STATISTICAL ANALYSIS PLAN

Protocol Title	A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of Molindone Hydrochloride Extended-Release Tablets for the Treatment of Impulsive Aggression in Pediatric Patients with Attention Deficit/Hyperactivity Disorder (ADHD) in Conjunction with Standard ADHD Treatment
Protocol Number	810P302
Protocol Version	V7.0
Protocol Date:	29 March 2019
Study Drug	Molindone Hydrochloride Extended-Release Tablets (SPN-810)
Study Phase	III
IND Number	106,515
Sponsor	Supernus Pharmaceuticals, Inc. [redacted]
SAP Version	2.1
Date	21 January 21, 2020

SIGNATURE PAGE

Prepared by:

Reviewed by

Approved by

Table of Contents

LIST OF ABBREVIATIONS

AE	Adverse event
AESI	Adverse event of Special Interest
BID	Bis in die, twice daily
CRO	Contract Research Organization
CRF	Case report form
ECG	Electrocardiogram
FDA	Food and Drug Administration
GCP	Good Clinical Practice
IA	Impulsive Aggression
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
IND	Investigational New Drug Application
ITT	Intent-to-Treat
IV	Intravenous
IWRS	Interactive web-based randomization system
kg	kilogram
mcg	Microgram
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
msec	Millisecond
PK	Pharmacokinetic(s)
PO	Oral
PP	Per Protocol
PT	Preferred term in MedDRA and WHODD
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SM	Study Medication
SOC	System Organ Class in MedDRA
TDD	Total Daily Dose
TEAE	Treatment-emergent adverse event
TESAE	Treatment-emergent serious adverse event
WHODD	WHO Drug Dictionary

REVISION HISTORY

This is Version 2.1 of the Statistical Analysis Plan (SAP).

Summary of changes to the SAP from Version 2 submitted to FDA on 17Dec2019.

Section	Page	Description of Change	Rationale
7.1.2	25	The modified ITT is defined to exclude subjects with baseline IA frequency <6 per week	Retrospective Modified Overt Aggression Scale (R-MOAS) score was used for inclusion criteria but the primary analysis was analyzed based on data collected using IA diary Instrument. Based on discussion with KOLs, subjects with baseline IA frequency <6 per week for primary endpoint are considered to be normative. The original definition of the primary efficacy analysis population is modified based on advise from KOLs to detect treatment difference. This population will be called mITT and will be used for the primary analysis.
7.1.3	25	The ITT population is moved to Section 7.1.3	Per FDA advise to submit results under the original definition of the primary efficacy analysis population
7.8.1	29	The first paragraph and the first sentence of 2nd paragraph re-phrased	Clarification
7.8.2	30	Added a sentence	Clarification
7.8.6	33	Subsection 7.8.6.1 and 7.8.6.2 for key secondary analysis and additional secondary analysis created.	Clarification

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to ensure the credibility of the study findings by specifying the statistical approaches to the analysis of the data prior to database lock. This SAP covers the planned analyses of all data collected on the eCRFs and diary pages, and will identify handling of data issues. The statistical analysis plan presented in this document will supersede the statistical analysis methods described in the clinical protocol. Any deviations/changes from the planned analyses described in this SAP will be identified, with justification, in the appropriate section of the clinical study report. This SAP is based on the clinical study protocol 810P302, version 7.0, dated 29 March 2019 and its associated electronic case report forms (eCRF).

2 STUDY OBJECTIVES

2.1 *Primary Objective*

The primary objective is to assess the efficacy and safety of SPN-810 in reducing the frequency of impulsive aggression (IA) behaviors in pediatric patients with attention-deficit/hyperactivity disorder (ADHD) when taken in conjunction with standard ADHD treatment.

2.2 *Secondary Objectives*

2.2.1 *Key Secondary Objective*

The key secondary objective of the study is to assess the effect of SPN-810 on the Clinical Global Impression – Severity Scale (CGI-S).

2.2.2 *Secondary Objective*

The secondary objectives of the study are to assess the following:

- The effect of SPN-810 on the Investigator-rated Clinical Global Impression – Improvement Scale (CGI-I)
- The effect of SPN-810 on the child's overall health as measured by the Child Health Questionnaire Parent Form 28-item (CHQ-PF28)
- The effect of treating the child with SPN-810 on the parent-child relationship as measured by the Parenting Stress Index – Short Form version 4 (PSI-4-SF)
- The effect of SPN-810 on the caregiver-completed CGI-I
- The effect of SPN-810 on inattention, hyperactivity-impulsivity and oppositional defiant disorder measured by the SNAP-IV Rating Scale
- The effect of SPN-810 on the responder rate (defined as ≥ 50 % in the reduction of the frequency of IA behaviors)
- The effect of SPN-810 on the responder rate (defined as ≥ 30 % in the reduction of the frequency of IA behaviors)

2.3 *Tertiary Objectives*

[redacted]

[redacted]

3 STUDY DESIGN

3.1 *General Description*

Protocol 810P302 is a randomized, placebo-controlled, double blind, multicenter, parallel group, fixed dose study to demonstrate the efficacy, safety, and tolerability of SPN-810 in the treatment of IA in patients aged 6-12 years with ADHD in conjunction with standard ADHD treatment. The study design is presented schematically in Figure 1 and Figure 2. The study is divided into three phases: Pre-Treatment, Treatment, and Conversion/Taper.

Figure 1: Treatment Schedule (Conversion)

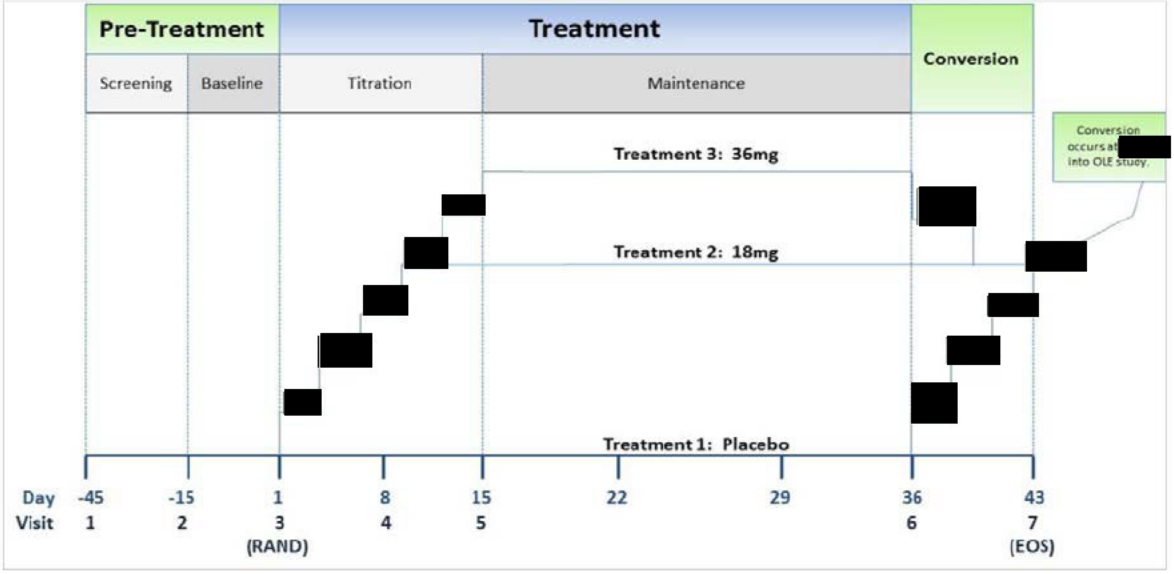
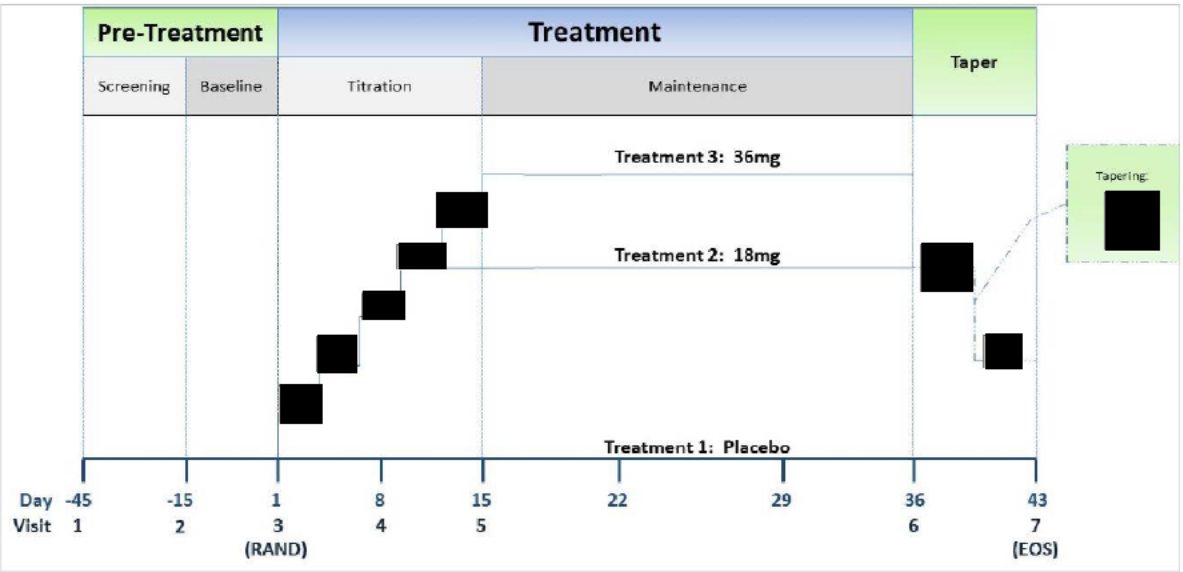


Figure 2: Treatment Schedule (Taper)



Following screening, eligible subjects will enter a flexible 15-day baseline period, at which time the IA diary will be issued to the subject's primary caregiver. At the end of the baseline period, eligible subjects whose primary caregiver has maintained at least 80% compliance with the IA diary will be randomized. Per the original randomization eligible subjects who complete the baseline period and meet the requirements for the double-blind study will be randomized at Visit 3 (Day 1) in a 1:1:1 ratio to receive 18 mg/day, 36 mg/day SPN-810 or placebo and proceed to the titration period, which will be two weeks. However, based on the 810P301 study Interim Analysis result decision, the 18 mg arm was dropped partway through the study. As a result, subjects planned to be randomized to the 18 mg arm were be re-allocated to the 36 mg or placebo arm in a ratio of 2:1.

Subjects will be titrated to maintenance dose over a period of 2 weeks. Following dose titration, subjects on active treatment will be maintained at their designated dose level for 3 weeks.

