolled onto dose optimization. The following will be summarized by treatment sequence group for all enrolled subjects:

- The number of subjects who are enrolled onto dose optimization
- The number of subjects who receive study medication
- The number of subjects in Safety population
- The number of subjects who are randomized into double-blind
- The number of subjects in FAS population
- The number of subjects in Completers population
- The number of subjects who complete the study
- The number of subjects who discontinued the study and reason for discontinuation

The reasons for all study discontinuations will be summarized by treatment sequence, by treatment arm and study period, and will also be provided as a by-subject listing. The number and percentage of randomized subjects in each of the three efficacy study populations will be summarized for all subjects combined and for each treatment sequence.

7.3 PROTOCOL DEVIATIONS AND VIOLATIONS

All major protocol deviations or violations (PDVs) for each subject will be reviewed on blinded data prior to database lock to evaluate whether the subject has been protocol compliant. Major PDVs will be assessed as to whether or not they have a significant impact on the assessment of efficacy.

Major protocol violations will include, but are not limited to the following:

- Dosing non-compliance
- Deviation from visit window during the double-blind phase
- Non-permitted concomitant medications
- Violation of Inclusion/Exclusion Criteria

The number of subjects with each type of major PDV will be tabulated by category and by deviation type for all subjects in the SP population. A listing of all PDVs including deviation date, deviation type, deviation description and any relevant comments will be generated.

7.4 TREATMENTS

7.4.1 Extent of Study Drug Exposure

The duration (days) of study drug dosing will be determined as follows for each subject and summarized by treatment arm and overall for both SP and Completers populations:

Duration of study drug dosing overall for the double blind period (days) = date of last dose – date of first dose of double blind treatment + 1.

The date of first dose of double blind treatment will be Visit 5 + 1 day according to the protocol.

Duration of study drug dosing per visit (days) = date of last dose at that visit – previous visit date

The percent compliance with study medication will be calculated at each visit and for the overall double-blind Treatment Period.

Percent compliance = {[(Total number of patches dispensed) – (Total number of patches returned unused)]/ number of days in the interval} * 100%

The duration of study drug dosing and percent compliance will be summarized by visit and overall for both Safety and Completers populations.

Treatment dose at each visit during the dose-optimization phase will be summarized for the Safety population. A summary table for treatment dose for each treatment and period for the double-blind phase will be also presented.

All dosing and compliance data also will be provided as data listings.

7.4.2 Concomitant Medications

All prior and concomitant medications recorded in the electronic case report form (eCRF) will be coded to generic term and all matching Anatomic Therapeutic Classification (ATC) codes using the current version of WHO Drug. Summaries will be prepared using the coded generic term.

The number and percentage of subjects who had prior medications will be provided by preferred drug name for the SP population. The numbers and percentages of subjects who had concomitant medications during the Dose-Optimization Period and during the double-blind Treatment Period will be presented separately for the SP population.

All prior and concomitant medications recorded in the eCRF will be listed.

7.5 DEMOGRAPHIC AND BASELINE CHARACTERISTICS

Demographics and baseline characteristic data will be summarized by treatment arm for both the SP population and FAS. Demographic and baseline characteristics include, but are not limited to, age, age group, gender, race, ethnicity, weight, height, BMI, ADHD subtype, ADHD-RS-IV total (continuous and group) and subscale scores at baseline. Age will be calculated based on the date of informed consent, and the weight and height measurements will be those measured at screening. A category of ‘Multiple’ will be defined for race to summarize those subjects selecting more than one; all selections will be listed.

7.6 DERMAL EVALUATIONS

Summaries of dermal evaluations will be based on the SP population. Summary statistics will be presented separately for the SP population on all subjects combined during the Dose-Optimization Period and by treatment arm during the double-blind Treatment Period.

No statistical analyses will be performed for the dermal evaluation data.

7.6.1.1 ADHESIONS

The source data are the adhesion scores of the patches. Data from the in-clinic assessments will be summarized for each treatment arm.

- Frequency table showing the number of patches with each adhesion score at each evaluation time point.
- The number of patches that detached will be summarized.
- For detached patches, the duration of patch wear before patch detachment will be summarized (time of patch detachment will be subtracted from time of application).

7.6.1.2 IRRITATION

The source data will be the actual irritation scores recorded following visual evaluation of the application sites. If a patch is removed due to irritation, Last Observation carried Forward (LOCF) will be applied for all remaining time points. LOCF is defined as the highest irritation score observed prior to discontinuation of the patch due to irritation.

The letter grades will be converted to numerical equivalents in the following way: N (none)=0, A=0, B=1, C=2, and F, G, and H=3. For each subject, a combined score will be derived by adding the numerical grade and the numerical equivalent of the letter grade at each evaluation time point (e.g., 2C=2+2=4). Missing values for 'Other Effects" when the applicable numerical value is "0" will be calculated as a score of "N" with a numerical value of "0". Analyses for this data will be identified as combined scores.

