Statistical Analysis Plan

Sponsor	Supernus Pharmaceuticals, Inc.
Protocol Title:	A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Multicenter, Parallel-Group Flexible-Dose Study of the Efficacy and Safety of SPN-812 in Adults with Attention-Deficit/Hyperactivity Disorder (ADHD)
Protocol Number:	812P306
Protocol Date:	06-Mar-2020
[redacted]	[redacted]
Document Version:	Final Version 3.0
Document Date:	20-Oct-2020

Signature Page

[redacted]

[redacted]

Document History

Version	Date	Section	Description of Change/Purpose
1.0	22-Jul-2020	N/A	N/A
2.0	18-Sep-2020	1.0	Protocol version was corrected.
2.0	18-Sep-2020	3.5.1, 3.5.3	Corrected description of timing of first dose
2.0	18-Sep-2020	4.0	The compliance calculation was corrected to include number of capsules that the subject was instructed to take daily. Added item 40 to SDQ-1 subscale derivation.
2.0	18-Sep-2020	5.3	Clarified that T-scores will now be derived instead of coming from an external vendor.
2.0	18-Sep-2020	6.1	Removed references to unused ADaM dataset names.
2.0	18-Sep-2020	6.10.3	Added SDQ total and mean scores to the demographic and baseline characteristic summary. [redacted]
2.0	18-Sep-2020	6.11	Added a summary of subjects in each dose group by week.
2.0	18-Sep-2020	6.12.2	MMRM will be used for sensitivity analysis

			by handling missing data with MI under MNAR in place of ANCOVA at the EOS based on FDA feedback.
2.0	18-Sep-2020	6.12.3	Added a description of how effect size will be calculated and presented.
2.0	18-Sep-2020	7.3	Vital signs normal ranges were updated to reflect medical safety review.
2.0	18-Sep-2020	6.13.1, 6.13.2, 6.14, 8.0	Verbiage added clarifying that MMRM will be used for secondary efficacy [redacted] analyses of endpoints except for BRIEF-A [redacted] endpoint, which will still use ANCOVA.
2.0	18-Sep-2020	10.0	[redacted].
2.0	18-Sep-2020	11.0	Removed table, figure, and listing titles since they are contained in a separate appendix to this SAP document.
3.0	20-Oct-2020	2.6	Added clarifying text with respect to Treatment Policy Estimand and population summary about handling missing data with respect to subjects who discontinued due to AE or lack of efficacy

3.0	20-Oct-2020	2.6	Listed the analysis datasets that will contain the imputed data for efficacy analyses
3.0	20-Oct-2020	6.7.1	Added clarifying text about how to impute data for subjects who discontinued due to AE or lack of efficacy
3.0	20-Oct-2020	6.12.2	Added pseudocode for MIANALYZE step of multiple imputation

Table of Contents

1. Introduction 11
2. Study Objectives 11
2.1. Primary Objective 11
2.2. Key Secondary Objective 11
2.3. Additional Secondary Objectives 12
2.4. Safety Objective 12
[redacted] 12
2.6. Estimand 12
3. Study Description 13
3.1. Study Design 13
3.2. Schedule of Visits and Study Procedures 14
3.3. Study Population 17
3.4. Completion and Discontinuation of Subjects 17
3.5. Study Treatments 18
3.5.1. Study Medication Identity, Packaging, and Labeling 18
3.5.2. Study Medication Administration 18
3.5.3. Method of Assigning Subjects to Treatment Groups 18
3.5.4. Blinding and Unblinding 19
3.5.5. Concomitant Medications 19
3.6. Sample Size and Power 19
4. Definitions and Derivations 20
5. Study Variables 24
5.1. Primary Efficacy Variable 24
5.2. Key Secondary Efficacy Variable 24
5.3. Additional Secondary Efficacy Variables 25
[redacted] 26
5.5. Safety Variables 26
[redacted] 27
6. Statistical Methods 27
6.1. General Principles 27
6.2. Analysis Populations 28
6.3. Interim Analysis and Data Monitoring 28
6.4. Adjustments for Covariates 28
6.5. Hypotheses 28
6.6. Multiple Comparisons 29

6.7. Handling of Dropouts or Missing Data 29
6.7.1. Missing Efficacy Data 29
6.7.2. Missing Safety Variables 30
6.8. Analysis Visit Windows 31
6.9. Pooling of Sites 31
6.10. Study Subjects and Demographics 31
6.10.1. Disposition of Subjects and Withdrawals 31
6.10.2. Protocol Violations and Deviations 32
6.10.3. Demographics and Other Baseline Characteristics 32
6.11. Exposure and Compliance 33
6.12. Efficacy Analysis 33
6.12.1. Primary Efficacy Analysis 33
6.12.2. Sensitivity Analysis 34
[redacted]
6.13. Analysis of Secondary Endpoints 35
6.13.1. Key Secondary Analysis 35
6.13.2. Additional Secondary Analyses 36
[redacted]
6.15. Supplementary Analyses 37
[redacted]
7. Safety and Tolerability Analysis 38
7.1. Adverse Events 38
7.1.1. Adverse Events Leading to Withdrawal 39
7.1.2. Deaths and Serious Adverse Events 39
7.2. Clinical Laboratory Evaluations 39
7.3. Vital Signs 40
7.4. Electrocardiograms 40
7.5. Physical Examination 41
7.6. Columbia Suicide Severity Rating Scale (C-SSRS) 41
7.7. Concomitant Medications 41
7.8. Caffeine Consumption 41
7.9. Other Safety Assessments 41
8. Changes from Planned Analysis 41
9. Validation of Analysis Reports 42
10. References 42
11. Planned Table, Figure, and Listing Descriptions 43

LIST OF ABBREVIATIONS

[redacted]	[redacted]
ADHD	Attention-deficit/hyperactivity disorder
AE	Adverse event
AESI	adverse events of special interest
AIC	Akaike information criterion
AISRS	ADHD Investigator Symptom Rating Scale
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
BP	blood pressure
bpm	beats per minute
BRIEF-A	Behavior Rating Inventory of Executive Function – Adult Version
BRI	Behavioral Regulation Index
CFB	change from baseline
CGI-I	Clinical Global Impression – Improvement
CGI-S	Clinical Global Impression – Severity of Illness
CI	confidence intervals
CRF	case report form
CRO	contract research organization
C-SSRS	Columbia Suicide Severity Rating Scale
CSR	clinical study report
DSM-5™	Diagnostic and Statistical Manual of Mental Disorders – 5th Edition
ECG	electrocardiogram
EOS	end of study
ET	early termination
FAS	Full Analysis Set
FOCP	females of childbearing potential
FPC	follow-up phone call
FSH	follicle stimulating hormone
GAD-7	Generalized Anxiety Disorder 7-Item scale
GEC	Global Executive Composite
HAM-A	Hamilton Anxiety Rating Scale
IC or ICF	informed consent or informed consent form

ICH	International Council for Harmonization
LS	least squares
MAR	missing at random
MedDRA	medical dictionary for regulatory activities
mg	milligram
MI	multiple imputation
MI	Metacognitive Index
MMRM	mixed model with repeated measures
MNAR	missing not at random
ms	millisecond
N	number
NA	not applicable
OLE	open-label extension
PK	pharmacokinetic
POC	point of care
PP	per protocol
PT	preferred term
QD	once daily
QOL	quality of life
REML	restricted maximum liklihood
SAE	serious adverse event
SAP	statistical analysis plan
SAS®	a software system used for data analysis
SCID-5-CT	Structured Clinical Interview for DSM-5 Clinical Trials version 5
SD	standard deviation
SDQ	Symptoms of Depression Questionnaire
SM	study medication
SOC	system organ class
TEAE	treatment-emergent adverse event
UDS	urine drug screen
WHO	World Health Organization
WHO-DDE	World Health Organization Drug Dictionary Enhanced

List of Tables

Table 1: Schedule of Events 14
Table 2: Study Medication Administration 18
Table 3: Vital Sign Normal Ranges 40

List of Figures

Figure 1: Study Schematic 14

1. Introduction

This document describes the statistical analyses and data presentations to be performed on Study 812P306, (A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Multicenter, Parallel-Group Flexible-Dose Study of the Efficacy and Safety of SPN-812 in Adults with Attention-Deficit/Hyperactivity Disorder (ADHD)), protocol version 4.0, dated 06Mar2020).

The purpose of this SAP is to ensure the credibility of the study findings by specifying detailed 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 describe handling of data issues. It describes the efficacy and safety variables, anticipated data manipulations, and other details of the analyses not provided in the study protocol. The analysis methods 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.

2. Study Objectives

2.1. Primary Objective

The primary objective is to evaluate the efficacy of SPN-812 200 mg to 600 mg compared to placebo for the treatment of ADHD in adults as measured by the adult ADHD Investigator Symptom Rating Scale (AISRS) total score.