3.2 *Schedule of Visits and Procedures*

All subjects who are randomized and take any study medication (SM) will be followed according to the protocol regardless of the number of doses of SM taken, unless consent for follow-up is withdrawn. Table 1 below presents the schedule of visits and procedures for the study.

Table 1: Schedule of Visits and Procedures

Phase	Pre-treatment		Treatment				Conversion/ Taper
Period	Screening	Baseline	Titration		Maintenance		
VISIT NUMBER	1	2	3	4	5	6	7
DAY	-45	-15	1	8	15	36	43
WINDOW (DAYS)	≤45d prior to Visit 3	≥15d prior to Visit 3	0[c]	7d±2d from Visit 3	14d±2d from Visit 3	21d±3d from Visit 5	7d±1d from Visit 6
Informed Consent/Assent [a]	X [b]						
R-MOAS, K-SADS-PL 2013 & Vitiello Aggression Scale	X						
Medical History	X		X [d]				
Demographics	X						
Physical Examination	X					X	
ECG (12-lead)	X			X		X	
Inclusion/Exclusion Criteria	X						
Randomization			X [b,e]				
Urine Drug Screen	X		X				
Urine Pregnancy Test [f]	X		X				X
Diary Training & Distribution or Evaluation		X	X [e]	X	X	X	
Vital Signs[g]	X						X
Weight, height, BMI	X		X	X	X	X	X
Hematology/chemistry/Urinalysis	X		X			X	X
PK Blood Sampling				X [h]	X [h]		
Columbia Suicide Severity Rating Scale (CSSRS)	X		X	X	X	X	X
Investigator CGI-S	X		X	X	X	X	
Caregiver and investigator CGI-I				X	X	X	

Phase	Pre-treatment		Treatment				Conversion/ Taper
Period	Screening	Baseline	Titration		Maintenance		
VISIT NUMBER	1	2	3	4	5	6	7
DAY	-45	-15	1	8	15	36	43
WINDOW (DAYS)	≤45d prior to Visit 3	≥15d prior to Visit 3	0[c]	7d±2d from Visit 3	14d±2d from Visit 3	21d±3d from Visit 5	7d±1d from Visit 6
Efficacy scales (SNAP-IV, CHQ-PF28, PSI-4-SF)			X			X	
Safety Scales (Simpson-Angus, Barnes, AIMS)			X	X	X	X	X
Infrequent Behaviors Checklist			X	X	X	X	
Adverse Events			X	X	X	X	X
Concomitant Medications	X	X	X	X	X	X	X
Drug Dispensation			X [b]		X	X	
Drug Return and Compliance					X	X	X
Diary Return							X

a Written consent must be obtained prior to performing any study-related procedure.

b Access IWRS.

c Visit 3 will occur at least 15 days following Visit 2

d Assess for any clinically significant change in Medical History since screening.

e Diary compliance must be at least 80% (minimum of 12 days out of 15) to qualify for randomization.

f To be performed for female subjects of childbearing potential prior to administration of first dose of SM and will have to be tested as negative for the subject to continue in the study.

g Heart Rate (HR), blood pressure (BP), temperature, and respiratory rate will be measured.

h Total of 5 PK blood samples will be obtained over one or two visits (Visit 4 and/or Visit 5)

3.3 *Study Treatment*

3.3.1 *Treatment, Dose, and Mode of Administration*

SPN-810 (molindone hydrochloride extended-release tablet) or matching placebo tablets will be administered orally (PO) twice daily (BID). Subjects will be force-titrated over a period of 2 weeks to their final randomized dose at Visit 3**.** Subjects will be titrated up to the final randomized total daily dose (TDD).

- Treatment 1: placebo
- Treatment 2: 18mg total daily dose (TDD)
- Treatment 3: 36mg TDD

Based on the 810P301 study Interim Analysis result decision, the 18 mg dose arm was dropped partway through the study.

3.3.2 *Duration of Treatment and Study Duration*

Total subject duration on study is approximately 13 weeks.

- Pre-treatment phase: 45 days
  - Screening period: up to 30 days
  - Baseline period: at least 15 days
- Treatment phase: 5 weeks
  - Titration period: 2 weeks
  - Maintenance period: 3 weeks
- Conversion/Tapering phase: 1 week

3.3.3 *Methods of Assigning Subjects to Treatment Group*

Allocation of study drug will be completed centrally using an interactive web response system (IWRS) that will determine which kit to assign to the subject. Separate schedules for subject randomization and drug list will be created. The original randomization scheme assigns treatment to each randomization number in a 1:1:1. However, based on the 810P301 study Interim Analysis result decision, the 18 mg arm was dropped partway through the study. As a result, subjects planned to be randomized to the 18 mg arm will be re-allocated to the 36 mg or placebo arm in a ratio of 2:1.

3.3.4 *Blinding*

Subjects, all site personnel, the Sponsor and its designated Contract Research Organizations (CROs) are blinded with the following exceptions. The CRO and Sponsor personnel who are

involved in the randomization generation, study drug supply, and PK concentration assays are unblinded. However, they are not involved in the routine study monitoring, data cleaning and review, the design of this SAP or data analysis.

3.4 *Hypotheses*

The null and the alternative hypotheses are as in the following.

- H_0: There is no difference between SPN-810 36 mg and placebo in the treatment of IA in subjects with ADHD in conjunction with standard ADHD treatment.
- H_a: There is a difference between SPN-810 36 mg and placebo in the treatment of IA in subjects with ADHD in conjunction with standard ADHD treatment.

3.5 *Determination of Sample Size*

The original sample size was based on the results from the Phase 2 study. It was assumed that the average treatment difference between SPN-810 dose groups and placebo is 15 with a common standard deviation (SD) =27.3. A sample size of 77 subjects per arm (1:1:1 allocation ratio) would yield 90% power to detect a non-zero difference between the median of 18 mg or 36 mg dose group and the placebo using the Wilcoxon rank-sum test with a 2-sided significance level of alpha=0.05.

After the interim analysis for Study 810P301, the 18 mg dose arm was discontinued for randomization. Subjects planned to be randomized to the 18 mg arm were re-randomized to the 36 mg or placebo arms at a 2:1 allocation ratio. A blinded review of the 810P301 data indicates that the SD for the interim data was actually 34.83 instead of the protocol assumed 27.3. This new SD leads to a total sample size of 306 (adjusted for 20% attrition) subjects. Approximately 167 subjects were randomized prior to the implementation of the 2:1 randomization. It follows that 139 subjects would be randomized in a 2:1 ratio of 36 mg to following the re-calculation of sample size with the new SD.

The sample size was calculated using the nQuery Advisor Software, Version 7.

3.6 *Changes to Analyses Planned in the Protocol*

This SAP will be updated as necessary if any changes in the protocol occur due to the expected protocol amendment impacting the SAP.

4 DEFINITIONS AND DERIVED DATASETS

Table 1. Terminology and Definition

Terminology	Definition
Study Medication	SPN-810 or the matching placebo.
Study Day	The first dose date (Day 1) is defined as the date on which a subject took the first dose of double-blind study medication (SM), Other study days are defined relative to the Study Day 1. For visits prior to the first dose of the SM, Study Day is calculated as Visit Date – Day 1. For visits after the first dose, Study Day is calculated as Visit Date – Day 1 +1.
Enrolled	Subject has provided signed informed consent.
Randomized	Subject is assigned with the kit number for the SM.
Completer for the Study	Subject who completes the maintenance period (Visit 6).
Baseline	Unless specified otherwise, baseline is defined as the non-missing value collected most recent to and before the time of the very first dose of the SM. For example, Visit 3 CGI-S, CHQ-PF28, PSI-4-SF, and SNAP-IV are collected prior to the first dose. Their values are considered baseline for the covariate analysis.
Baseline Period	For subject daily IA diaries, on and prior to Visit 3 (inclusive, ≤Day 1).
Prior Medication	Medication collected on the Prior/Concomitant Medication CRF, with the start date prior to Study Day 1.
Concomitant Medication	Medication collected on the Prior/Concomitant Medication CRF, with end date on/after Study Day 1.
Titration Period	For subject daily IA diaries from Visit 3 through Visit 5 (v3< to ≤ V5)
Maintenance Period	For subject daily IA diaries from Visit 5 through Visit 6 (V5< to ≤ V6)
Treatment Period	Titration and Maintenance Periods combined

Terminology	Definition
Derived Dataset	Derived dataset (analysis dataset) is any permanently stored collection of data containing raw and/or derived variables created to support the production of statistical summary tables. Derived variables will be created from the "raw" data. Analysis datasets to be created will include ADEFF (efficacy analysis dataset), ADSL (subject-level analysis dataset), ADAE (Adverse event analysis dataset) and other as appropriate. In deriving ADEFF, the source data include both the episodic diary and evening diary data files. Specifications for derived datasets will be developed which will include the names and definitions of derived variables in the derived SAS datasets.

5 EFFICACY AND SAFETY VARIABLES

5.1 *Primary Efficacy Variable*

The primary efficacy endpoint is percent change (PCH_T) in the frequency (unweighted score) of IA behaviors per 7 days in the treatment (titration and maintenance) period relative to the Baseline period calculated over the number of days with non-missing IA diary data.

PCH_T will be calculated as 100 x (T – B)/B, where T and B are IA behavior frequencies per 7 days during the treatment period and baseline period, respectively. The IA behavior frequency per 7 days is defined as (SUM/DAY) x 7, where SUM is the total of the IA behaviors reported in the subject IA diary, and DAY is the number of days with non-missing IA score in the subject IA diary during the specified study period.

This primary efficacy variable will be based on a checklist of 15 IA behaviors collected in an electronic IA diary. The IA diary comprises two parts: 1) an episodic diary that will be used by the primary caregiver (or alternate) to enter events as soon as possible after they are observed; and 2) an evening diary that will prompt the caregiver to review events for the day and to enter any events that were not previously captured. Each event will be characterized by a checklist of 15 observed behaviors: Yelling, Screaming, Threatening, Scratching, Throwing, Slamming, Hitting Self, Arguing, Cursing, Name Calling, Shoving, Hair Pulling, Fighting, Hitting Others, and Kicking Others. The checklist will indicate whether each behavior was observed (coded 1) or was not observed (coded 0) during the incidence of an event. Each day can have multiple events. A day can have no event, as can be attested in the evening diary. In this case, if no event is reported

during a day, and the evening diary confirms this, the daily event score for that subject will be 0 for each of the 15 IA variables. Behaviors not on this list will not be captured.

5.2 *Secondary Efficacy Variables*

5.2.1 *Key Secondary Variables*

The key secondary efficacy variable is the change from Baseline (Visit 3) to Visit 6 in Investigator CGI-S score.