The following data from the in-clinic assessments will be provided for each treatment arm:

- Frequency table with each combined "Dermal Response" and "Other Effect" score using worst case for all subjects.
- Frequency table with each combined "Dermal Response" and "Other Effect" score using worst case for subjects who discontinued due to irritation.
- Summary of number of patches that were removed due to irritation and number of patches that were removed due to irritation split by irritation combined score assessed at 30 minutes after detachment/removal

All irritation scores will be listed, including those captured in the dermal evaluation diary.

7.6.1.3 DISCOMFORT

The source data will be the discomfort scores of the patches. A summary of discomfort Scores recorded in clinic will present the discomfort scores and a summary of Description of Discomfort with the number and percent of subjects by each treatment and time point.

The Wong Baker Faces Scores (WBS) and Visual Analog Scale (VAS) scores will be presented separately and combined, by treatment for each timepoint. For the separate analyses the WBS and VAS will each be summarized by presenting mean, SD, median, minimum, maximum and IQR. For the combined analyses the numerical value of the scales will be taken and used to present combined continuous summary statistics. Additionally the WBS and VAS scores will be converted to the Numeric Rating Scale (NRS-11) and presented in combination under the Pain

Level with the number and percentage of subjects falling into each category. Subjects that do not provide a WBS or VAS score due to the discomfort assessment being 0 (no discomfort) will be include in the Rating level 0 and Pain Level of No Pain.

Numeric Rating Scale (NRS-11):	
Rating	Pain Level
0	No Pain
1-3	Mild Pain
4-6	Moderate Pain
7-10	Severe Pain

All discomfort scores, including the WBS and VAS scores, will be listed, including those captured in the dermal evaluation diary.

7.6.1.4 ADHESIVE RESIDUE

The source data will be the adhesive residue scores of the patches. A summary of Adhesive Residue Scores recorded in clinic will present the adhesive residue scores with the number and percent of subjects by each treatment and time point. All adhesive residue data will be listed, including those captured in the dermal evaluation diary.

7.7 SAFETY ANALYSES

Safety data for the Dose-Optimization Period will be summarized using combined data from all subjects in the SP population.

Safety data from the double-blind Treatment Period will be summarized using data from each treatment arm where applicable in the SP population.

No inferential statistics are planned.

7.7.1 Adverse Events

Treatment-emergent adverse events will be assigned to either the open-label Dose-Optimization Period or the double-blind Treatment Period. TEAEs occurring in the safety follow up after the double blind Treatment Period will be reported under the double blind Treatment Period for the last treatment they received. TEAEs for which a missing or incomplete start and/or stop date makes it impossible to determine which period of the study they started in will be counted as starting in the Dose-Optimization Period. The subject's dose of study drug at the onset of the event will be determined for all TEAEs. If a missing or incomplete start date makes it impossible to determine the dose at onset, then the dose at onset will be set to the lowest dose in the study (i.e., 5 mg/ 4.76 cm^2). Adverse events that start before the first patch application will be summarized separately. For summaries, TEAEs with missing severity are assumed to be severe, and TEAEs with missing relationship to study drug are assumed to be related.

Summary statistics for TEAEs will be presented separately for the SP population on all subjects combined during the Dose-Optimization Period and by treatment arm during the double-blind Treatment Period. All summaries of adverse events will consist of the number of events and the numbers and percentages of subjects who had the events being summarized.

All TEAEs will be classified by SOC and PT using the latest version of MedDRA dictionary. All TEAEs that occur after first patch application up to 30 days following last patch application will be summarized. Subjects may have more than one TEAE per SOC or PT. Unless otherwise specified, at each level of subject summarization in reporting the incidence of the AE, a subject will be counted once if the subject reported one or more events.

The following summaries will be presented separately for the Dose-Optimization Period and the double-blind Treatment Period by the dose being taken at the onset of the events and by MedDRA SOC, and preferred term:

- Overall summary of TEAEs
- TEAEs
- Treatment-related TEAEs
- TEAEs leading to premature discontinuation of study drug
- TEAEs by intensity.

7.7.2 Deaths and Serious Adverse Events

The following summaries will be presented separately for the Dose-Optimization Period and the double-blind Treatment Period by the dose being taken at the onset of the events and by MedDRA SOC, and preferred term:

- Serious adverse events (SAEs)
- Treatment-related SAEs

All data for serious adverse events and deaths will be listed

7.7.3 Laboratory Data

A central laboratory will be used to analyze all blood and urine samples with the exception of the urine pregnancy tests. Laboratory data is only collected at Screening and Visit 7/End of Study (or Early Termination visit). For analyses of laboratory data, 'baseline' refers to the data collected at screening. The clinical laboratory assessment will consist of the following:

Hematology: hemoglobin, hematocrit, white blood cell count, red blood cell count, platelet count, basophils, neutrophils, bands, lymphocytes, monocytes, and eosinophils.