2.2. Key Secondary Objective

The key secondary objective is to evaluate the efficacy of SPN-812 200 mg to 600 mg compared to placebo on the global assessment of severity for ADHD as measured by the Clinical Global Impression – Severity of Illness (CGI-S) scale.

2.3. Additional Secondary Objectives

The additional secondary objectives are to evaluate the efficacy of SPN-812 200 mg to 600 mg compared to placebo on:

1. Clinical response rate of severity of illness as measured by the categorical CGI-S Responder Rate (CGI-S score 1 or 2)
2. Global assessment of improvement as measured by the Clinical Global Impression – Improvement scale (CGI-I) for ADHD
3. Clinical response rate of improvement as measured by the categorical CGI-I Responder Rate (CGI-I score 1 or 2)
4. Anxiety symptoms as measured by the Generalized Anxiety Disorder 7-Item scale (GAD-7)
5. Clinician-rated ADHD symptoms as measured by the AISRS subscales of Inattention and Hyperactivity/Impulsivity
6. Clinician response rate of ADHD symptom reduction as measured by the 50% Responder rate in AISRS total score
7. Clinician response rate of ADHD symptom reduction as measured by the 30% Responder rate in AISRS total score
8. Executive functioning as measured by Behavior Rating Inventory of Executive Function – Adult Version (BRIEF-A; Self Report)
9. Aspects of executive function and problems of self-regulation as measured by the BRIEF-A Summary Index Scales and BRIEF-A scales

2.4. Safety Objective

The safety objective is to evaluate the safety and tolerability of SPN-812 200 mg to 600 mg in adult subjects with ADHD.

[redacted]

2.6. Estimand

Consistent with the ICH E9 Addendum, definition of the attributes of the estimand (target of estimation) is provided in this section.

1. Population: The population targeted for the scientific question is defined via the inclusion and exclusion criteria in adults diagnosed with ADHD per Diagnostic and Statistical

Manual of Mental Disorders – 5th Edition (DSM-5™) criteria.

2. Variable (or endpoint): Change from baseline to Week 6 of the double-blind study period in AISRS 18-item, each item scored on a scale of 0 to 4, where 0 = none and 3 = severe, total score ranging from 0 to 54.

3. Intercurrent Events: A treatment policy estimand will be followed. All observed values will be used regardless of occurrence of an intercurrent event. If a subject is discontinued due to AE or lack of efficacy, missing data will be handled as stated in item 4 below. i

4. Population level summary: The change from baseline in AISRS Total Score to Week 6 (EOS), will be analyzed using a Mixed Model for Repeated Measures (MMRM), which assumes that missing data are missing at random (MAR). The model will include fixed effect terms for baseline AISRS total score, treatment, study visit, and treatment-by-study visit interaction as independent variables. Missing data will be accounted for by the MMRM model except for subjects discontinued due to AE or lack of efficacy, which will be explicitly imputed using multiple imputation under Missing not at Random assumption.

3. Study Description

3.1. Study Design

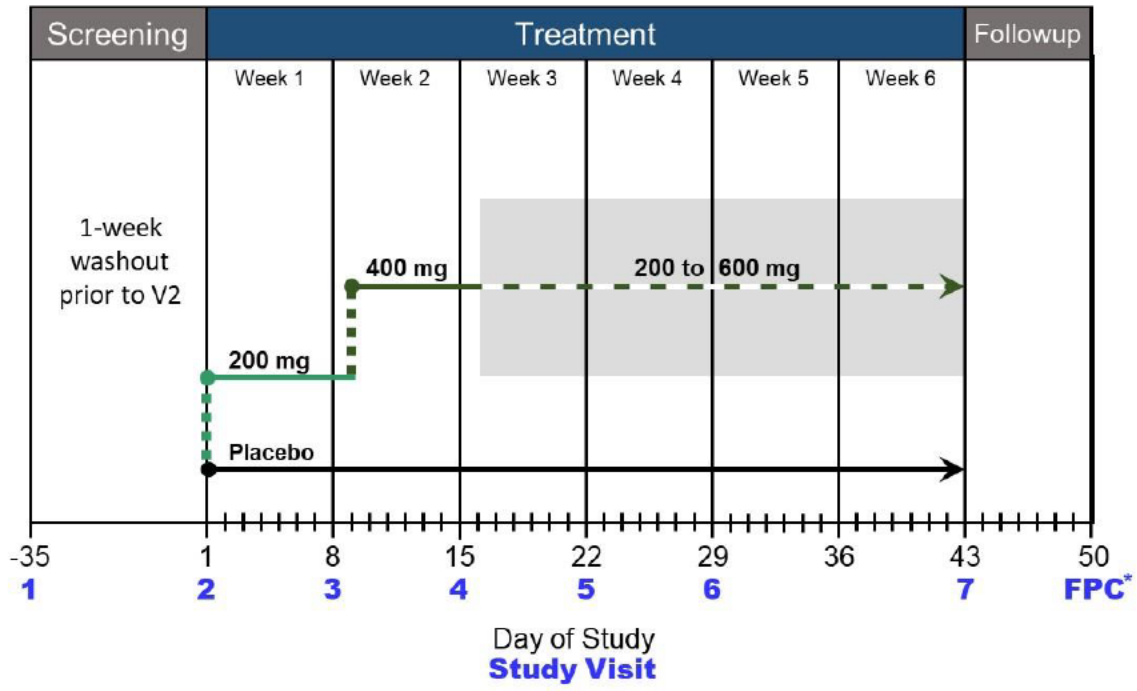
This is a randomized, double-blind, placebo-controlled, multicenter, parallel group, flexible dose study of SPN-812 in adults diagnosed with ADHD per Diagnostic and Statistical Manual of Mental Disorders – 5th Edition (DSM-5™) criteria. Approximately 366 subjects will be randomized in a 1:1 ratio (183 subjects per arm): SPN-812 (200 mg to 600 mg) or placebo. Following up to 5 weeks of screening, subjects will be treated with study medication (SM) for 6 weeks.

At the Screening Visit (Visit 1), after informed consent is obtained, subjects will undergo initial screening evaluations/scales, and inclusion/exclusion criteria will be reviewed to confirm the subject's eligibility. Subjects taking ADHD medication at screening will undergo a washout of at least 1 week (or 5 half-lives of the medication, whichever is longer) before the Baseline Visit (Visit 2; Day 1). At the Baseline Visit, subjects who meet inclusion/exclusion criteria will be randomized. The Treatment Period will consist of 6 weeks of treatment starting Day 2 (morning after Visit 2) until Day 43 (Visit 7; Week 6, EOS). At study visits, subjects will undergo efficacy and safety evaluations, compliance evaluations, return the previous dosing card, and receive the next card. The total study duration from the Screening Visit to the end of the Treatment Period is up to 11 weeks.

Subjects who complete the study may enroll in a separate open-label extension (OLE) safety study, 812P311. Those subjects who do not enroll in the OLE safety study will receive a phone call 1 week after their EOS visit for safety.

A study schematic is provided in Figure 1 below.

Figure 1: Study Schematic



*A safety follow-up phone call will be performed 1 week after EOS Visit only for those subjects who do not enroll/rollover into the Open-Label Extension study

3.2. Schedule of Visits and Study Procedures

All subjects who are randomized and take the initial dose of 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 Events

Study Period	Screening	Baseline	Treatment				EOS/ET	Follow-up[a] (phone call)
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**FPC[a]**
Day of Study	**-35 to -1**	**1**	**8**	**15**	**22**	**29**	**43**	**50**
Week of Study	**-5 to -1**	**–**	**1**	**2**	**3**	**4**	**6**	**–**
Study Visit Window	**–**	**–**	**±2**	**±2**	**±2**	**±2**	**±2**	**±2**
Signed informed consent	√							
SCID-5-CT	√							
HAM-A	√							
Relevant histories (social, medical, psychiatric, family psychiatric, neurological)	√							
Demographics	√							

Study Period	Screening	Baseline	Treatment				EOS/ET	Follow-up[a] (phone call)
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**FPC[a]**
Day of Study	**-35 to -1**	**1**	**8**	**15**	**22**	**29**	**43**	**50**
Week of Study	**-5 to -1**	**–**	**1**	**2**	**3**	**4**	**6**	**–**
Study Visit Window	**–**	**–**	**±2**	**±2**	**±2**	**±2**	**±2**	**±2**
Smoking, alcohol consumption use/history	√							
Physical examination	√						√	
Height	√							
Blood sample for FSH (post-menopausal females only)	√							
Serum pregnancy test (FOCP only)	√							
Serology	√							
Subject training video module	√	√		√		√		
Review eligibility criteria	√	√						
Randomize		√						
Review adverse events			√	√	√	√	√	√
Review concomitant medications and caffeine use	√	√	√	√	√	√	√	√
Hematology/serum chemistry	√	√					√	
Urinalysis	√	√					√	
ECG	√	√					√	
Serum drug screen	√							
Urine drug screen [c, g]	√ [c, g]	√ [h]	√ [h]	√ [h]	√ [h]	√ [h]	√ [h]	
Urine pregnancy test (FOCP)		√	√	√	√	√	√	
Orthostatic BP/HR [b]	√	√	√	√	√	√	√	
Vital signs [d] and weight	√	√	√	√	√	√	√	
C-SSRS	√	√	√	√	√	√	√	
AISRS [f]	√	√	√	√	√	√	√ [e]	
CGI-S [f]	√	√	√	√	√	√	√ [e]	
CGI-I [f]			√	√	√	√	√ [e]	
GAD-7		√		√		√	√ [e]	
BRIEF-A		√					√ [e]	
SDQ	√	[redacted]						
[redacted]								
Blood sample for PGx (optional)		√						