5.2.2 *Additional Secondary Variables*

1. Investigator CGI-I score at Visit 6
2. CHQ-PF28 score at Visit 6
    - Physical functioning summary score
    - Psychosocial health summary score
3. PSI-4-SF (Parenting Stress Index – Short Form) score at Visit 6 in
    - Parental Distress
    - Parent-Child Dysfunctional Interaction
    - Difficult Child
4. Caregiver CGI-I score at Visit 6
5. SNAP-IV ADHD (Swanson, Nolan and Pelham-IV Rating Scale for Inattention, hyperactivity-impulsivity and oppositional defiant disorder measured) score at Visit 6 in
    - Inattention ratings
    - Hyperactivity/Impulsivity ratings
    - Oppositional Defiant Disorder
    - Combined subscale ratings
6. % of responders with ≥50% reduction in the frequency of IA behaviors from baseline to the treatment period
7. % of responders with ≥30% reduction in the frequency of IA behaviors from baseline to the treatment period

The CGI scale was developed to provide a brief, stand-alone assessment of the clinician's view of a subject's global functioning prior to and after administration of a SM (Guy 1976). Severity of illness (CGI-S) and global improvement (CGI-I) are both rated on a scale of 1 to 7 with 7 being "extremely ill" or "very much worse", respectively. Successful therapy is indicated by a lower overall score in subsequent testing. Investigators should consider their total clinical experience with children who have IA associated with ADHD and rate how severe the subject's condition is at the time.

- **CGI-I**, relative to the condition at Visit 3, will be evaluated by the caregiver and by the Investigator at each post baseline visit on a 7-point scale with 1=Very much improved, 2=Much improved, 3=Minimally improved, 4=No change, 5=Minimally worse, 6=Much worse, and 7=Very much worse.
- **CGI-S** will be evaluated by the Investigator at –each visit on a 7-point scale with 1=Normal, 2=Borderline ill, 3=Mildly ill, 4=Moderately ill, 5=Markedly ill, 6=Severely ill, and 7=Extremely ill.

The Child Health Questionnaire Parent Form 28-item (**CHQ-PF28**) (Protocol Appendix 10.11) is evaluated by the primary caregiver at Visit 3 and Visit 6 for the subject's health status and health related quality of life. CHQ-PF28 items have four, five, or six response options, divided over eight multi-item scales (Physical Functioning [PF, Q2.1a–2.1c], Behavior [BE, Q5.1a-5.1c, Q5.2], Mental Health [MH, Q6.1a–6.1c], Self-Esteem [SE, Q7.1a-7.1c], General Health Perceptions [GH, Q1.1, Q8.1a-8.1c], Parental Impact - Emotional [PE, Q9.1a-9.1b], Parental Impact - Time [PT, Q9.2a-9.2b], and Family Activities [FA, Q9.3a-9.3b]) and five single item concepts (Global Health [GGH, Q1.1], Role/Social Limitations - Emotional/Behavior [REB, 3.1], Role/Social Limitations - Physical [RP, Q3.2], Bodily Pain/Discomfort [BP, Q4.1], Global Behavior [GBE, Q5.2], Family Cohesion [FC, Q9.4], and Change in Health [CH, Q8.2]). In addition, the individual scale scores will be aggregated to derive 2 summary component scores: the physical functioning and psychosocial health summary scores. Range on subscales and the overall scale is 0–100, where 0 is the worst possible health state and 100 the best possible health state. Details for item and multi-item scale scoring, calculation of standardized scores and Aggregate Physical and Psychological Summary Scores can be found in the Child Health Questionnaire (CHQ) Scoring and Interpretation Manual (HealthActCHQ).

The Parenting Stress Index – Short Form (**PSI-4-SF**) is a 36-item self-report measure of parenting stress at Visit 3 and Visit 6. Three subscales (Parental Distress [PD: Q1-12], Parent-Child Dysfunctional Interaction [P-CDI: Q13-24], and Difficult Child [DC: Q25-36]) consist of 12 items each. Parents use a 5-point scale to indicate the degree to which they agree with each statement. The subscale score is the sum of the respective item scores. "Total Stress" score is the sum of the 3 subscale scores. The subscale will be set to missing if 2 or more (>1) item scores in the subscale are missing. Total Stress score is set to missing if 1 or more subscale score is missing.

The **SNAP-IV** (Protocol Appendix 10.7) rating scale includes 18 ADHD and 8 oppositional defiant disorder (ODD) symptoms as specified in the DSM-IV-TR and International Statistical Classification of Diseases and Health Related Problems 10th Revision (ICD-10) Classification of Mental and Behavioral Disorders. The symptoms are scored by assigning a severity estimate for each symptom on a 4-point scale. The SNAP-IV rating is performed by the parent or legal

representative at each visit. The ratings from the SNAP-IV scale are grouped into the following 4 subscales:

- ADHD-Inattention (items #1-9),
- ADHD-Hyperactivity/Impulsivity (items #10-18)
- ODD (items #19-26)
- ADHD-Combined subscale: (items #1 – 18)

Each subscale score is the average of the available item scores in the subscale. The SNAP-IV rating scale will be administered at Visit 3 and Visit 6.

5.3 *Pharmacokinetic Variables*

Not applicable

5.4 *Safety Variables*

5.4.1 *Medical History and Adverse Events (AEs)*

AEs that occur on/after signing ICF will be recorded on the AE CRF. Medical history and AEs will be coded for preferred term (PT) and system organ class (SOC) using Medical Dictionary for Regulatory Activities (MedDRA). AEs are evaluated by the investigators for seriousness, relationship, and severity.

5.4.2 *Prior, Concomitant, and Prohibited Medications*

Prior medications and concomitant medications will be extracted from the Prior/Concomitant Medication CRF. Medications taken prior to Study Day 1 will be considered as prior medications and medications taken on or after Study Day 1 will be considered as concomitant medications. Prior and concomitant medications will be coded using WHO Drug Dictionary (WHODD) for PT and Anatomical Therapeutic Chemistry (ATC) classification.

Prohibited medications include, but are not limited to,

- α_2-adrenergic agonists (e.g. clonidine and guanfacine) used for any other reason except for monotherapy treatment for ADHD
- Anti-psychotics including aripiprazole, risperidone, quetiapine, and ziprasidone
- Anticonvulsants including carbamazepine and valproic acid, antidepressants, mood stabilizers including lithium, benzodiazepines, cholinesterase inhibitors or any drug known to inhibit CYP2D6 activity
- Herbal supplements

5.4.3 *Clinical Laboratory Tests (Hematology, Chemistry, and Urinalysis)*

Laboratory tests will be converted to follow the International System (SI) units. Laboratory tests will be collected at Visits 1, 3, 6, and 7. The tests include hematology, serum chemistry and urinalysis panels. Laboratory values will be categorized according to the appropriate laboratory reference ranges.

5.4.4 *Vital Signs, Height and Weight*

Vital signs include body temperature, heart rate, respiratory rate, systolic blood pressure, and diastolic blood pressure. They will be collected at Visits 1 and 7. Subject height, weight, and BMI will be collected at Visits 1 and 3 – 7.

5.4.5 *12-Lead ECG*

The following ECG variables will be collected at Visits 1, 4, and 6: HR (heart rate), PR (respiratory rate), QRS interval, QT interval, and QTcF (QT interval corrected using the Fridericia's method).

5.4.6 *Physical Examinations*

Physical examinations will be conducted at Visit 1 and Visit 6. Clinically significant findings will be reported as adverse events.

5.4.7 *Columbia Suicide Severity Rating Scale (C-SSRS)*

The C-SSRS (Protocol Appendix 10.6) is a questionnaire assessing suicidal ideation and behavior at baseline and since last visit. This assessment will be conducted at Visits 1, and 3 – 7. The following outcomes are C-SSRS categories and have binary responses (yes/no).

Category 1 – Wish to be dead
Category 2 – Non-specific Active Suicidal Thoughts
Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5 – Active Suicidal Ideation with Specific Plan and Intent
Category 6 – Preparatory Acts or Behavior
Category 7 – Aborted Attempt
Category 8 – Interrupted Attempt
Category 9 – Actual Attempt (non-fatal)
Category 10 – Completed Suicide

Three composite binary endpoints and one severity score are derived from the above categories.

Suicidal ideation: A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS.

Suicidal behavior: A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS.

Suicidal ideation or behavior: A "yes" answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.

Suicidal Ideation score: The maximum suicidal ideation category (1-5 on the C-SSRS) present at the assessment. Assign a score of 0 if no ideation is present.

5.4.8 *Infrequent Behaviors Checklist*

The infrequent behaviors checklist is a checklist of 15 behaviors that (along with the 15 IA Diary behaviors) were qualitatively linked to IA during the development of the IA diary. These behaviors include teasing, spitting, biting, weapons, ripping, breaking, vandalizing, destroying, fire setting, hitting animal, kicking self, kicking animal, severe injury self, severe injury others, severe injury animal. Caregivers will be asked which, if any, of these behaviors have been observed since the patient's last visit. This assessment will be administered at Visits 3 – 6 by the site coordinator.

5.4.9 *Extrapyramidal symptoms (EPS) scales*

The EPS symptom and neuroleptic malignant syndrome will be evaluated using the Simpson-Angus scale (Protocol Appendix 10.8), the Barnes Akathisia scale (Protocol Appendix 10.9), and the Abnormal Involuntary Movement scale (AIMS) (Protocol Appendix 10.10). The Simpson-Angus scale is a 10-item rating scale, consisting of 1 item measuring gait, 6 items measuring rigidity, and 3 items measuring glabella tap, tremor and salivation. The Barnes Akathisia scale is a rating scale for drug-induced akathisia. It includes Objective Akathisia, Subjective Awareness of Restlessness, Subjective Distress Related to Restlessness, and Global Clinical Assessment of Akathisia subscales. Additionally, total score will be calculated by summing up the Objective and two Subjective scale scores (Barnes 1989). The objective and two subjective symptoms are rated on a 4-point scale ranging from 0 to 3. The global clinical assessment is rated on a 5-point scale ranging from 0 to 4. The AIMS test is a rating scale used to measure tardive dyskinesia (Munetz 1988). It includes 12 items to rate involuntary movements of various areas of the subject's body. These assessments will be administered at Visit 3 and all subsequent visits.

6 STATISTICAL METHODS

6.1 *General Methodology*

All statistical analyses will be conducted using SAS Version 9.4 or later on the Microsoft® Windows Operating System.

All tabulations of results will include summaries for the three treatment groups: SPN810 36mg, SPN810 18mg and placebo, however, statistical test will only include SPN810 36mg and placebo.

Continuous data will be summarized using n (number of subjects with non-missing observations), mean, median, standard deviation (SD), minimum value, and maximum value. Categorical data will be summarized using the frequency count and percentage (n, %) of subjects in each category. Number of subjects with non-missing values or number of subjects with missing values (e.g., Not Done) will be presented, where appropriate. Subjects with missing values will not contribute to the denominator of percentage calculations. Counts of zero in any category will be presented without percentage.