Serum Chemistry: total bilirubin, alkaline phosphatase, alanine transaminase, aspartate transaminase, creatine phosphokinase, lactate dehydrogenase, blood urea nitrogen, creatinine, glucose, total protein, albumin, sodium, potassium, chloride, calcium, inorganic phosphorus, and Hemoglobin A_{1c}.

Urinalysis: specific gravity, pH, protein, glucose, ketones, bilirubin, and blood.

Urine drug screen and cotinine (screening only)

Summary statistics of raw data and change from baseline to End of Study values for each laboratory parameter will be presented by treatment arm. Data will be summarized as appropriate to the variable type.

For change from baseline summaries, subjects with an undefined change from baseline, because of missing baseline data, will be excluded. Individual Subject Changes will be identified through shift tables. Shift tables will be presented for each laboratory parameter with counts and percentages of subjects, by treatment arm, for shift (change) from baseline, using the normal ranges from the central laboratory.

Where multiple results are provided for a parameter the first set of observations will be included in the summaries, all data will be listed. Similarly, if both automated and manual differential results are provided for a visit then the automated differentials will be summarized, all will be listed.

Laboratory data from unscheduled visits will be listed but will not be included in any summaries.

Results of urine pregnancy results will be listed but will not be included in any summaries.

7.7.4 Vital Signs

Vital signs, including pulse, systolic blood pressure, diastolic blood pressure, temperature, and respiratory rate will be summarized. Summary statistics of raw data and change from baseline values for each vital signs parameter will be presented by treatment arm and time point. Data will be summarized as appropriate to the variable type. For change from baseline summaries, the results collected at baseline (visit 0) will be used in these calculations or the last data available prior to Visit 0 as per Section 5 definitions. Subjects with an undefined change from baseline, because of missing baseline data, will be excluded. Shift tables will be presented for each vital sign parameter with counts and percentages of subjects, by treatment arm and time point, for shift (change) from baseline, using the normal ranges for vital signs parameters. Normal ranges for vital signs data are defined in Appendix 2.

Vital signs data from unscheduled visits will be listed but will not be included in any summaries.

7.7.5 Columbia Suicide Severity Rating Scale (Children's Version)

Summaries will be presented for each of the subscales (i.e. Suicidal Ideation and Suicidal Behavior).

Suicidal Ideation will be summarized for each of the following 5 ideation categories:

1. Wish to be dead
2. Non-specific active thoughts
3. Active suicidal ideation with any methods (not plan) without intent to act
4. Active suicidal ideation with some intent to act, without specific plan
5. Active suicidal ideation with specific plan and intent.

Descriptive summary of the number and percentages of subjects with the responses of "Yes" for each of the five categories will be presented by treatment arms and visit. The intensity and frequency for ideation will also be summarized descriptively by treatment arm and visit.

Suicidal Behavior will be summarized for each of the following 5 behavior categories:

1. Actual attempt
2. Interrupted intent
3. Aborted attempt
4. Preparatory Acts or Behavior
5. Suicidal behavior present during assessment period.

Descriptive summary of the number and percentages of subjects with the responses of "Yes" for each of the five categories will be presented by treatment arms and visit.

Additionally, the number of subjects who reported at least one occurrence of suicidal ideation and at least one occurrence of suicidal behavior, and subjects who reported at least one occurrence of suicidal ideation or at least one occurrence of suicidal behavior will be summarized by treatment arm and visit. For these summaries a subject will only be counted once in each in each sub-category.

7.7.6 Physical Examinations, ECGs, and Other Observations Related to Safety

There will be no separate physical examination listings since screening physical examination abnormalities were recorded as Medical History and post-randomization physical examination abnormalities were recorded as AEs.

All quantitative 12-lead ECG interval measurements (QRS, QT, QTc, etc.) will be summarized with descriptive statistics by visit, and treatment arm. Summary statistics of raw data and change from baseline values will be presented by treatment arm and time point. Data will be summarized as appropriate to the variable type. For change from baseline summaries, subjects with an undefined change from baseline, because of missing baseline data, will be excluded.

ECG assessments at each study visit will be evaluated to determine whether the ECG findings are: 1) Normal; 2) Abnormal, Not Clinically Significant and 3) Abnormal, Clinically Significant. Descriptive statistics (numbers and percentages of subjects in each of the three categories) will be presented by treatment arms for each visit. In addition shift tables will be also presented based on the above three categories of ECG interpretations, by treatment arm and time point.