Study Period	Screening	Baseline	Treatment				EOS/ET	Follow-up[a] (phone call)
Visit Number	1	2	3	4	5	6	7	FPC[a]
Day of Study	-35 to -1	1	8	15	22	29	43	50
Week of Study	-5 to -1	–	1	2	3	4	6	–
Study Visit Window	–	–	±2	±2	±2	±2	±2	±2
SM dispensed		√	√	√	√	√		
SM returned and accountability			√	√	√	√	√	

[redacted]; ADHD = attention-deficit/hyperactivity disorder; AISRS = Adult ADHD Investigator Symptom Rating Scale; BP = blood pressure; BRIEF-A = Behavior Rating Inventory of Executive Function – Adult Version (Self); CGI-I = Clinical Global Impression – Improvement scale; CGI-S = Clinical Global Impression – Severity of Illness scale; C-SSRS = Columbia Suicide Severity Rating Scale; ECG = electrocardiogram; EOS = end of study; ET = early termination; FOCP = females of childbearing potential; FPC = follow-up phone call; FSH = follicle stimulating hormone; GAD-7 = Generalized Anxiety Disorder 7-item scale; HAM-A = Hamilton Anxiety Rating Scale; HR = heart rate; PGx = pharmacogenomics; [redacted]; POC = "Point of Care'; SCID-5-CT = Structured Clinical Interview for DSM-5 Clinical Trials version; SDQ = Symptoms of Depression Questionnaire; SM = study medication; UDS = urine drug screen.

a. Follow-up assessments (via phone calls) are only required for subjects who do not enroll in the open-label extension study.
b. Orthostatic blood pressure and heart rate should be measured after the subject has been seated for at least 5 minutes and within 3 minutes of standing.
c. With sponsor approval, subjects who exhibit a positive UDS for cannabis at the Screening Visit and, per Investigator judgement, are not considered a habitual/chronic cannabis user, may undergo an additional UDS urine drug screen at the Baseline Visit to determine eligibility; see Exclusion Criterion 14 in Section 3.3.3 of the Protocol for details.
d. Blood pressure, heart rate, respiratory rate, and oral temperature.
e. If subject's ET visit is conducted >7 days after the date of subject's last dose, do not perform/collect efficacy assessments at ET visi[redacted].
f. Investigator-rated efficacy assessments should be performed/conducted/collected prior to administering any self-report efficacy assessments to subjects.
g. Perform Standard UDS and Point of Care UDS Test 2 only (V1; Table 3 of the Protocol).
h. Perform Point of Care UDS both Test 1 and Test 2 (V2-V7; Table 3 of the Protocol).

3.3. Study Population

The study population will comprise adults who are 18 to ≤65 years of age diagnosed with ADHD with a Visit 2 AISRS total score of ≥26 and a CGI-S score of ≥4.

3.4. Completion and Discontinuation of Subjects

Subjects will be considered to have completed the study if they complete all visits up to and including Visit 7 (EOS).

Subjects who are randomized and dosed with SM, but who withdraw or are withdrawn from participation in the study by the Investigator before he/she finishes the study (i.e., after Visit 2 but to prior Visit 7), should complete an ET Visit. Procedures listed for Visit should be completed at ET visit, with the following exception:

If subject's Early Termination (ET) visit occurs:

- >7 days after the date of subject's last dose, efficacy assessments **should not** be performed/collected at ET visit.
- ≤7 days after the date of subject's last dose, efficacy assessments should be performed/collected at ET visit;

All reasons for screening failure will be recorded. If the subject passes screening but fails eligibility at Visit 2 (Baseline Visit), the reason(s) will also be recorded.

The Investigator(s) or subjects themselves may stop SM treatment at any time for safety or personal reasons. A subject is free to withdraw from the study at any time for any reason without prejudice to their future medical care by the physician or at the institution. The Sponsor may also withdraw the subject at any time in the interest of subject safety. The withdrawal of a subject from the study should be discussed where possible with the Medical Monitor and Clinical Research Associate (CRA) before the subject stops SM. Subjects removed from the study for any reason will not be replaced.

Reasons for a subject's early discontinuation may include:

- Withdrawal of consent
- Noncompliance
- Occurrence of unmanageable AEs
- Lost to follow-up
- Other

The primary reason for the subject's early discontinuation, including specific reason why subject withdrew consent or PI ended subject's study, must be recorded in the subject's medical record and on the electronic case report form (eCRF). If the subject withdraws consent or the Investigator discontinues the subject's participation in the study, the reason for the subject's withdrawal or Investigator's discontinuation of the subject should also be documented and captured on the eCRF. If a subject is withdrawn for more than one reason, each reason should be documented in the source document and the most medically significant reason should be entered on the eCRF.

If a subject misses doses of SM during this study, the Investigator shall counsel the subject on the importance of compliance. If the subject has consistently missed doses, he or she may be discontinued from the study at the discretion of the Investigator and in consultation with the Medical Monitor; all procedures for discontinuation will be followed.

3.5. Study Treatments

3.5.1. Study Medication Identity, Packaging, and Labeling

Study medication is supplied as capsules packaged in a double-blind configuration and supplied by the sponsor in labeled blister cards. Each SM blister card will include identical-looking capsules that contain 200 mg of SPN-812 or matching placebo to provide daily dose levels of 0 mg, 200 mg, 400 mg, or 600 mg of SPN-812. Each blister card will supply a subject with 7 days of dosing plus 2 extra days if needed. Each card will be labeled with the protocol number, at a minimum.

3.5.2. Study Medication Administration

Study medication will be administered orally once daily (QD) as intact capsules in the morning. The subject should take first dose of study medication on Day 2 (morning after Visit 2). During the first 2 weeks of the Treatment Period, regardless of treatment arm, subjects will take 2 capsules QD. During the remaining 4 weeks of the Treatment Period, subjects may take a minimum of 1 capsule daily up to a maximum of 3 capsules QD. See Table 2 below for the SM administration schedule. Each SPN-812 capsule contains 200 mg SPN-812.

Table 2: Study Medication Administration

Treatment Arm	Number and Identity of Capsules to be Taken QD During the Treatment Period		
	Week 1	**Week 2**	**Week 3 to Week 6**
A: Placebo	2 Placebo	2 Placebo	1 to 3 Placebo
B: SPN-812	1 Placebo 1 SPN-812	2 SPN-812	1 to 3 SPN-812
Abbreviation: QD = once daily. Each SPN-812 capsule contains 200 mg SPN-812			

3.5.3. Method of Assigning Subjects to Treatment Groups

Eligible subjects will be randomized at Visit 2 (Baseline) in a 1:1 ratio to SPN-812 (200 mg to 600 mg) or placebo.

Treatment A: Placebo

Treatment B: SPN-812 (200 mg to 600 mg)

Allocation of study treatment will occur centrally via an interactive web response system (IWRS) using a randomization schedule to determine the SM assignment for each subject being randomized. A dosing card(s) will be given to the subject at each study visit starting at Visit 2.

3.5.4. Blinding and Unblinding

The subject and all personnel involved with the conduct and interpretation of the study, including the Investigators, study site personnel, the sponsor and contract research organization (CRO) clinical staff, including the Medical Monitor, will be blinded to SM. [redacted] Randomization schedule data will be kept strictly confidential, filed securely, and accessible only to authorized persons until the time of unblinding.

3.5.5. Concomitant Medications

Subjects may not be on any prohibited medication as indicated in the inclusion/exclusion criteria. SPN-812 is a strong CYP1A2 inhibitor. Substrates with a narrow therapeutic window are prohibited during the study. Specific prohibited concomitant medications for this study include known CYP1A2 substrates (e.g., theophylline, melatonin). Subjects receiving prior ADHD medication at screening will undergo a washout period of at least 1 week (or 5 half-lives of the medication, whichever is longer) before the Baseline Visit (Day 1).