Treatment differences will be evaluated as SPN-810 dose minus placebo. Unless otherwise specified, all statistical hypothesis tests will be 2-sided with a significance level of α=0.05. All confidence interval will be calculated 2-sided from the unadjusted 5% significance level. P-values will be reported to 4 decimal places. P-values <0.0001 will be presented as "<0.0001". Similarly, p-values >0.9999 will be presented as ">0.9999"

The precision rules for the presentation of summary statistics will be:

- Sample size (n, N) and number of missing responses (if displayed) – Integer
- Mean, median, percentiles, and confidence interval – One additional decimal place than reported/collected
- Standard deviation – Two additional decimal places than reported/collected
- Minimum, maximum – Same number of decimal places as reported/collected
- Ratios – two decimal places
- Percentages – one decimal place generally, or two decimal places for <0.1%

The data summaries will be accompanied by individual subject data listings. All data available from diary, eCRF, and external transfer (labs) will be listed and will include relevant and pertinent subject information, e.g., treatment, center, age/gender/race, and study day. The listings will be sorted in the order of treatment group, unique subject ID, and assessment date/time and assessment name. The unique subject ID consists of 3-digit study number followed by [redacted] digit site

number and [redacted] digit subject number separated by hyphen (e.g., [redacted]).

For each summary table, a supporting listing will be identified in the footnote.

Dates will be presented in the ISO-8601 format YYYY-MM-DD. Times will be displayed in 24-hour clock format. Numbering for tables, figures and listings will follow ICH E3 Guidelines.

Alternative methods of analysis of the data may be considered prior to un-blinding when some of the assumptions underlying the proposed analyses not be met. However, the reason for any departure from the planned approach and methods will be fully documented as an amendment to the SAP or in the clinical study report.

6.2 *Visit Window and Unscheduled Assessments*

There is no visit window for the by visit analysis. Data will be analyzed according to the visit they are associated with. Data collected at unscheduled time points will not be summarized at the unscheduled time points. However, the unscheduled data will be used for analysis if data are missing at the next scheduled visit due to early termination.

6.3 *Adjustment for Covariates*

An exploratory analysis of the primary efficacy endpoint will be performed using a randomization-based approach (Dmitrienko and Koch 2017). This analysis will include age (categorical and continuous), gender, pooled site and baseline IA adjusted to 7 days as covariates.

6.4 *Handling of Dropouts, Missing Data, and Data Discrepancies*

For the primary analysis of the primary efficacy endpoint, the frequency of IA behaviors during the treatment (titration + maintenance) period will be calculated over the number of days with non-missing IA diary data in the treatment period. No explicit imputation of missing data will be used, but this approach is implicitly equivalent to using the frequency of IA behaviors during the days with non-missing IA diary data to impute the frequency for days after study discontinuation or days with missing IA diary data. In addition to this implicit imputation, multiple imputations under the assumption of missing at random (MAR) and missing not at random (MNAR) will be used for handling missing data for sensitivity analysis. Details of imputation methods are specified in Section 7.8.2

For analysis of frequency of IA behaviors during the treatment period (titration or maintenance), the following methods will be followed.

1) If subjects discontinued the study during or prior to the maintenance period, their maintenance period will be counted from Visit 5 (exclusive) through the date of last dose as documented on the End of Study eCRF page (inclusive). If the date of last dose is not collected, it will be imputed by the subjects' last visit for the study.

2) If subjects have ≥7 daily diaries in the maintenance period, their maintenance IA frequencies will be based on maintenance daily diaries.

3) If subjects have <7 daily diaries in the maintenance period, their maintenance IA frequencies will be calculated pooling the titration period diaries with the maintenance period diaries.

4) If subjects do not have any maintenance diaries, their maintenance IA frequencies will be entirely derived from the titration period diaries.

Missing data for secondary efficacy endpoints (investigator CGI-S, investigator CGI-I and caregiver CGI-I) will be handled by using the Mixed Model for Repeated Measures method, which is implemented via SAS® PROC MIXED procedure. MMRM assumes data are "Missing At Random (MAR)". That is, the probability that an observation is missing may depend on the observed values but not the missing values in the previously scheduled visits.

6.4.1 *Adverse Event or Concomitant Medications Dates or Times*

For AEs with missing or partially missing start date/time, the following imputation rules will be applied for the determination of treatment-emergent status:

For partial start date/time:

- If the year is unknown, then the date will be assigned the date and time of first dose of study treatment.
- If the year is known to be different from the year of the first dose, then missing month and day will be imputed as the first month and first day of the month.
- If the year is known to be the year of the first dose,
    a) If the month is unknown or is the same as the month of the first dose, then the missing month and day will be imputed by the month and day of the first dose.
    b) If the month is known to be different from the month of the first dose, then the missing day is imputed as 01 (first day of the month).
- If the time is unknown, then:

a) If the date (day, month, and year) matches the date of the administration of study drug, then the time of the study treatment will be used to impute the missing time.

b) Otherwise, '00:00' will be assigned.

For the determination of the prior and concomitant status, the follow rules will be followed for incomplete dates.

- If the medication stop date is partially missing,
  - If the year and month indicate the stop date is before study drug administration, it is Prior medication.
  - Otherwise, it is concomitant medication.
- If the medication stop date is completely missing, it is concomitant medication.

6.5 *Interim Analysis and Adaptive Design*

There is no planned interim analysis.

6.6 *Timing of Final Analyses*

Data summary and statistical analysis will commence after all subjects have completed study Visit 7 (or discontinued prior to Visit 7) and the study database is cleaned and locked.

6.7 *Multicenter Study*

This is a multicenter study. The primary analysis will have all sites pooled. Where applicable, analysis adjusting for site effect or analysis evaluating homogeneity across sites may be carried out. If conducted, small sites (defined as those with sample size <9 ITT subjects/site) will be pooled with larger sites based on the geographic proximity.

6.8 *Multiple Comparisons/Multiplicity Adjustment*

Because of the early termination of the 18 mg arm based on the external data, there will be no formal hypothesis test and statistical comparison between the 18 mg arm and the placebo arm for any of the efficacy endpoints. Thus, there is no multiplicity adjustment due to the multiple dose arms to placebo comparisons with respect to the primary and key secondary endpoint analysis. Multiplicity adjustment for the multiple endpoints with respect to additional secondary end points will be carried out in a fixed sequential gate-keeping approach. The test of the null hypothesis of 36 mg vs placebo for additional secondary endpoints from one endpoint to the next endpoint will be carried out in such a way that the test of the subsequent endpoint will be conducted if and only if the test rejects the null hypothesis for the previous additional secondary endpoint. If any test fails to reject the null hypothesis with respect to a certain additional

secondary endpoint then there will be no test of hypothesis with respect to all subsequent additional secondary endpoints and these tests will all be considered failure.

6.9 *Examination of Subgroups*

In addition to the analysis adjusted for the confounding effects by the covariates (see Section 6.3), primary and key secondary efficacy variables will be summarized by subgroup to explore the heterogeneity of the treatment effect across subgroups. Subgroups are defined as follows:

- Gender (Male, Female)
- Age (6-9, 10-12)
- Race (White, Non-white)

7 STATISTICAL ANALYSIS

7.1 *Analysis Populations*

The population of "all enrolled subjects" consists of all those screened subjects who meet the requirements for study participation and are entered in the Baseline period (i.e., distributed IA diary) of the study. The population of "all randomized subjects" consists of all those enrolled subjects who complete the Baseline Period, meet the inclusion/exclusion criteria and are randomized.

7.1.1 *Safety Population*

The safety population will include all randomized subjects who received at least 1 dose of study drug and have at least one post-randomization safety measurements.

7.1.2 *Modified Intent-to-treat (mITT) Population*

The modified intent-to-treat population will include all subjects who received at least 1 dose of study drug and have a baseline IA frequency ≥ 6 per week and at least 1 valid post-randomization assessment of frequency of IA behaviors based on IA diary entry. Subjects will be analyzed according to the treatment they were randomized to.

7.1.3 *Intent-to-treat (ITT) Population*

The intent-to-treat population will include all subjects who received at least 1 dose of study drug and at least 1 valid post-randomization assessment of frequency of IA behaviors based on IA diary entry. Subjects will be analyzed according to the treatment they were randomized to.

7.1.4 *Per Protocol (PP) Population*

The per protocol population will include all of the subjects in the ITT population who completed the treatment period (Completed Maintenance Visit 6) with 80% diary completion compliance and who did not have major protocol deviations. Diary compliance is calculated as

$$\text{diary compliance} = \frac{\text{number of days diary completed within the treatment period}}{\text{Number of treatment days}} \times 100\%$$

Detailed specification of the PP population will be provided prior to the database lock for the final analysis.

7.2 *Analysis for the Discontinued Treatment Arm*

As stated in Section 3.5, the enrollment for the 18 mg arm was discontinued after the interim analysis for Study 810P301, a study separate from this one. As a result, there will be no formal hypothesis tests and statistical comparison between the 18 mg arm and placebo for any efficacy endpoints. However, for completeness in the presentation, 18 mg arm will remain in the table with summary statistics presented.

7.3 *Disposition of Subjects*

Subject disposition and reasons for screening failure will be summarized for screen-failed subjects (Table 14.1.2.2).

Subject randomization schedule will be listed. Subjects whose randomization codes are unblinded during the conduct of the study will be listed along with the reason for un-blinding (Listing 16.1.7).

The following summaries and listings will be based on all randomized subjects.

Subjects (n and %) who completed and discontinued from the study will be tabulated by the treatment group and study period (titration, maintenance, titration and maintenance combined, and all period combined) for randomized population (Table 14.1.1). The primary reasons for study discontinuation will be included in the tabulation. The primary reasons may include any of the following:

- Subject withdrew consent
- Lost to follow-up
- Administrative reason
- Adverse event
- Investigator decision

- Failure to follow required study procedures
- Other

Also, subject's enrollment will be summarized by study site among ITT population (Table 14.1.2.3).

A data listing will be presented for subject's disposition and reasons for discontinuation (Listing 16.2.1.1).

7.4 *Protocol Deviations*

Protocol deviations will include, but are not limited to

- Non-compliance with any scheduled study visit
- Non-compliance with study treatment
- Prohibited medications
- Non-compliance with study inclusion or exclusion criteria
- Non-compliance with study assessment procedures

Subjects with major protocol deviations will be listed based on all randomized subjects. The listing will include a brief description of the deviation, deviation category, and if applicable, study day when deviation occurred along with other pertinent information (Listing 16.2.1.2).

If warranted, subjects (n and %) with major protocol deviations will be tabulated by treatment group and deviation category for ITT population (Table 14.1.2.4).

Subject's inclusion/exclusion from each analysis population will be listed along with the applicable reasons for exclusion (Listing 16.2.1.3).