ECG data from unscheduled visits will be listed but will not be included in any summaries.

Screening results from the Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II) and M.I.N.I. International Neuropsychiatric Interview (M.I.N.I. 6.0) are used as to assess inclusion/exclusion criteria and are not recorded. These data will not be listed or included in any summaries.

8. INTERIM ANALYSES

No interim analysis is planned for this study.

9. DATA REVIEW

Final data for analysis will be cleaned prior to receipt for statistical programming. The purpose of this section is to indicate the history of the data and the process used to ensure that the data are acceptable for statistical analysis prior to database lock.

9.1 DATA HANDLING AND TRANSFER

All of the data will come from the [redacted] data management group in SAS® dataset format (SAS version 9.1 or later). Coexistent diseases and AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), version 15.0 or later, to assign a system organ class (SOC) and preferred term (PT) to each AE. Prior and concomitant medications will be coded to preferred drug names using the latest available World Health Organization Drug Dictionary (WHODD).

More details about the data handling can be found in the Data Management Plan.

9.2 DATA SCREENING

Beyond the data screening built into the Data Management Plan, the programming of analysis datasets, tables, figures, and listings (TFL) provides additional data screening. Presumed data issues will be output into SAS logs identified by the word "Problem" and extracted from the logs by a SAS macro and sent to Data Management.

Prior to database lock and unblinding, the database will be frozen for a period of time during which no changes will be made to the database. A review of pre-freeze TFLs run on clean subjects and post-freeze, blinded TFLs run on the frozen database allow for further data screening prior to lock. The post-freeze blinded TFLs will be discussed with the sponsor in a data review meeting to identify any final data issues and seek corrections prior to database lock. The [redacted] statistician and the sponsor must approve database lock.

10. VALIDATION

[redacted] goal is to ensure that each TFL delivery is submitted to the highest level of quality. Quality control procedures will be documented separately in the study-specific quality control plan. The following describes the procedure typically followed at [redacted].

Derived datasets will be independently reprogrammed by a second programmer and the separate datasets produced must match 100%.

Tables will be independently reprogrammed by a second programmer for numeric results. Statisticians will be involved in the process of programming and validating tables that include inferential statistical results.

Figures will be checked for consistency against corresponding tables and listings, or independently reprogrammed if there are no corresponding tables or listings.

Listings will be checked for consistency against corresponding tables, figures, and derived datasets.

The entire set of TFLs will be checked for completeness and consistency prior to its delivery to the client by the lead analysis programmer, the lead statistician, and a senior level statistician, or above, who is not a member of the project team.

The validation process will be repeated any time TFLs are redelivered using different data. Execution of this validation process will be documented through the study Table of Programs which is provided to the client at study conclusion.

APPENDIX 1 GLOSSARY OF ABBREVIATIONS

Glossary of Abbreviations:	
ADHD	Attention Deficit Hyperactivity Disorder
ADHD-RS-IV	Attention-Deficit/Hyperactivity Disorder Rating Scale, Version IV
AE	Adverse Event
ATC	Anatomic Therapeutic Classification
BMI	Body Mass Index
C-SSRS	Columbia Suicide Severity Rating Scale
CGI	Clinical Global Impression
CGI-I	Clinical Global Impression-Improvement
CGI-S	Clinical Global Impression-Severity
CI	Confidence Interval
CPRS-R:S	Conners' Parent Rating Scale-Revised Short Form
CRF	Case Report Form
C-SSRS	Columbia Suicide Severity Rating Scale
CSR	Clinical Study Report
d-ATS	d-Amphetamine Transdermal Drug Delivery System
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, Text Revision
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
FAS	Full analysis set
ICH	International Conference on Harmonization
ITT	Intent-To-Treat
IVRS	Interactive Voice Response System
LOCF	Last Observation carried Forward
LS	Least squares
MAR	Missing at random
MCAR	Missing completely at random
MedDRA	Medical Dictionary for Regulatory Activities
MNAR	Missing not at random
NRS-11	Numeric Rating Scale (11 point scale)
PDV	Protocol Deviation or Violation
PERMP	Permanent Product Measure of Performance
PERMP-A	Permanent Product Measure of Performance , Number of Problems Attempted

Glossary of Abbreviations:	
PERMP-C	Permanent Product Measure of Performance , Number of Problems Correct
PT	Preferred term
SA	Safety
SAP	Statistical Analysis Plan
SAE	Serious Adverse Event
SD	Standard Deviation
SKAMP	Swanson, Kotkin, Agler, M-Flynn and Pelham Scale
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TFL	Tables, Figures, and Listings
VAS	Visual Analog Scale
WASI-II	Wechsler Abbreviated Scale of Intelligence - Second Edition
WBS	Wong Baker Faces Scale
WHODD	World Health Organization Drug Dictionary