No concomitant medications are allowed during the study, with the following exceptions:

- Nutritional supplements (e.g., multivitamins, fish oil) (herbal supplements are prohibited)
- EMLA® or other numbing cream for venipuncture
- Common over-the-counter therapies for minor transient ailments (e.g., acetaminophen for headache, ibuprofen for fever).

Additional concomitant medications are allowable on a case-by-case basis at the discretion of the Investigator and sponsor approval.

All concomitant medications will be recorded in the eCRF.

Caffeine use is permitted during the study and will be recorded in the eCRF.

3.6. Sample Size and Power

Assuming an effect size of 0.407, 128 subjects per treatment group (256 total subjects for 2 arms) in the Full Analysis Set (FAS) will yield 90% power at a significance level of 0.05 (2-sided) to reject the equality of treatment means between the placebo and the SPN-812 treatment group. Assuming approximately 30% of subjects drop out before the completion of the study, an adjusted sample size of 366 subjects (183 per arm) will be randomized to obtain 128 subjects per arm in the FAS at the completion of the study.

4. Definitions and Derivations

- Baseline is the last non-missing assessment recorded before receiving the first dose of study medication, will be used as the baseline observation for all calculations of change from baseline.
- **AISRS Total Score** = sum of the 18 individual items. Each individual item is scored on a scale of 0 to 4, where 0 = none and 3 = severe, so the total score ranges from 0 to 54. Higher scores indicate more severe symptoms. If more than 3 items of AISRS are missing then the total score will be set to missing. If ≤3 items are missing then the values for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer.
- **AISRS Inattention Subscale Score** = sum of the odd numbered individual items. Each individual item is scored on a scale of 0 to 3, where 0 = none and 3 = severe, so the total score ranges from 0 to 27. Higher scores indicate more severe symptoms. If more than 3 items of AISRS are missing then the subscale score will be set to missing. If ≤3 items are missing then the values for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer.
- **AISRS Hyperactivity/Impulsivity Subscale Score** = sum of the even numbered individual items. Each individual item is scored on a scale of 0 to 3, where 0 = none and 3 = severe, so the total score ranges from 0 to 27. Higher scores indicate more severe symptoms. If more than 3 items of AISRS are missing then the subscale score will be set to missing. If ≤3 items are missing then the values for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer.
- **Percent Reduction in AISRS Total Score** = 100 * (AISRS Total Score at Week 6 (EOS) - Baseline AISRS Total Score) / Baseline AISRS Total Score.
- **GAD-7 Total Score** = sum of the 7 individual items. Each individual item is scored on a scale of 0 to 3, where 0 = not at all and 3 = nearly every day, so the total score ranges from 0 to 21. Higher scores indicate worse anxiety. If 1 of the individual item scores is missing, the total score will be set to missing.
- **BRIEF-A Global Executive Composite (GEC) Total Raw Score** = sum of the BRI and MI, with a range of 70 to 210. Higher scores indicate poorer executive function. If the BRI or MI is missing, the total score will be set to missing.
- **BRIEF-A Behavioral Regulation Index (BRI) Raw Score** = sum of the Inhibit, Shift, Emotional Control, and Self-Monitor Scale Scores, with a range of 30 to 90. Higher scores indicate poorer executive function. If 1 of the scale scores is missing, the BRI will be set to missing.
- **BRIEF-A Metacognitive Index (MI) Raw Score** = sum of the Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials Scale Scores with a range of

40 to 120. Higher scores indicate poorer executive function. If 1 of the scale scores is missing, the MI will be set to missing.

- **BRIEF-A Inhibit Scale Score** = sum of items 5, 16, 29, 36, 43, 55, 58, and 73. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the total score ranges from 8 to 24. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Self-Monitor Scale Score** = sum of items 13, 23, 37, 50, 64, and 70. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the scale score ranges from 6 to 18. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Plan/Organize Scale Score** = sum of items 9, 15, 21, 34, 39, 47, 54, 63, 66, and 71. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the scale score ranges from 10 to 30. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Shift Scale Score** = sum of items 8, 22, 32, 44, 61, and 67. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the scale score ranges from 6 to 18. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Initiate Scale Score** = sum of items 6, 14, 20, 25, 45, 49, 53, and 62. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the scale score ranges from 8 to 24. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Task Monitor Scale Score** = sum of items 2, 18, 24, 41, 52, and 75. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the scale score ranges from 6 to 18. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Emotional Control Scale Score** = sum of items 1, 12, 19, 28, 33, 42, 51, 57, 69, and 72. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the scale score ranges from 10 to 30. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Working Memory Scale Score** = sum of items 4, 11, 17, 26, 35, 46, 56, and 68. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often, so the scale score ranges from 8 to 24. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

- **BRIEF-A Organization of Materials Scale Score** = sum of items 3, 7, 30, 31, 40, 60, 65, and 74. Each individual item is scored on a scale of 1 to 3, where 1 = never and 3 = often,

so the scale score ranges from 8 to 24. Higher scores indicate poorer executive function. If 1 of the individual item scores is missing, the scale score will be set to missing.

[redacted]

- **SDQ Average (Mean) Score** = average of the 44 individual item scores. Each individual item is scored on a scale of 1 to 6, so the average score ranges from 1 to 6. Higher scores indicate worse depression. If 1 of the individual item scores is missing, the average (mean) score will be set to missing.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

- **HAM-A Total Score =** sum of the 14 individual items. Each individual item is scored on a 5-point Likert scale, where 0 = Not present, 1 = Mild, 2 = Moderate, 3 = Severe, and 4 = Very severe. The total score ranges from 0 to 56, where ≤17 indicates "mild anxiety severity", 18 to 24 indicates "moderate anxiety severity", 25 to 30 indicates "moderate to severe anxiety severity", and ≥31 indicates "severe anxiety". If 1 of the individual item scores is missing, the total score will be set to missing.

- **Change from Baseline** = value at current time point – value at baseline.

- **Percent Change from Baseline** = (change from baseline / value at baseline) * 100.

- **Body Mass Index (BMI) kg/m^2** =

$$\left(\frac{weight\ (kg)}{\frac{height(cm)}{100} * \frac{height(cm)}{100}}\right)$$

- **Treatment Emergent Adverse Event (TEAE)** = any adverse event with an onset date/time after first dose of study medication.
- **Duration of Adverse Event** = AE end date – AE start date + 1
- **Duration of Treatment Exposure** = date of last dose – date of first dose + 1
- **Percent of Study Medication Compliance** =

$$\sum_{i=2}^{i=7}\left(\frac{Number\ of\ capsules\ dispensed\ at\ visit\ (i-1)-Numberof\ capsules\ returned\ at\ visit\ i}{X*(Number\ of\ days\ from\ Visit\ i-1\ to\ \ Visit\ i)}\right)*100$$

Where X is equal to the number of capsules that the subject was instructed to take daily during the treatment period (between visits) i.e. 1 (200 mg) or 2 (2 x 200 mg), or 3 (3 x 200 mg).

5. Study Variables

5.1. Primary Efficacy Variable

The primary efficacy endpoint of this study is the change from baseline (CFB) at end of study (EOS) in the AISRS total score.

The adult ADHD Investigator Symptom Rating Scale (AISRS) was developed to better measure the presence and severity of ADHD symptoms based on DSM-IV diagnostic criteria in adult patients[4]. It is a semi-structured clinical interview with suggested prompts for each item to improve interrater reliability. The scale consists of 18 items that directly correspond to the 18 symptoms of ADHD and are further subdivided into two subscales: Inattention (9 items) and Hyperactivity/Impulsivity (9 items). During the interview with the subject, the clinician/investigator rates the frequency and severity of each symptom on a 4-point Likert-type scale, where 0 = None, 1 = Mild, 2 = Moderate, and 3 = Severe, with a maximum total score of 54 points and maximum subscale score of 27 points. Derivations for scores are described in Section 4.

The scale allows the assessment of functional impairments linked to each symptom dimension. The AISRS total score is the sum of the Inattention and Hyperactivity/Impulsivity subscale scores.

The AISRS is used to assess drug efficacy in the treatment of ADHD in adults, and the AISRS total score is the primary outcome measure for this study.

5.2. Key Secondary Efficacy Variable

The key secondary efficacy endpoint is the CFB at EOS in the CGI-S score.

The CGI-S scale is a single item clinician rating of the clinician's assessment of the severity of the ADHD symptoms in relation to the clinician's total experience with patients with ADHD.

The CGI-S is evaluated on a 7-point scale, where 1 = Normal, not at ill, asymptomatic, 2 = Borderline Ill, 3 = Mildly Ill, 4 = Moderately Ill, 5 = Markedly Ill, 6 = Severely Ill, and 7 = Extremely Ill. Successful therapy is indicated by a lower overall score in subsequent testing.

The CGI-S 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 SM[5].