7.5 *Demographics and Baseline Subject Characteristics*

Subject age, sex, ethnicity, race, height at screening, weight at screening and baseline will be summarized based on the ITT population by treatment group. Age will be summarized as a continuous variable as well as a categorical variable (6 to 9 [i.e., <10], and 10 to 12 [i.e., ≥10 and ≤ 12] years, inclusive) (Table 14.1.2.1).

All demographic and baseline characteristics will be included in the subject data listing (Listing 16.2.1.4).

7.6 *Baseline Disease Characteristics*

Summaries of baseline disease characteristics will be based on the ITT population, except for medical history, which will be based on the Safety population.

Subject incidence (n and %) of medical history will be tabulated by MedDRA System Organ Class (SOC) and Preferred Term (PT) for each treatment group as per Section 7.9.1 (Table 14.1.3.1).

R-MOAS (item, and domain scores [Table 14.1.5.1.1and Table 14.1.5.1.2]) and Vitiello Aggression Scale (item, and domain scores [Table 14.1.5.2.1 and Table 14.1.5.2.2]) will be summarized by treatment group as per Section 6.1.

All baseline disease characteristics will be listed (Listing 16.2.1.5 – Listing 16.2.1.7).

7.7 *Treatments*

7.7.1 *Treatment Compliance and Extent of Exposure*

Summaries of treatment compliance and exposure will be based on the ITT population.

Duration of exposure is defined as the total number of days a subject is exposed to any study treatment. This will be calculated for each subject by taking the difference between the date of last dose minus the date of the first dose, plus 1 (date of last dose – date of first dose +1).

Duration of Treatment exposure will be summarized using descriptive statistics (n, mean, SD, median, minimum, and maximum).

Percent of SM compliance is defined as

$$\text{Treatment Compliance} = \frac{\text{Number of Tablets Dispensed} - \text{Number of Tablets Returned}}{4\text{x}(\text{Date of Last Dose} - \text{Date of First Dose} + 1)} \cdot 100\%.$$

Each subject is expected to take 4 tablets per day as in the following. Subjects randomized to placebo group will take 2 placebo tablets in the morning and 2 placebo tablets in the evening. Subjects randomized to SPN-810 18 mg will take one 9 mg active tablet and 1 placebo tablet in the morning and one 9 mg active tablet and 1 placebo tablet in the evening. Subjects randomized to SPN-810 36 mg will take two 9 mg active tablets in the morning and two 9 mg active tablets in the evening. For each treatment, SM compliance will be summarized with number of subjects (n and %) by compliance category (<80%, 80-120%, and >120%). SM compliance will also be summarized as a continuous variable using descriptive statistics (n, mean, standard deviation, median, minimum, and maximum) for each treatment.

Summaries of treatment compliance and exposure will be presented separately for the Titration Period, Maintenance Period, and combined Titration and Maintenance Periods (Table 14.1.4 and Listing 16.2.1.8).

7.7.2 *Prior and Concomitant Medications*

Prior and concomitant medications will be tabulated separately based on the Safety Population. The number (n and %) of subjects taking at least one medication and the number (%) of subjects taking each medication at the WHODD preferred term level will be tabulated by treatment group, Anatomical-Therapeutic-Chemical classification 4 (ATC4), and PT. Subjects taking the same PT medication twice will only be counted once (Table 14.1.3.2 and Table 14.1.3.3). In addition, the number and percent for subjects with ≥5% allowed concomitant medications by ATC code and preferred term will be reported (Table 14.1.3.4).

A subject data listing will be provided to include the reported medication name, the WHODD PT, ATC4, study day and pertinent subject information (Listing 16.2.1.9.1 and Listing 16.2.1.9.2).

A separate data listing will include subjects who received prohibited medications (Listing 16.2.1.9.3).

7.8 *Efficacy Analyses*

7.8.1 *Primary Efficacy Analysis*

The primary analysis will be based on the mITT population and the primary efficacy endpoint is percent change (PCH_T) in the frequency (unweighted score) of IA behaviors per 7 days in the treatment period relative to the Baseline period calculated over the number of days with non-missing IA diary. The analysis will be repeated for the ITT population, which follows the original primary efficacy analysis population definition.

Both of the above analyses will be performed using the Wilcoxon rank-sum test to compare the median SPN-810 36 mg with the median of the Placebo. The Hodges-Lehmann estimate of the difference (SPN-810 36 mg dose minus placebo) and the associated 95% confidence interval (CI) around the difference will be calculated. The statistical analysis will be presented in the table along with the summary of the change from baseline frequency (Table 14.2.1.1). The daily IA behavior diary will be listed. The daily frequency, the derived 7-day frequency for each week, and the derived 7-day frequency for the treatment period will be included (Listing 16.2.2.1.1 – Listing 16.2.2.1.2). Histogram of median percent change from baseline in the frequency of IA Behaviors in treatment will be created for 36 mg dose and Placebo (Figure 14.2.1.1).

7.8.2 *Sensitivity Analyses*

Sensitivity analyses will be performed both on the mITT and ITT populations.

7.8.2.1 Multiple Imputation Assuming Missing at Random

This sensitivity analysis assumes that the diary data are missing at random (MAR), that is, the probability that an observation is missing may depend on the observed values but not the missing values (e.g., the probability of missing Week 5 diary is related to the observed Week 4 or Week 3 value but not the unobserved Week 5 value). In this sensitivity analysis, missing data is imputed multiple times, thus creating the multiply-imputed diary data. The final analysis incorporates the between-imputation variability with the within-imputation variability.

The MI is implemented using the following three steps.

1. SAS PROC MI is applied with input dataset containing some missing values for all days during the titration and maintenance period to create 100 datasets. The data sets will include separate columns for the frequency of incidences during each day starting from baseline. The Markov Chain Monte Carlo (MCMC) method will be used to complete the missingness pattern to a monotone pattern separately by treatment arm. The monotone patterns will be achieved by applying sequential imputation based on Bayesian regression with the treatment arm included as a covariate. All copies contain identical values of the non-missing data items, but different values imputed for missing values.

2. For each of these MI data sets, the percent change will be computed as in the observed data set and the primary analysis based on the Wilcoxon rank-sum test will be conducted and asymptotic 95% confidence intervals will be constructed.

3. To produce a single confidence interval for each dose vs placebo comparison (e.g., Dose 1 versus placebo), PROC MIANALYZE will be used and Rubin's combination rules will be applied to the treatment effect estimates and associated asymptotic standard errors from the MI data sets (Rubin 1987). The treatment effect estimates will be defined as the midpoints of the asymptotic confidence intervals and the standard errors will be defined as the asymptotic standard errors (based on the width of the associated 95% confidence intervals) from the Hodges-Lehmann estimate of the individual datasets.

An example of SAS code is as follows.

```
PROC MI DATA=adqs NIMPUTE=100 SEED=123456789 OUT=adqs_imp ;
    BY trtp ;
    VAR age sex bwk  twk1 twk2 twk3  twk4  twk5 ;
    MCMC CHAIN=multiple IMPUTE=full INITIAL=em PRIOR=jeffreys ;
```

```
RUN ;
```

where trtp is treatment group, bwk, twk1, twk2, twk3, twk4, and twk5 are variables of weekly IA scores during the baseline, and treatment (titration + maintenance) periods. These are derived as (SUM/DAY)*7, where SUM is the total of the IA scores reported on the IA diary in that analysis week, and DAY is the number of days with non-missing IA score in the IA diary in that analysis week. Analysis week is defined as: 1st week is the 1st 7 days after visit 3 (exclusive) during titration, 2nd week is the rest of the days up to Visit 5 (inclusive) in titration period, 3rd week is the 1st 7 days after Visit 5 (exclusive), 4th week is the following 7 days after 3rd week, the 5th week is the rest of the days in maintenance period up to Visit 6 (inclusive). Adqs is the input data set with missing values, and adqs_imp is the complete imputed data set after 100 imputations. For the weekly IA score calculation, if the week has <3 daily diaries, the weekly IA score for this week will be set to missing and imputed using the above procedure.

From the above multiply-imputed IA data, the Hodges-Lehmann confidence interval will be calculated for each imputed data set. The point estimate for the between group difference and the associated standard error (sem) will be derived from the confidence interval. These point estimates and the sem's will be summarized by the SAS MIANALYZE procedure. An example SAS code is as follows.

```
PROC MIANALYZE DATA=parmest ;
    MODELEFFECTS estim ;
    STDERR sem ;
    ODS OUTPUT PARAMETERESTIMATES=parmest2(KEEP= paramcd label estimate
    stderr lclmean uclmean probt) ;
RUN ;
```

where estim and sem are the point estimate and standard error from each imputed set, estimate, stderr, lclmean, uclmean, and probt are the final summarized point estimate, standard error, lower and upper confidence limits, and the p-value for the between group comparison (Table 14.2.1.2.1).

7.8.2.2 Multiple Imputation Assuming Missing Not at Random

This sensitivity analysis assumes that the diary data are missing not at random (MNAR), that is, the probability that an observation is missing may depend on its underlying unobserved value.

This approach may be considered "worst-case" sensitivity analyses as it assumes that after discontinuation subjects from the dosing arms would adopt the outcome model estimated from the placebo arm. To generate missing values from this "placebo-based" imputation model, PROC MI with the MNAR statement will be used (Table 14.2.1.2.2). An example of SAS code is as follows.

```
PROC MI DATA=adqs NIMPUTE=100 SEED=123456789 OUT=adqs_imp ;
    CLASS trtp ;
    VAR age sex bwk twk1 twk2 twk3 twk4 twk5;
    FCS REG ;
    MNAR MODEL(bwk twk1 twk2 twk3 twk4 twk5 / MODELOBS=(trtp='PLACEBO')) ;
RUN ;
```

where multivariate imputation is carried out by the fully conditional specification (FCS) method.

Histogram of median percent change from baseline in the frequency of IA Behaviors by imputation status in treatment and maintenance will be created for 36 mg dose and Placebo (Figure 14.2.1.2.1 – Figure 14.2.1.2.2).

7.8.2.3 Analysis of PCH Based on Data from Maintenance Period

For this sensitivity analysis, the primary analysis will be repeated using data restricted to the maintenance period only (Table 14.2.1.2.3).

7.8.3 *Supplementary Analysis*

A supplementary analysis will be performed based on the Per Protocol population and will be presented in Table 14.2.1.2.4.

7.8.4 *Analysis Based on Adjustment of Covariates*

An exploratory analysis of the primary efficacy endpoint will be performed using a randomization-based approach (Dmitrienko and Koch 2017). Separate models will be produced for the following key prognostic baseline characteristics (age [categorical and continuous] and gender [categorical]). In addition, pooled site and baseline IA adjusted to 7 days will be used as covariates (Table 14.2.1.3).

This analysis will be done by calling the following macro:

/* Linear Model with Adjustment for Covariates for Treatment period*/

%NPARCOV4(OUTCOMES = PCH_T, COVARS = gender age blsum, C=1, STRATA = center, TRTGRPS = trtp, HYPOTH = NULL, TRANSFORM = NONE, COMBINE = FIRST, DSNIN = XXXX, DSNOUT = XXXX, ALPHA = 0.05, SEED = XXXX, NREPS = 500, EXACT = NO, DETAILS = YES);

The above macro call will provide estimates of treatment effect and p-value. To obtain the 95% confidence interval, the macro will be called again specifying HYPOTH = ALT.

7.8.5 *Examination of Subgroups*

The primary efficacy variables will be summarized by subgroup to explore the heterogeneity of the treatment effect across subgroups. This analysis will be done for the following subgroups: gender (male, female), age (6-9 and 10-12), and race (White, Non-white) and baseline body weight (approximate median baseline weight: <median vs >=median). (Table 14.2.1.4.1 - Table 14.2.1.4.4).

7.8.6 *Analysis of Secondary Efficacy Endpoints*

The analyses of all secondary endpoints except CHQ-PF28, PSI-4-SF and SNAP-IV ADHD will be based on both the mITT and ITT population but the analysis of CHQ-PF28, PSI-4-SF and SNAP-IV ADHD will be based on the ITT population only.

7.8.6.1 Analysis of Key Secondary Efficacy Endpoint

The key secondary efficacy endpoint is the change from baseline (Visit 3) to Visit 6 in the Investigator CGI-S score.

This endpoint will be analyzed using Mixed-Effect Model for Repeated Measure (MMRM) for the mITT and ITT populations. The model includes treatment, visit, and interaction between treatment and visit as fixed factors, and baseline as covariate. The model parameters will be estimated using restricted maximum likelihood method with unstructured variance-covariance matrix and Kenward-Roger approximation to estimate denominator degrees of freedom. In case there is a convergence problem in the MMRM model with the unstructured variance covariance matrix, the first (co)variance structure which does not have convergence problem will be used for the analysis from the following ordered list: 1) Toeplitz [TOEP], 2) Autoregressive of Order 1 [AR(1)], and 3) Compound Symmetry [CS].

The least squares means of SPN-810 36 mg dose and placebo, the difference in the least squares mean (SPN-810 36 mg dose minus placebo), and the 2-sided 95% CI for the difference will be calculated (Table 14.2.2.1, Listing 16.2.2.2).

An example of SAS code is provided as follows.

```
ODS output DIFFS=DIF LSMEANS=LSMN;
PROC MIXED DATA=XXXX ;
      CLASS trtp (ref='Placebo') usubjid visit;
      MODEL Response = base trtp visit trtp*visit / DDFM=kr ;
      REPEATED visit / SUBJECT=usubjid TYPE=UN ;
      LSMEANS trtp*visit/PDIFF ALPHA=0.05;
RUN ;
```

Note: usubjid is subject ID, trtp is treatment variable, visit is the nominal study visit, and Response is respective endpoints.

In addition, the mean change from baseline by visit will be presented in Figure 14.2.4.1.

7.8.6.2 Analyses of Additional Secondary Efficacy Points

The following summaries and listings will be presented.

1. Investigator CGI-I score (continuous and binary) at Visits 4, 5, and 6 (Table 14.2.2.2.1 – Table 2.2.2.2, Figure 14.2.3.1, and Listing 16.2.2.3)
2. PSI-4-SF score at visit 6 in total score and domains will be summarized in Table 14.2.2.4 (Listing 16.2.2.5.1 – Listing 16.2.2.5.2).
    a. Total Score
    b. Parental Distress
    c. Parent-Child Dysfunctional Interaction
    d. Difficult Child
3. Caregiver CGI-I score (continuous and binary) by Visit (Table 14.2.2.5.1 – Table 14.2.2.5.2, Figure 14.2.2.1, and Listing 16.2.2.3)
4. SNAP-IV ADHD scores at visit 6 will be summarized in Table 14.2.2.6 (Listing 16.2.2.6.1 - Listing 16.2.2.6.2):
    a. Inattention ratings
    b. Hyperactivity/Impulsivity ratings
    c. Oppositional Defiant Disorder
    d. Combined Scale ratings
5. Percentage of responders with 50% or more reduction in the frequency of IA from baseline to treatment period (Table 14.2.2.7)
6. Percentage of responders with 30% or more reduction in the frequency of IA from baseline to treatment period (Table 14.2.2.8)
1.

Scores of CGI-I (investigator and caregiver) will be analyzed using a Mixed-Effect Model for Repeated Measure (MMRM) similar to the key secondary outcome. CGI-S will be used as baseline covariate.

Scores for CHQ-PF28, PSI-4-SF, and SNAP-IV will be analyzed using analysis of covariance (ANCOVA) on the ITT population. The model includes treatment and baseline as fixed independent covariates and Visit 6 value as a response variable. The least squares mean of each treatment group, the difference in the least squares mean (SPN-810 36 mg dose minus placebo), and the 2-sided 95% CI for the difference will be obtained.

The SAS code for the ANCOVA will be:

```
PROC MIXED DATA=XXXX ;
    CLASS trtp usubjid;
    MODEL Response = base trtp / DDFM=kr ;
    LSMEANS trtp t/PDIFF=control ('PLACEBO') ALPHA=0.05;
RUN ;
```

Where usubjid is subject ID, trtp is treatment variable and Response is CHQ-PF28, PSI-4-SF, and SNAP-IV at Visit 6.

The Percentage of responders defined as 50% or more reduction (and 30% or more reduction) in the frequency of IA from baseline to treatment period will be analyzed using the logistic regression model with treatment as explanatory variables. Odds ratio (36 mg Dose / Placebo), and 95% CI for the odds ratio and p-value will be presented. In addition, the number and percentage of responders will be tabulated.

An example of SAS code follows.

```
ODS OUTPUT OddsRatios=or /* ODDS RATIO AND 95% CL*/
        LSMeans=lsm(keep=trtpn mu) /* PROB OF RESPONDERS */ ;
PROC LOGISTIC DATA=ef1 PLOTS=none ;
    CLASS siteid trtpn / PARAM=glm ;
    MODEL avalc(EVENT="Yes")=trtpn;
    LSMEANS trtpn / PDIFF=controll('2') ILINK ;
RUN;
```

To control family-wise error rate at 0.05 for multiple endpoints, the key secondary endpoint and additional secondary endpoints will be ordered as follows: CGI-S, investigator CGI-I, CHQ-PF28, PSI-4-SF, caregiver CGI-I, SNAP-IV ADHD, Percentage of responders with 50% or more reduction in the frequency of IA, percentage of responders with 30% or more reduction in the frequency of IA. The testing will proceed in the following order.

1. Test the null hypothesis H_0: There is no difference between the 36 mg dose SPN-810 and placebo will be performed with respect to CGI-S. If not rejected then the hypothesis will not be tested with respect to any additional secondary endpoint.
2. If H_0 is rejected above, then proceed to test the same null hypothesis with respect to the 1st additional secondary endpoint (CGI-I). If not rejected then the hypothesis will not be tested with respect to any additional secondary endpoint after CGI-I.
3. If H_0 is rejected above for CGI-I, then test of H_0 for the 2nd additional secondary endpoint (CHQ-PF28) will follow. If not rejected then the hypothesis testing will stop with respect to any subsequent additional secondary endpoint after CHQ-PF28. Otherwise, proceed

to test the same null hypothesis with respect to the 3rd additional secondary endpoint (PSI-4-SF).

The process will be repeated until all the secondary endpoints are gone through in the pre-defined order above in the same manner.

Subgroup (age, gender and race) analyses will be performed for the key secondary endpoint to explore the heterogeneity of the treatment effect across subgroups (Table 14.2.3.1.1-Table 14.2.3.1.3).

Supportive analyses will also be performed by summarizing the descriptive statistics for primary and secondary endpoints (Table 14.2.4.1 – Table 2.4.7).

7.9 *Safety Analyses*

Subject safety is assessed by AEs, labs, vital signs, 12-lead ECGs, and findings from the physical examinations. The occurrence of neurological side effects will be assessed by looking at any worsening in scores for each of the Simpson-Angus scale, Barnes Akathisia scale, and AIMS. Suicidal ideation and suicidal behavior will be measured by C-SSRS. All summaries related to safety analyses will use the safety population and will be presented by treatment groups.

7.9.1 *Adverse Events*

AEs will be summarized using subject incidence table by treatment period (titration, maintenance, combined titration and maintenance, and tapering). An overview of TEAE will be presented. This table will include numbers and percentages of subjects experiencing at least 1 TEAE, study drug-related TEAE, AE leading to study drug discontinuation, severe TEAE, serious TEAE, related TESAE, AESI, and TEAE with outcome of death (Table 14.3.1.1).

Additionally, the number and percent (%) are calculated based on the number of unique subjects within each category (e.g., preferred term) by treatment group. A subject reporting multiple events of the same category will be counted only once for that category. For summary purpose, AE relationship to the study drug will be grouped into “Unrelated” for “unrelated” or “unlikely related” and “Related” for “possibly”, “probably”, or “definitely related”. For subjects with more than one event coded to the same PT, the subjects will be counted for the categories with the strongest relationship and the greatest severity. The following subject incidence tables will be presented.

- TEAEs by PT (Preferred Term) sorted by the decreasing order of subject incidence in the combined active dose group (Table 14.3.1.6)
- TEAEs by SOC (System Organ Class) and PT sorted alphabetically (Table 14.3.1.2)

- Study Drug-Related TEAEs by SOC and PT (Table 14.3.3)
- Severity of TEAEs by SOC and PT (Table 14.3.1.4)
- TESAEs by SOC and PT when applicable (Table 14.3.1.5)
- Number and percent of patients reporting common AEs (≥5% in any group) by PT (Table 14.3.1.7)
- TEAEs leading to study drug discontinuation by SOC and PT (Table 14.3.1.8)
- Treatment Emergent AESI by SOC and PT (Table 14.3.1.9)

Where warranted by the subject incidence, these same tables may be presented by treatment period (Titration, Maintenance and Conversion/Tapering).

A subject data listing will be provided for all adverse events (Listing 16.2.3.1.1). Included in the listing are the reported term, PT, SOC, TEAE flag, study day when AE starts, duration, relationship, severity, action taken, outcome, and seriousness category.

Data listings will be provided for subjects who die on study, experience SAEs, have TEAEs leading to study drug discontinuation, or AEs of special interest (Listing 16.2.3.1.2 – Listing 16.2.3.1.5).

7.9.2 *Laboratory Test*

Clinical laboratory values will be summarized by treatment group and visit using descriptive statistics for hematology, biochemistry, and urinalysis. For quantitative laboratory parameters, both actual values and change from baseline values will be summarized (Table 14.3.2.1.1, Table 14.3.2.2.1, and Table 14.3.2.3.1).