5.3. Additional Secondary Efficacy Variables

The additional secondary efficacy endpoints of this study include the following:

1. Percentage of subjects with a CGI-S score of 1 or 2 at EOS
2. CGI-I score at EOS
3. Percentage of subjects with a CGI-I score of 1 or 2 at EOS

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 SM[5]. The CGI-I scale is an assessment of how much the patient's illness has improved or worsened relative to a baseline state at the beginning of treatment. The CGI-I is evaluated by the investigator at each post-baseline study visit during treatment relative to the subject's condition at baseline on a 7-point scale where 1 = Very much improved, 2 = Much improved, 3 = Minimally improved, 4 = No change, 5 = Minimally worse, 6 = Much worse, and 7 = Very much worse. Successful therapy is indicated by a lower overall score in subsequent testing. The CGI-S assessment obtained at the baseline visit should serve as a basis for the investigator's CGI-I assessment of improvement in the subject's conditions at each post-baseline study visit during the treatment period.

4. CFB at EOS in the GAD-7 total score

Generalized Anxiety Disorder 7 scale (GAD-7) is a self-reported 7-item questionnaire for screening and measuring the severity of generalized anxiety disorder[6]. The GAD-7 measures the severity of various symptoms of generalized anxiety disorder over the past 2 weeks according to reported response categories with assigned points. The patient scores each GAD-7 item on 4-point Likert scale, where 0 = Not at all, 1 = Several days, 2 = More than half the days, and 3 = Nearly every day. GAD-7 total scores range from 0 to 21, where a total score of 1 to 4 = None/Minimal anxiety, 5 to 9 = Mild anxiety, 10 to 14 = Moderate anxiety, and ≥15 = Severe anxiety. Derivations for scores are described in Section 4.

5. CFB at EOS in the AISRS Inattention subscale score and the Hyperactivity/Impulsivity subscale score
6. AISRS 50% Responder rate (defined as the percentage of subjects with a ≥50% reduction in the CFB AISRS total score) at EOS
7. AISRS 30% Responder rate (defined as the percentage of subjects with a ≥30% reduction in the CFB AISRS total score) at EOS

8. CFB at EOS in the BRIEF-A Global Executive Composite (GEC) T-score
9. CFB at EOS in the BRIEF-A T-score by each Summary Index Scale and by individual BRIEF-A scale.

The BRIEF-A is a standardized 75-item rating scale in 9 scales and 2 summary index scales that assesses aspects of executive function and problems with self-regulation from the perspective of

the individual. The subject rates each item on a 3-point Likert scale (Never, Sometimes, or Often) based on their experiences within the last month. Higher scores indicate poorer executive function.

The Global Executive composite *(GEC)* raw score is the sum the Behavioral Regulation Index *(BRI)* raw score and Metacognition Index *(MI)* raw score. The BRI raw score is the sum of raw scores of subscales Inhibit, Shift, Emotional Control, and Self-Monitor scales. To calculate the Metacognition Index *(MI)* raw score, sum the raw scores obtained for the Initiate, Working Memory, Plan/Organize, Task Monitor; and Organization of Materials scales. Raw GEC scores, BRI, MI, and their component subscale (individual scales) will be converted to T-scores and the CFB in T-scores will be calculated. Derivations for raw scores are described in Section 4. The raw GEC, BRI, MI and individual subscale scores will be converted to T-scores using Appendix A of the BRIEF-A Professional Manual.

[redacted]

5.5. Safety Variables

The safety endpoints of this study include the following:

- Adverse events (AEs)
- Clinical safety laboratory tests
- Vital signs
- Weight

- Electrocardiograms (ECGs)
- Physical examination
- Columbia Suicide Severity Rating Scale (C-SSRS)

[redacted]cted [redacted]

[redacted]

6. Statistical Methods

6.1. General Principles

All statistical analysis, data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will be performed using SAS version 9.4 or higher by Premier research, which is the designated CRO. Premier will be responsible for creating table, listing, and figure (TLF) reports via SAS programming. The reports and SAS programs will be delivered to Supernus at the completion of the study.

All tabulations of analysis results will include summaries for the two treatments: SPN-812 and placebo.

Continuous (quantitative) variable summaries will include the number of subjects (n) with non-missing values, mean, standard deviation (SD), median, minimum, and maximum.

Categorical (qualitative) variable summaries will include the frequency and percentage of subjects who are in the particular category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of subjects in the study population for the treatment groups, unless otherwise specified. The denominator for by-visit displays will be the number of subjects in the relevant study population with non-missing data at each study visit.

The minimum and maximum will be reported with the same degree of precision (i.e., the same number of decimal places) as the observed data. Measures of location (mean and median) will be reported to 1 degree of precision more than the observed data and measures of spread (SD) will be reported to 2 degrees of precision more than the observed data.

Percentages will be presented to 1 decimal place, unless otherwise specified and with the following exceptions: a percentage of 0% will not be displayed (i.e., only the count of 0 will be displayed), and a percentage of 100% will be displayed with 0 decimal places.

Derived analysis datasets will be produced from the SDTM data. Specifications for derived datasets will be developed. Analysis datasets to be created will include ADQS (Questionnaire analysis dataset), ADEF1 (Primary analysis efficacy dataset), ADEF2 (Sensitivity analysis efficacy dataset), ADSL (subject-level analysis dataset), ADAE (Adverse event analysis dataset) and others, as appropriate.

Data collected at unscheduled time points will not be summarized but will be presented in subject listings.

Unless otherwise indicated, all statistical tests will be conducted at the 0.05 significance level using 2-tailed tests, and *P* values will be reported. Corresponding 95% confidence intervals (CIs) will be presented for statistical tests, where applicable.

A *P* value of ≤0.10 but >0.05 will be considered evidence of a trend.

All *P* values will be displayed in four decimals and rounded using standard scientific notation (e.g., 0.XXXX). If a *P* value less than 0.0001 occurs it will be shown in tables as <0.0001

All collected data will be presented in listings. Data not subject to analysis according to this plan will not appear in any tables or graphs but will be included only in the data listings

6.2. Analysis Populations

The following analysis populations are planned for this study:

- **Randomized Population:** The Randomized Population includes all subjects who complete the Baseline Period, meet inclusion/exclusion criteria, and are randomized.
- **Full Analysis Set:** The FAS is a subset of subjects in the Randomized Population who took at least 1 dose of SM, and had a Baseline and at least one post-Baseline assessment of AISRS. Subjects in the FAS will be analyzed according to the treatment to which they were randomized. The efficacy analyses will be conducted using the FAS.
- **Per Protocol (PP) Population:** The PP Population is a subset of subjects in the FAS who complete all 7 visits through EOS with no missing AISRS assessments and no major protocol violations. Subjects in the PP Population will be analyzed according to the treatment received.
- **Safety Population:** The Safety Population includes all subjects randomized into the study who receive at least one dose of study medication. Subjects in the Safety Population will be analyzed according to the treatment received. The safety analyses will be conducted using the Safety Population.

6.3. Interim Analysis and Data Monitoring

No interim analyses are planned.

6.4. Adjustments for Covariates

If there is a statistical difference among treatment groups with respect to baseline characteristics, that variable may be added to the statistical models as a blocking factor or covariate to determine the effect on treatment. Baseline efficacy scale scores are planned for inclusion in statistical models, unless otherwise indicated. See Section 6.12 for more details.

6.5. Hypotheses

The primary objective of this study is to test the null hypothesis (H_0) is that there is no difference between SPN-812 and placebo groups with respect to LS means. The alternative hypothesis (H_a) is that the LS mean of SPN-812 is greater than the LS Mean of Placebo in absolute value. In other words, 1 daily dose of SPN-812 is superior to placebo in the treatment of Adults with

Attention-Deficit/Hyperactivity Disorder (ADHD) in adult patients.

6.6. Multiple Comparisons

The type I error rate for the study will be preserved at the 5% significance level by using sequential (hierarchical) testing for the primary and key secondary endpoints. The key secondary endpoint will only be tested if SPN-812 treatment is significantly different from placebo for the primary endpoint.

The testing will be done in the following order:

1. Primary: change from baseline in AISRS total score at EOS (Week 6).

2. Key Secondary: change from baseline in CGI-S score at EOS (Week 6).

Because the testing is sequential, the type I error rate of $\alpha = 0.05$ is maintained. Failure at any stage in the sequence implies automatic failure at all subsequent stages. All *P* values for each comparison without adjustments will be provided in the summary tables for informational purposes.

No adjustments will be made for multiple comparisons for other endpoints.

6.7. Handling of Dropouts or Missing Data

6.7.1. Missing Efficacy Data

All possible efforts will be made to ensure that subjects stay in the study and all data are collected as scheduled. Any subject who has received at least 1 dose of study medication who discontinues from the study for any reason will be encouraged to complete the end of study (EOS) visit.

If more than 3 items of AISRS are missing then the total score will be set to missing.
If ≤3 items are missing then the values for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer.

1. Missing endpoint values in this study may result from subjects discontinuing from the study prematurely or missing intermediate assessments while remaining on study. The primary efficacy analysis method will be based on a mixed model with repeated measures (MMRM) which utilizes all available data (complete and partial) from subjects included in an analysis set. The MMRM-based approach assumes that missing AISRS total scores are missing at random (MAR). MAR refers to a missingness mechanism that is independent of missing responses, conditionally on observed response history and covariates, that is, given the observed data, the reason for the missing data does not depend on the unseen data. This assumption inherently implies that the treatment effect is similar for those who discontinue prematurely and for those who complete the study in their respective treatment groups. Under MAR, the propensity for a data point to be missing is not related to the missing data, but is related to some of the observed data. Missing data for subjects discontinued due to AE or lack of efficacy will be explicitly imputed using multiple imputation under Missing not at Random assumption. .

.

Sensitivity analyses to investigate robustness to missing primary endpoint data will be performed by assuming that missing AISRS total scores are missing not at random (MNAR), meaning that the probability that an observation is missing may depend on its underlying unobserved value. See Section 6.12.2 for more details.

For analysis of secondary endpoints, missing values will be assumed to be MAR.

6.7.2. Missing Safety Variables

Missing dates for AEs and non-study concomitant medications will be imputed as described in using the following rules:

If partial AE or medication dates occur, the convention for replacing missing dates for the purpose of statistical analysis is as follows.

For partial AE and medication start dates:

- If the year is unknown, then do not impute the date but assign a missing value.
- If the year is known, but the month or month and day is/are unknown, then:
    - If the year matches the year of first dose date and the end date (if present) is after first dose date, then impute as the month and day of the first dose date.
    - Otherwise, assign 01 January.
- If the year and month are known, but the day is unknown, then:
    - If the month and year match the month and year of the first dose date, then impute as the day of the first dose date.
    - Otherwise, assign 01.

For partial AE and medication end dates:

- If the year is unknown, then do not impute the date but assign as missing value.
- If the year is known, but the month or month and day is/are unknown, then:
    - If the year matches the year of the last date of the study (date of last contact if subject lost to follow-up; date of completion or early termination), then impute as the month and day of the last date of the study.
    - Otherwise, assign 31 December.
- If the year and month are known, but the day is unknown, then:

- If the month and year match the month and year of the last date of the study, then impute as the day of the last date of the study.

- Otherwise, assign the last day of the month.

If partial times occur, the convention is as follows:

- if the missing time occurs on the day of the first dose and both the hour and minute are missing, then the time assigned is the time of the first dose, otherwise if both the hour and minute are missing and the date is not the date of first dose the time assigned is 12:00;

- if the date is the same as the date of the first dose and

  - only hour is missing, the hour assigned is 12 or the hour of first dose, whichever is later;

  - only the minute is missing, the minute assigned is 30 or the minute of first dose, whichever is later;

- Otherwise if the date is not the same as the date of first dose, the hour assigned is 12 if the hour is missing and the minute assigned is 30 if the minute is missing.

Missing data for all other safety endpoints will not be imputed.

6.8. Analysis Visit Windows

Data from scheduled visits will be analyzed. Visits will be analyzed as scheduled. Unscheduled measurements will be excluded from the descriptive statistics and statistical analyses but will be included in listings.

6.9. Pooling of Sites

This is a multicenter study. The primary analysis will be performed without site as a factor. However, if applicable, for certain [redacted] analyses small sites (defined as those with sample size <12 FAS subjects per site with uneven distribution of treatment groups) will be pooled based on geographic proximity.

6.10. Study Subjects and Demographics

6.10.1. Disposition of Subjects and Withdrawals

Subject disposition summary will presented for the randomized population. Disposition will include tabulations of the number and percentage of subjects in each of the following categories by treatment group and overall:

- Subjects in the Randomized Population
- Subjects in the FAS
- Subjects in the PP Population

- Subjects in the Safety Population

Within each of the previous categories, the number and percentage of subjects who completed and discontinued from the study and primary reason for early discontinuation will be summarized. The reason for early discontinuation may include any of the following:

- Withdrawal of consent
- Noncompliance
- Occurrence of unmanageable AEs
- Lost to follow-up
- Other

The number of subjects in each of the above analysis populations who will be continuing in the 311 OLE safety study will also be tabulated.

All disposition information as well as inclusion and exclusion criteria met/not met and reasons for screen failures will be listed.

6.10.2. Protocol Violations and Deviations

Protocol deviations will be listed and major protocol deviations will be summarized.

6.10.3. Demographics and Other Baseline Characteristics

Summary statistics for age, age group, sex, ethnicity, race, height and weight at screening, BMI, HAM-A score, AISRS Total score at screening, SDQ [redacted] mean score at Screening and CGI-S score at screening will be presented by treatment group and overall.

Age will be broken up into quartiles to determine age group. If the distribution of age is such that the use of quartiles does not lead to distinct intervals of values, other cut-off points will be examined (e.g. tertiles, above/below median).

Baseline comparability among the treatment groups will be summarized using a Chi-square test for categorical variables and one-way ANOVA (F-test) for continuous variables. *P*-values will be used for descriptive purposes only.

These analyses will be conducted for the FAS.

The number and percent of subjects reporting various medical histories, grouped by MedDRA system organ class (SOC) and preferred term (PT), will be tabulated by treatment group and overall. This analysis will be conducted for the Safety Population.

Prior medications will be summarized by treatment group, by the number and percentage of subjects taking each medication, classified using WHO-DDE ATC Class Level 4 and preferred term (ATC Class Level 5) and will be presented by treatment group and overall. This analysis will be conducted for the Safety Population. Prior medications are defined in Section 7.7.

The Structured Clinical Interview for DSM-5 Clinical (SCID) and the Hamilton Anxiety Rating Scale (HAM-A), as well as the alcohol, tobacco, and substance use collected at screening will be listed. Subject training module data will also be listed. See Section 4 for the derivation of the HAM-A total score.

6.11. Exposure and Compliance

Duration of treatment exposure will be summarized by treatment group and overall using descriptive statistics. Additionally, duration of treatment exposure will be summarized by duration category.

Duration of treatment exposure will be categorized as follows: 1-7 days, 8-14 days, 15-21 days, 22-35 days, 36-42 days, and >42 days

Subject exposure will be summarized categorically by study week, dose (200 mg, 400 mg and 600 mg), and treatment group. The number and percent of subjects in each dose group will be presented for each treatment.

For each treatment, SM compliance will be summarized by compliance category (<80%, 80 to 120%, and >120%) and number of subjects in each compliance category. Study medication compliance will also be summarized as a continuous variable using descriptive statistics.

This analysis will be conducted for the Safety Population. Derivations of duration of treatment exposure and compliance are defined in Section 4. All study medication administration, exposure, and accountability data collected will be listed.

6.12. Efficacy Analysis

All efficacy endpoints will be summarized using the FAS. All efficacy data will be presented in data listings. The mean profiles of AISRS Total Score, CGI-S score, and CGI-I score will be presented graphically by treatment group and scheduled study visit. In addition, plots of the Cumulative Distribution Functions of Change from Baseline in AISRS Total score will be presented. Bar charts of CGI-S and CGI-I response by treatment group and study visit will also be presented. Furthermore, bar charts for the absolute mean of GAD-7, BRIEF-A, [redacted] total scores by treatment group and study visit be presented.

To ensure the overall type I error rate of $\alpha = 0.05$ will be maintained for the study, hierarchical testing, based on a fixed sequence procedure, will be used for the primary endpoint and continuing for the key secondary endpoints. The MMRM primary analysis, as described in Section 6.12.1 will start the hierarchical testing. If statistical significance is declared for the primary MMRM analysis, formal hypothesis testing will be done for the key secondary endpoint (see Section 6.6 for sequence of testing) until a non-significant result is reached. The key secondary endpoint will only be tested if SPN-812 treatment is significantly different from placebo for the primary endpoint. All other *P* values, after a non-significant *P* value is reached, will be considered nominal.

6.12.1. Primary Efficacy Analysis

The primary efficacy endpoint, change from baseline in AISRS total score to Week 6, will be analyzed using MMRM, which assumes that missing data are MAR. The model will include change from baseline in AISRS score as the dependent variable, fixed effect terms for baseline AISRS total score, treatment, study visit, and treatment-by-study visit interaction as independent variables. All post-baseline study visits will be included in the model; however, the primary comparison will be between SPN-812 and placebo at Week 6. The model parameters will be estimated using the restricted maximum likelihood (REML) method with unstructured variance-covariance matrix and Kenward-Roger approximation to estimate denominator degrees of

freedom. If the unstructured covariance model fails to converge, the first (co)variance structure that does not have convergence problem will be used for the analysis from the following ordered list: 1) Toeplitz, 2) Autoregressive of order 1, and 3) Compound Symmetry.