Laboratory test result will be assigned a low, normal, high (LNH) classification according to whether the value is below (L), within (N), or above (H) the reference range provided by the laboratory. Within-treatment comparisons will be based on three by three tables (shift tables), that, for a particular laboratory test, compare the LNH classification at baseline to the LNH classification at visit. By subject-listings of all laboratory values will be provided for subjects who have L or H classification in any of their laboratory tests (Table 14.3.2.1.2, Table 14.3.2.2.2, and Table 14.3.2.3.2).

Note the summary and shift tables will only use the planned/scheduled tests. However, both scheduled and repeat tests will be in the data listing.

A complete lab data listing, including hematology, biochemistry, and urinalysis will be provided for all subjects. Investigator assessment of clinical significance is based on the local laboratory reference ranges (Listing 16.2.3.2.1 – Listing 16.2.3.2.3).

7.9.3 *Vital Signs, Height, Weight, and BMI*

Vital signs (body temperature, heart rate, respiration rate, systolic blood pressure, and diastolic blood pressure), and their changes from baseline will be summarized by treatment group and visit (Table 14.3.3.1, Listing 16.2.3.3.1 - Listing 16.2.3.3.2). Subject body size (weight, height, and BMI) will be summarized similarly (Table 14.3.3.2, Listing 16.2.3.3.3).

7.9.4 *12-Lead ECG*

By-visit tabular summaries of the quantitative ECG parameters and the overall ECG findings (normal, abnormal not clinically significant, or abnormal clinically significant) will be presented. The QT interval will be corrected using Fridericia's method.

ECG results will be summarized by visit by treatment group using descriptive statistics (for quantitative ECG parameters) and frequency tables (for qualitative ECG parameters, including the overall ECG finding). For quantitative ECG parameters, both actual values and change from screening values will be summarized (Table 14.3.4.1 and Table 14.3.4.2).

Data listings will be presented for ECG parameters by visit. QTcF meeting the criteria of >450 but ≤480, >480 but ≤500 or >500 msec for the absolute value, or >30 msec or >60 msec increase from baseline will be identified (Listing 16.2.3.4.1 and Listing 16.2.3.4.3).

7.9.5 *Columbia Suicide Severity Rating Scale (C-SSRS)*

C-SSRS outcomes will be summarized using number and percent of subjects by categories for suicidal ideation only, suicidal behavior only and suicidality (ideation and behavior combined). The summary will be presented by treatment groups. (Table 14.3.5 and Listing 16.2.3.5).

7.9.6 *Infrequent Behaviors Checklist*

Infrequent behaviors will be listed for each subject by treatment group (Table 14.3.7 and Listing 16.2.3.7).

7.9.7 *Other Special Tests*

The occurrence of neurological side effects will be assessed by looking at the changes in scores from baseline to post-baseline visits for each of the Simpson-Angus scale (Table 14.3.6.1.1 - Table 14.3.6.1.2 and Listing 16.2.3.6.1), Barnes Akathisia scale (Table 14.3.6.2.1 - Table 14.3.6.2.2 and Listing 16.2.3.6.2), and AIMS (Table 14.3.6.3.1 - Table 14.3.6.3.2 and Listing 16.2.3.6.3). For each item on each of these scales, the number (and percentage) of subjects with a worse score at any post-baseline visit, compared to baseline, will be presented. A listing of these subjects will also be provided.

8 VALIDATION

The Clinical Operations, Clinical Research, and Statistics groups at Supernus will work with [redacted] and [redacted] to ensure that the data collected for the study are of the highest quality possible. The study monitor will be responsible for reviewing and verifying the accuracy of the data recorded on the electronic case report forms (eCRFs) direct from source documents at the investigative site. DSG will be responsible for performing edit checks and reviewing all data entered into the electronic database to identify discrepant and/or inconsistent values and to send queries to the clinical sites. The Investigator will be responsible for answering queries about discrepant data and providing electronic signatures to confirm data integrity.

The programming of Tables, Listings and Figures (TLFs) based on the clinical data is outsourced to [redacted] Supernus seeks to ensure the quality of the reports provided by [redacted] in the form of TLFs must pass a rigorous validation process involving the following processes.

- Derived datasets must be independently reprogrammed by a second programmer. The separate datasets produced by the 2 programmers must match 100%.
- Tables must be independently reprogrammed by a second programmer for and the results from both programs must match.
- Figures must be checked for consistency against corresponding tables and listings, or independently reprogrammed if there are no corresponding tables or listings.
- Listings must be checked for consistency against corresponding tables, figures, and derived datasets.

The entire set of TLF must be checked for completeness and consistency prior to its delivery to Supernus.

The above validation process must be repeated any time TLFs are redelivered using different data. Execution of this validation process must be documented through the study Table of Programs that will be provided to Supernus at study conclusion.

9 REFERENCES

Brannath W, M Posch, P Bauer. Recursive combination tests. Journal of the American Statistical Association. 2002; 97: 236-244.

Chang, M. Adaptive Design Theory and Implementation Using SAS and R. Chapman & Hall/CRC, Boca Raton, Florida. 2008a.

Change, M. Classical and Adaptive Clinical Trial Designs Using ExpDesign Studio. Wiley, New Jersey. 2008b.

Chen, YH, DL DeMets, and KKG Lan. "Increasing the sample size when the unblinded interim result is promising." Statistics in Medicine, 2004. 23: 1023-1038.

HealthActCHQ. The CHQ Scoring and Interpretation Manual. Boston, MA: HealthActCHQ, 2013. 2013.

Dmitrienko, A, RB D'Agostino. Tutorial in Biostatistics: Traditional multiplicity adjustment methods in clinical trials. Statistics in Medicine. 2013; 32:5172-5218.

Dmitrienko A, AC Tamhane, and F Bretz. Multiple Testing Problems in Pharmaceutical Statistics. Chapman & Hall/CRC, Boca Raton, Florida. 2010.

Dmitrienko, A and GG Koch. Analysis of Clinical Trials Using SAS: A Practical Guide. 2nd Ed. Cary, NC: SAS Institute Inc. 2017.

Hintze, J. PASS 12, NCSS, LLC. Kaysville, Utah, USA. 2013.

Hommel, G. Adaptive modifications of hypotheses after an interim analysis. Biometrical Journal. 2005; 5: 581-589.

Jennison C and Turnbull BW. Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC, Boca Raton, Florida. 2000.

Mehta CR and SJ Pocock. "Adaptive increase in sample size when interim results are promising: a practical guide with examples." Statistics in Medicine, 2000. 00:1-6.

Wang, SJ, HMJ Hung, Y Tsong, L Cui. Group sequential test strategies for superiority and non-inferiority hypotheses in active controlled clinical trials. Statistics in Medicine. 2001; 20: 1903-1912.

Appendix 1. Exclusion Criteria for the Per Protocol Populations

Appendix 2. List of Tables, Listings, and Figures

Tables
Table 14.1.1 Summary of Subject Disposition and Reason for Discontinuation by Treatment Group and Period (Randomized Population)
Table 14.1.2.1 Demographics and Baseline Characteristics (ITT Population)
Table 14.1.2.2 Subject Demographic Characteristics and Reason for Screening Failure (Screening Failure Subjects)
Table 14.1.2.3 Subject Enrollment Distribution by Study Site (ITT Population)
Table 14.1.2.4 Summary of Major Protocol Deviations During the Study, by Treatment Group (ITT Population)
Table 14.1.3.1 Medical History by System Organ Class and Preferred Term (Safety Population)
Table 14.1.3.2 Prior Medications by ATC Code and Preferred Term (Safety Population)
Table 14.1.3.3 Concomitant Medications by ATC Code and Preferred Term (Safety Population)
Table 14.1.3.4 Subjects with ≥5% Allowed Concomitant Medications by ATC Code and Preferred Term (Safety Population)
Table 14.1.4 Exposure to Study Medication and Compliance (ITT Population)
Table 14.1.5.1.1 Retrospective Modified Overt Aggression Scale at Baseline – Item Scores (ITT Population)
Table 14.1.5.1.2 Retrospective Modified Overt Aggression Scale at Baseline – Domain and Total Scores (ITT Population)
Table 14.1.5.2.1 Vitiello Aggression Scale (VAS) at Baseline – Item Scores (ITT Population)
Table 14.1.5.2.2 Vitiello Aggression Scale (VAS) at Baseline – Domain and Total Scores (ITT Population)
Section 14.2 – Efficacy Tables
Table 14.2.1.1 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Treatment Period [1] - (mITT and ITT Population)
Table 14.2.1.2.1 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Treatment Period [1] - Sensitivity Analysis, Multiple Imputation with MAR Assumption (mITT and ITT Population)
Table 14.2.1.2.2 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Treatment Period [1] – Sensitivity Analysis, Multiple Imputation (mITT and ITT Population)
Table 14.2.1.2.3 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Maintenance Period - Sensitivity Analysis (mITT and ITT Population)
Table 14.2.1.3 Percent Change from Baseline to Treatment Period in Frequency of IA Behaviors: Adjustment for Covariates using Randomization-Based Approach of Dmitrienko and Koch (mITT and ITT Population)
Table 14.2.1.4.1 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Treatment Period by Gender - Subgroup Analysis (mITT and ITT Population)
Table 14.2.1.4.2 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Treatment Period by Age - Subgroup Analysis (mITT and ITT Population)
Table 14.2.1.4.3 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Treatment Period by Race - Subgroup Analysis (mITT and ITT Population)