The adjusted mean (LS Mean) of CFB to EOS of SPN-812 treatment group and placebo for AISRS Total Score will be presented, along with the corresponding standard error. The SPN-812 treatment group will be compared with the placebo by presenting the difference between the LS Means of SPN-812 and placebo (SPN-812 minus placebo), and 95% confidence intervals for the treatment difference and the p-value.

The observed value and change from baseline in AISRS score will also be summarized descriptively by study visit and treatment group using descriptive statistics in addition to the above noted model. Results from the MMRM primary analysis, described above, will be used to determine study success.

6.12.2. Sensitivity Analysis

The sensitivity analysis assumes that missing AISRS Total Scores are missing not at random (MNAR), that is, the probability that an observation is missing may depend on its underlying unobserved value. For example, the probability of missing AISRS Total Score at Week 4 is not related to the observed AISRS Total Score at Visit 5. Placebo-based multiple imputation will be used to fill in missing values. This approach may be considered "worst-case" sensitivity analysis as it assumes that after discontinuation, subjects from the SPN-812 treatment group would adopt the outcome model estimated from the placebo arm. The placebo-based imputation will be implemented by adopting the following three steps.

1. SAS PROC MI (SEED=220877) is applied to the input dataset containing all by-visit AISRS Total Score during the baseline and treatment period. Multivariate imputation will be carried out by the fully conditional specification (FCS) method. One hundred (100) multiply-imputed datasets will be created. SAS® PROC MI with the MNAR statement will be implemented by using the following SAS code.

[redacted]

```
RUN;
```

   In the above SAS code, EFF denotes the input efficacy analysis dataset, MNAROUT denotes the output dataset, Baseline, Week1, Week 2, Week 3, Week 4, and Week 6 are analysis values of AISRS total scores at baseline and at each of the treatment weeks from Week 1 to Week 6, where modelobs=(treatment=("placebo"))) means model observations will be from the placebo arm only.

Imputed values will be rounded to the closest integer. In case imputed values are below or above the range of the AISRS scale, values will be imputed to respectively the minimum or maximum value of the scale.

2. For each of the imputations from Step #1, CFB to study visit will be computed on the completed dataset.

   For each imputation the CFB will be analyzed using an MMRM ANCOVA model, which will include change from baseline in AISRS total score as the dependent variable, treatment, study visit, study visit-by-treatment interaction as fixed effect factors, and baseline AISRS total score as a covariate.

3. Finally, to combine estimates (LS means, LS mean treatment differences, 95% CI around the difference and *P* value) from the 100 datasets, SAS® PROC MIANALYZE will be used as in the following.

[redacted]

```
run;
```

The adjusted mean (LS Mean) of CFB to EOS of SPN-812 treatment group and placebo, along with the corresponding standard error, difference between the LS Means of SPN-812 and placebo (SPN-812 minus placebo), 95% confidence intervals for the treatment difference and the p-value will be presented from the combined estimates.

[redacted]

6.13. Analysis of Secondary Endpoints

6.13.1. Key Secondary Analysis

The key secondary efficacy endpoint, change from baseline in CGI-S score to Week 6 (EOS), will be analyzed using a MMRM. The model will include change from baseline in CGI-S score to Week 6 (EOS), as the dependent variable, baseline in CGI-S score as a covariate, and

treatment, study visit, and treatment-by-study visit interaction as independent fixed effect variables.

The adjusted mean (LS Mean) of CFB to EOS of SPN-812 treatment group and placebo, along with the corresponding standard error, difference between the LS Means of SPN-812 and placebo (SPN-812 minus placebo), 95% confidence intervals for the treatment difference and the p-value will be presented.

The null hypothesis is that the difference in change from baseline in SPN-812 minus placebo is 0 with a 2-sided alternative considering a difference in either direction.

The observed value and change from baseline in CGI-S score will also be summarized descriptively by study visit and treatment group using descriptive statistics in addition to the above noted model.

6.13.2. Additional Secondary Analyses

1) A CGI-S "responder" is defined as having a CGI-S score of 1 or 2. The number and percentage of CGI-S responders along with 95% CIs will be summarized by treatment group and study visit. A 2-sided Pearson's Chi-Square Test, or Fisher's Exact Test, as appropriate, will be performed at the 5% significance level to determine whether there is a significant difference between treatment groups. The null hypothesis is that there is no difference in proportion of CGI-S responders between SPN-812 and placebo, with a 2-sided alternative considering a difference in either direction. The difference in proportions between SPN-812 and placebo (SPN-812 minus placebo), 95% CIs for the differences, and *P*-values for the differences in treatment will also be presented.

2) The absolute value of in CGI-I score at each study visit including Week 6, will be analyzed using a MMRM with absolute CGI-I score as the dependent variable, baseline CGI-S score as a covariate, and treatment, study visit, treatment-by-study visit interaction as independent fixed effect factors. SPN-812 and placebo will be compared; the LS Means and the LS mean differences in CGI-I score will be presented along with associated 95% CIs of the treatment difference and *P*-values will be computed. The null hypothesis is that the difference in CGI-I score in SPN-812 minus placebo is 0 with a 2-sided alternative considering a difference in either direction. Additionally, Absolute CGI-I scores will be summarized by treatment group and study visit using descriptive statistics. CGI-I scores are collected at Week 1 through Week 6.

3) A CGI-I "responder" is defined as having a CGI-I score of 1 or 2. The analysis in number 1 above will be repeated for this endpoint.

4) Change from baseline in GAD-7 score to Week 6 will be analyzed using the same MMRM in number 2 above.

5) Change from baseline in AISRS Inattention and Hyperactivity/Impulsivity subscale score to Week 6, will be analyzed using the same MMRM in number 2 above.

6) A 50% AISRS "responder" is defined as having at least a 50% reduction from baseline AISRS total score. For each subject, the percent reduction will be calculated as defined in Section 4. The number and percentage of 50% AISRS responders along with 95% CIs will be summarized by treatment group and study visit. A 2-sided Pearson's Chi-Square Test, or Fisher's Exact Test, as appropriate, will be performed at the 5% significance level to determine whether there is a significant difference between treatment groups. The null hypothesis is that there is no difference in proportion of 50% AISRS responders between SPN-812 and placebo, with a 2-sided alternative considering a difference in either direction. The difference in proportions between SPN-812 and placebo (SPN-812 minus placebo), 95% CIs for the differences, and *P*-values for the differences in treatment will also be presented.

7) Analysis of 30% AISRS responder rates, defined as having at least a 30% reduction from baseline AISRS total score, will mimic that in number 6 above.

8) With respect to BRIEF-A GEC T-score, data are collected at baseline and Week 6. Hence, MMRM is not applicable since there is only one post randomization value to be analyzed. The change from baseline in BRIEF-A GEC T-score to Week 6 will be analyzed using ANCOVA model with change from baseline in BRIEF-A GEC T-score as dependent variable and baseline BRIEF-A GEC T-score and treatment as fixed independent variables.

9) The change from baseline in the BRIEF-A BRI, MI, and each individual BRIEF-A subscale (BRIEF-A -Inhibit, BRIEF-A -Self-Monitor, BRIEF-A -Plan/Organize, BRIEF-A -Shift, BRIEF-A -Initiate, BRIEF-A -Task Monitor, BRIEF-A -Emotional Control, BRIEF-A -Working Memory, and BRIEF-A -Organization of Materials) will be analyzed using the same ANCOVA model in number 8 above.

[redacted]

6.15. Supplementary Analyses

The primary and key secondary analyses will be repeated in the PP Population.

7. Safety and Tolerability Analysis

Safety assessments include monitoring, evaluation, and recording of all concomitant medications, and the evaluation of AEs, clinical laboratory test results, vital signs, weight, 12-lead ECGs, C-SSRS, and the performance of physical examinations.

All safety analyses will be performed on the Safety population. No inferential statistical tests will be performed.

7.1. Adverse Events

All AEs, TEAEs, and SAEs will be coded using the MedDRA Version 22.0.

A treatment-emergent adverse event (TEAE) is defined as an AE with a start date on or after the date of the first dose of study medication, or that worsened in severity following the first dose of study medication. All AEs in this study will be recorded after administration of study medication, therefore all will be considered treatment-emergent.

The causal relationship of the AE to the study medication is determined by the investigator as Not Related, Unlikely Related, Possibly Related, and Definitely Related. These will be mapped to Unrelated (*Not Related or Unlikely Related*) and Related *(Possibly Related or Definitely Related*).

Adverse event severity grades are reported as mild, moderate, or severe.

Summaries of incidence rates (frequencies and percentages) of individual TEAEs will be presented by SOC, PT, and treatment group. Such summaries will be displayed for all TEAEs, TEAEs by maximum severity, and TEAEs by relationship.

Each subject will be counted only once within each summation level (SOC and PT). If a subject experiences more than 1 TEAE within each summation level, the TEAE with the strongest relationship or the maximum severity, as appropriate, will be included in the summaries of relationship and severity. If a particular event is missing the severity and/or relationship, then the strongest possible severity or relationship will be assumed for analysis (severity = severe, relationship = definitely related).

Incidences will be presented by descending frequency of SOC and PT within SOC, and then alphabetically within PT where the incidence is the same; this is based on overall subjects then alphabetically in case of a tie.

In addition, number and percent of patients reporting common AEs (≥5% in any treatment group) will be presented by PT.

Missing and partially missing AE start and/or stop dates and times will be imputed, for the purpose of statistical analysis, according to the specifications described in Section 6.7.2.

In the AE data listings, all AEs will be displayed. AEs that are treatment emergent will be flagged.

7.1.1. Adverse Events Leading to Withdrawal

A summary of incidence rates (frequencies and percentages) of TEAEs leading to withdrawal of study medication, by treatment group, SOC, and PT will be prepared for the Safety Population.

A data listing of AEs leading to withdrawal of study medication will also be provided, displaying details of the event(s) captured on the CRF.

7.1.2. Deaths and Serious Adverse Events

Any deaths that occur during the study will be listed.

Serious adverse events will be listed and also tabulated by SOC and PT and presented by treatment.

7.2. Clinical Laboratory Evaluations

Laboratory tests (clinical chemistry, hematology, and urinalysis) will be performed at Screening, Baseline, and Week 6 and results will be summarized descriptively by treatment group and study visit as both observed values and change from baseline values. See Section 4 for the definition of baseline. Number and percentage of subjects with abnormal qualitative urinalysis results will be provided by treatment group and study visit.

Laboratory values for clinical chemistry and hematology will be presented at Week 6 using descriptive statistics (n, mean, SD, median, minimum, and maximum) by treatment group for the actual and change from baseline values will be presented in Tables.

Laboratory values for clinical chemistry and hematology will be flagged as abnormally low (L) if the value < lower limit of the normal range, normal (N) if the value is within normal range or abnormally high (H) if the value > upper limit of the normal range. In addition, shift tables for the change from baseline to Week 6 (EOS) will be presented.

Laboratory values will be displayed in the data listings and those that are outside the normal range will be flagged and presented along with corresponding normal ranges (if available). A separate listing of abnormal laboratory values will be provided. All study visits within a parameter for a subject will be presented if at least 1 study visit within that parameter has an abnormal result.

Serology, follicle stimulating hormone (FSH), serum drug screen results, standard urine drug screen results, point of care urine drug screen results, and pregnancy test results will be listed separately.

7.3. Vital Signs

Vital signs will be collected at all study visits. Descriptive summaries of actual values and changes from baseline will be calculated for weight (kg), BMI (kg/m^2), oral body temperature (ºC), respiration rate (breaths per minute), sitting heart rate (bpm), standing heart rate (bpm), sitting systolic blood pressure (mmHg), standing systolic blood pressure (mmHg), sitting diastolic blood pressure (mmHg), and standing diastolic blood pressure (mmHg). These summaries will be presented by study visit and treatment group. See Section 4 for the definition of baseline.

The number of subjects with vital signs below, within, or above normal ranges, by study visit will be tabulated (shift tables) for each parameter by treatment group.

The number of subjects with vital signs values below, within, or above normal ranges, by study visit will be tabulated (shift tables) for each vital sign parameter by treatment group. Normal ranges are presented in Table 3.

Table 3: Vital Sign Normal Ranges

Measurement	Normal Range
Body Mass Index	18 – 35 kg/m^2
Temperature	95-100 °F (35-37.8 °C)
Diastolic blood pressure	60 – 90 mmHg
Systolic blood pressure	90 – 140 mmHg
Heart rate	50 – 100 bpm
Respiration rate	10 – 25 breaths per minute

7.4. Electrocardiograms

12-Lead ECGs will be collected at Screening, Baseline, and Week 6. Descriptive summaries will be presented for heart rate (bpm), PR interval (msec), QRS duration (msec), uncorrected QT interval (msec), and QTcF (msec). These summaries will be presented by study visit and treatment group.

The number of subjects with ECG results below, within, or above normal ranges, by study visit will be tabulated (shift tables) for each parameter by treatment group.

Normal, abnormal but not clinically significant, and abnormal and clinically significant ECG investigator interpretation results will flagged in the listings.

Additionally, the number and percentage of subjects with actual QT and QTcF values ≤450 msec, 450 msec to ≤480 msec, 480 to ≤ 500 msec, >500 msec and with changes from baseline <= 30 msec, 30 msec to ≤ 60 msec, and >60 msec will be presented.

7.5. Physical Examination

The physical examination conducted at screening will include assessments of all body systems except genitourinary. Any findings during screening will be recorded as medical history and any clinically significant abnormal findings during treatment will be recorded as an AE. At the EOS physical examination, only changes from the screening visit will be noted.

All physical examination results will be listed.

7.6. Columbia Suicide Severity Rating Scale (C-SSRS)

Assessment of suicidal ideation and behavior will be conducted using the Columbia-Suicide Severity Rating Scale (C-SSRS) which classifies suicidal ideation and behavior events into 11 preferred categories including 5 levels of suicidal ideation, 5 levels of suicidal behavior, and the category of self-injurious behaviors with no suicidal intent. The C-SSRS will be performed at all study visits. At the Screening visit, the "baseline" version of the C-SSRS will be administered. This version assesses Suicidal Ideation and Suicidal Behavior during the subject's lifetime and during the past 6 months. At the subsequent study visits, the "since last visit" version will be administered.

The number and percentages of subjects with a response of "Yes" at any point on the suicidal ideation only, suicidal behavior only, and suicidality (ideation and behavior combined) items will be summarized by treatment group.

7.7. Concomitant Medications

Prior and concomitant medications, coded using World Health Organization-Drug Dictionary Enhanced (WHO-DDE) (March 2019), will be summarized descriptively by Anatomical Therapeutic Chemical (ATC) classification Level 4 and PT (i.e., ATC classification Level 5), if applicable, using counts and percentages for the Safety Population.

If a medication starts prior to the first dose and continues after the first dose it will be considered both prior and concomitant. Prior and concomitant medications will also be listed. Prior and concomitant medications will be summarized descriptively by treatment using counts and percentages.

7.8. Caffeine Consumption

Caffeine consumption will be assessed at all study visits and listed.

7.9. Other Safety Assessments

Follow-up phone call data and subject training video module data will be listed.

8. Changes from Planned Analysis

A Mixed Model for Repeated Measure (MMRM) will be used instead of an ANCOVA model for the multiply imputed data using SAS PROC MI for primary, secondary, [redacted] efficacy

endpoints except for BRIEF-A, [redacted]. Since BRIEF-A, [redacted] data are only collected at baseline and Week 6 or ANCOVA analysis will be used as per the protocol.

9. Validation of Analysis Reports

Supernus seeks to ensure the quality of the reports provided by Premier in the form of TLFs and derived datasets must pass a rigorous validation process involving the following processes.

- Derived datasets must be independently reprogrammed by a second programmer based on analysis data specifications. The separate datasets produced by the 2 programmers must match 100% or in case of any differences, an explanation should be given and documented.
- 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 above validation process must be repeated any time TLFs are redelivered using updated or different data. Execution of this validation process must be documented through the study in a tracking sheet which contains TLF number and title, name of the production programmer, name of the QC programmer and comments. The validation tracking sheet must be submitted at each delivery of TFLs.

10. References

1. US Federal Register. (1998) International Conference on Harmonization; Guidance on Statistical Principles for Clinical Trials. Department of Health and Human Services: Food and Drug Administration [Docket No. 97D-0174]. Federal Register Volume 63, Number 179, pages 49583-49598. September 16, 1998.
2. ASA. (2016) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2016. http://www.amstat.org/about/ethicalguidelines.cfm
3. RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014. http://www.rss.org.uk/Images/PDF/join-us/RSS-Code-of-Conduct-2014.pdf.
4. Spencer TJ, Adler LA, Meihua Q, Saylor KE, Brown TE, Holdnack JA, Schuh KJ, Trzepacz PT, Kelsey DK. Validation of the adult ADHD investigator symptom rating scale (AISRS). J Atten Disord. 2010;14(1):57-68.
5. Guy W. *Clinical global impressions*. *ECDEU Assessment manual for Psychopharmacology*, Rockville, MD: US National Institutes of Health, Psychopharmacology Research Branch, 1976.
6. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10):1092-1097.
7. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

8. Glass GV, McGaw B, Smith ML (1981). Meta-Analysis in Social Research. Sage: Beverly Hills.

11. Planned Table, Figure, and Listing Descriptions

The planned summary tables, figures, and listings for protocol number 812P306 are contained in a separate appendix to this SAP.