Table 14.2.1.4.4 Percent Change from Baseline in the Average Weekly Frequency of IA Behaviors During the Treatment Period by Baseline Body Weight - Subgroup Analysis (mITT and ITT Population)
Table 14.2.2.1 Investigator Clinical Global Impression – Severity (CGI-S) Score and Change from Baseline by Visit (mITT and ITT Population)
Table 14.2.2.2.1 Investigator Clinical Global Impression – Improvement (CGI-I) Score by Visit (mITT and ITT Population)
Table 14.2.2.2.2 Summary of Investigator Clinical Global Impression – Improvement (CGI-I) Score by Visit (mITT and ITT Population)
Table 14.2.2.3.1 Child Health Questionnaire Parent Form 28 (CHQ-PF28)- Scale Scores and Change from Baseline to Visit 6 (ITT Population)
Table 14.2.2.3.2 Child Health Questionnaire Parent Form 28 (CHQ-PF28)-Component Scores and Change from Baseline to Visit 6 (ITT Population)
Table 14.2.2.4 Parenting Stress Index - Short Form (PSI-4-SF) Observed Scores and Change from Baseline to Visit 6 (ITT Population)
Table 14.2.2.5.1 Caregiver Clinical Global Impression – Improvement (CGI-I) Score by Visit (mITT and ITT Population)
Table 14.2.2.5.2 Summary of Caregiver Clinical Global Impression – Improvement (CGI-I) Score by Visit (mITT and ITT Population)
Table 14.2.2.6 SNAP-IV ADHD Inattention Subscale Score and Actual Value at Visit 6 (ITT Population)
Table 14.2.2.7 Percentage of responders with 50% or more reduction in the frequency of IA from baseline by the Treatment Period (mITT and ITT Population)
Table 14.2.2.8 Percentage of responders with 30% or more reduction in the frequency of IA from baseline by the Treatment Period (mITT and ITT Population)
Table 14.2.3.1.1 Subgroup Analysis of Investigator Clinical Global Impression – Severity (CGI-S) Score by Age Group and Visit (mITT and ITT Population)
Table 14.2.3.1.2 Subgroup Analysis of Investigator Clinical Global Impression – Severity (CGI-S) Score by Gender and Visit (mITT and ITT Population)
Table 14.2.3.1.3 Subgroup Analysis of Investigator Clinical Global Impression – Severity (CGI-S) Score by Race and Visit (mITT and ITT Population)
Table 14.2.4.1 Descriptive Statistics for Frequency of IA Behaviors Per 7 Days and Change from Baseline by Study Period (mITT and ITT Population)
Table 14.2.4.2 Descriptive Statistics of Observed Investigator Clinical Global Impression – Severity (CGI-S) by Visit - Supportive Analysis (mITT and ITT Population)
Table 14.2.4.3 Descriptive Statistics of Observed Investigator Clinical Global Impression (CGI-I) by Visit - Supportive Analysis (mITT and ITT Population)
Table 14.2.4.4 Descriptive Statistics of Child Health Questionnaire Parent Form 28 (CHQ-PF28) by Visit - Supportive Analysis (ITT Population)
Table 14.2.4.5 Descriptive Statistics of Parenting Stress Index - Short Form (PSI-4-SF) by Visit - Supportive Analysis (ITT Population)
Table 14.2.4.6 Descriptive Statistics of Observed Caregiver Clinical Global Impression – Improvement (CGI-I) by Visit - Supportive Analysis (ITT Population)
Table 14.2.4.7 Descriptive Statistics of SNAP-IV ADHD Observed Scores by Visit - Supportive Analysis (ITT Population)

Table 14.3.1.1 Overall Incidence of Treatment-Emergent Adverse Events by Study Period (Safety Population)
Table 14.3.1.2 Subject Incidence of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Treatment Period (Safety Population)
Table 14.3.1.3 Subject Incidence of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, Treatment-Relatedness, and Treatment Period (Safety Population)
Table 14.3.1.4 Subject Incidence of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, Severity, and Treatment Period (Safety Population)
Table 14.3.1.5 Subject Incidence of Treatment-Emergent Serious Adverse Events by System Organ Class, Preferred Term and Treatment Period (Safety Population)
Table 14.3.1.6 Subject Incidence of Treatment-Emergent Adverse Events by Preferred Term and Treatment Period (Safety Population)
Table 14.3.1.7 Incidence of Common Treatment-Emergent Adverse Events Reported for >=5% of Subjects in Any Group by Preferred Term and Treatment Period (Safety Population)
Table 14.3.1.8 Subject Incidence of Treatment-Emergent Adverse Events Leading to Study Drug Discontinuation by System Organ Class, Preferred Term and Treatment Period (Safety Population)
Table 14.3.1.9 Subject Incidence of Treatment-Emergent Adverse Events of Special Interest by System Organ Class, Preferred Term and Treatment Period (Safety Population)
Table 14.3.2.1.1 Hematology: Change from Baseline by Visit (Safety Population)
Table 14.3.2.1.2 Hematology: Shift from Baseline Classification by Visit (Safety Population)
Table 14.3.2.2.1 Blood Chemistry: Change from Baseline by Visit (Safety Population)
Table 14.3.2.2.2 Blood Chemistry: Shift from Baseline Classification by Visit (Safety Population)
Table 14.3.2.3.1 Urinalysis: Change from Baseline by Visit (Safety Population)
Table 14.3.2.3.2 Urinalysis: Shift from Baseline Classification by Visit (Safety Population)
Table 14.3.3.1 Vital Signs: Change from Baseline by Visit (Safety Population)
Table 14.3.3.2 Subject Body Size, Change from Baseline by Visit (Safety Population)
Table 14.3.4.1 12-Lead ECG Intervals: Change from Baseline by Visit (Safety Population)
Table 14.3.4.2 12-Lead ECG Overall Findings: Shift from Baseline Classification by Visit (Safety Population)
Table 14.3.5 Columbia Suicide Severity Rating Scale by Visit (Safety Population)
Table 14.3.6.1.1 Simpson-Angus Scale (SAS) Item Score and Change from Baseline by Visit (Safety Population)
Table 14.3.6.1.2 Simpson-Angus Scale (SAS) Item Score: Shift from Baseline by Visit (Safety Population)
Table 14.3.6.2.1 Barnes Akathisia Rating Scale (BAS) Item Score and Change from Baseline by Visit (Safety Population)
Table 14.3.6.2.2 Barnes Akathisia Rating Scale (BAS): Shift from Baseline by Visit (Safety Population)
Table 14.3.6.3.1 Abnormal Involuntary Movement Scale (AIMS) Item Score and Change from Baseline by Visit (Safety Population)
Table 14.3.6.3.2 Abnormal Involuntary Movement Scale (AIMS): Shift from Baseline by Visit (Safety Population)
Table 14.3.7 Infrequent Behaviors Checklist by Visit (Safety Population)

Figures
Figure 14.2.1.1 Median Percent Change from Baseline in the Frequency of IA Behaviors in Treatment Period by Stage (ITT Population)
Figure 14.2.1.2.1 Median Percent Change from Baseline in the Frequency of IA Behaviors by Imputation Status (ITT Population)
Figure 14.2.1.2.2 Median Percent Change from Baseline in the Frequency of IA Behaviors in the Maintenance Period (ITT Population)
Figure 14.2.2.1 Mean of Caregiver Clinical Global Impression – Improvement (CGI-I) Score by Visit (ITT Population)
Figure 14.2.3.1 Mean of Investigator Clinical Global Impression – Improvement (CGI-I) Score by Visit (ITT Population)
Figure 14.2.4.1 Mean Change from Baseline of Investigator Clinical Global Impression – Severity (CGI-S) Score by Visit (ITT Population)
Figure 14.2.5.1 Median Change from Baseline of Child Health Questionnaire Parent Form 28 (CHQ-PF28)- Physical Functioning Summary Score by Visit (ITT Population)
Figure 14.2.5.2 Median Change from Baseline of Child Health Questionnaire Parent Form 28 (CHQ-PF28)- Psychosocial Health Summary Score by Visit (ITT Population)

Listings
Listing 16.1.7 Subject Randomization Schedule(Randomized Population)
Listing 16.2.1.1 Subjects Disposition and Reason for Disposition (Randomized Population)
Listing 16.2.1.2 Subjects Who had Major Protocol Deviations (Randomized Population)
Listing 16.2.1.3 Subjects Inclusion and the Reason for Exclusion from Analysis Populations (Randomized Population)
Listing 16.2.1.4 Demographic and Baseline Characteristics (Randomized Population)
Listing 16.2.1.5 Medical History (Randomized Population)
Listing 16.2.1.6 R-MOAS Scale at Baseline (Randomized Population)
Listing 16.2.1.7 Viteillo Aggression Scale at Baseline (Randomized Population)
Listing 16.2.1.8 Study Drug Accountability and Compliance (Randomized Population)
Listing 16.2.1.9.1 Prior Medications (Randomized Population)
Listing 16.2.1.9.2 Concomitant Medications (Randomized Population)
Listing 16.2.1.9.3 Prohibited Medications (Randomized Population)
Listing 16.2.2.1.1 Impulsive Aggression Behavior Daily Diary (ITT Population)
Listing 16.2.2.1.2 Derived Primary Endpoint (ITT Population)
Listing 16.2.2.2 Investigator Clinical Global Impression – Severity (CGI-S) (ITT Population)
Listing 16.2.2.3 Investigator and Caregiver Clinical Global Impression – Improvement (CGI-I) (ITT Population)
Listing 16.2.2.4.1 Child Health Questionnaire Parent Form 28 (CHQ-PF28)– Question Text
Listing 16.2.2.4.2 Child Health Questionnaire Parent Form 28 (CHQ-PF28) – Individual Questions (ITT Population)

Listing 16.2.2.4.3 Child Health Questionnaire Parent Form 28 (CHQ-PF28) – Items, Components and General Health Scores (ITT Population)
Listing 16.2.2.5.1 Parenting Stress Index – Short Form (PSI-4-SF) – Individual Questions (ITT Population)
Listing 16.2.2.5.2 Parenting Stress Index – Short Form (PSI-4-SF) – Domain and Total Scores (ITT Population)
Listing 16.2.2.6.1 SNAP-IV ADHD Teacher and Parent Rating Scale – Individual Questions (ITT Population)
Listing 16.2.2.6.2 SNAP-IV ADHD Teacher and Parent Rating Scale – Subscales (ITT Population)
Listing 16.2.3.1.1 All Adverse Events (Safety Population)
Listing 16.2.3.1.2 Serious Adverse Events (Safety Population)
Listing 16.2.3.1.3 Adverse Events with Outcome of Death (Safety Population)
Listing 16.2.3.1.4 Adverse Events Leading to Study Drug Discontinuation (Safety Population)
Listing 16.2.3.1.5 Adverse Events of Special Interest (Safety Population)
Listing 16.2.3.2.1 Clinical Laboratory Tests-Hematology (Safety Population)
Listing 16.2.3.2.2 Clinical Laboratory Tests - Blood Chemistry (Safety Population)
Listing 16.2.3.2.3 Clinical Laboratory Tests – Urinalysis (Safety Population)
Listing 16.2.3.3.1 Vital Signs (Safety Population)
Listing 16.2.3.3.2 Derived Vital Signs Variables (Safety Population)
Listing 16.2.3.3.3 Subject Body Size (Safety Population)
Listing 16.2.3.4.1 12-Lead ECG Intervals (Safety Population)
Listing 16.2.3.4.2 12-Lead ECG Overall Interpretation (Safety Population)
Listing 16.2.3.4.3 Derived 12-Lead ECG Variables (Safety Population)
Listing 16.2.3.5 Columbia Suicidal Severity Rating Scale (Randomized Population)
Listing 16.2.3.6.1 Simpson-Angus Rating Scale (Safety Population)
Listing 16.2.3.6.2 Barnes Akathisia Rating Scale (Safety Population)
Listing 16.2.3.6.3 Abnormal Involuntary Movement Scale (Safety Population)
Listing 16.2.3.7 Infrequent Behaviors (Safety Population)

Appendix 3. Table, Figures and Listing Shells

Shells are presented as a separate document.