

Table of Contents

Table of Contents.....	1
Statistical Analysis Plan - 26May2023.....	2

CCI



Statistical Analysis Plan: 18-AVP-786-207 Study for Negative Symptoms of Schizophrenia

Study Title: A MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-ARM STUDY TO ASSESS THE EFFICACY, SAFETY, AND TOLERABILITY OF AVP-786 (DEUDEXTROMETHORPHAN HYDROBROMIDE [D6-DM]/QUINIDINE SULFATE [Q]) FOR THE TREATMENT OF NEGATIVE SYMPTOMS OF SCHIZOPHRENIA

Study Number: 18-AVP-786-207

Study Phase: Phase 2/3

Sponsor: Otsuka Pharmaceutical Development & Commercialization, Inc.

Version: 2.0

Date: 26 May 2023

Confidentiality Statement

This document is a confidential and proprietary communication of Otsuka Pharmaceutical Development & Commercialization, Inc. (OPDC) and it is for the sole information and use of the investigator's research team and the appointed Institutional Review Boards/Independent Ethics Committees.

Statistical Analysis Plan: 18-AVP-786-207

1 TABLE OF CONTENTS

1	TABLE OF CONTENTS.....	2
	LIST OF TABLES.....	4
	LIST OF FIGURES	4
	LIST OF APPENDICES.....	4
2	SIGNATURE PAGE	5
3	LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS.....	6
4	INTRODUCTION	9
5	STUDY OBJECTIVES.....	9
5.1	Primary Objectives.....	9
6	STUDY OVERVIEW	9
6.1	Study Design.....	9
	CCI	
6.3	Data Monitoring.....	11
6.4	Sample Size Estimation and Power Justification.....	11
7	STUDY ENDPOINTS	11
7.1	Scales and Questionnaires.....	11
7.1.1	Positive and Negative Syndrome Scale (PANSS)	11
7.1.2	16-Item Negative Symptom Assessments (NSA-16)	12
7.1.3	Patient Global Impression – Severity (PGI-S).....	14
7.1.4	Patient Global Impression – Change (PGI-C)	14
7.1.5	Calgary Depression Scale for Schizophrenia (CDSS)	14
7.1.6	Columbia Suicide Severity Rating Scale (C-SSRS)	15
7.1.7	Abnormal Involuntary Movements Scale (AIMS)	15
7.1.8	Barnes Akathisia Scale (BAS)	15
7.1.9	Simpson Angus Scale for Extrapyramidal Symptoms (SAS).....	16
7.1.10	Mini International Neuropsychiatric Interview (M.I.N.I.) Version 7.0.2	16
7.2	Efficacy Endpoints.....	17
7.3	Safety Endpoints	17
7.4	Pharmacokinetic, Genotype and Biomarker Parameters	17
8	ANALYSIS POPULATIONS	17
8.1	Safety Population.....	18
8.2	Modified Intent-to-Treat (mITT) Population.....	18
9	OVERALL STATISTICAL CONSIDERATIONS.....	18
9.1	Placebo Run-in Period and Study Treatment Period	18
9.2	Study Baseline	18
9.3	End of Treatment	18
9.4	Change from Baseline.....	18

Statistical Analysis Plan: 18-AVP-786-207

9.5	Visit Windows	18
9.6	Handling of Multiple Observations	19
9.7	Handling of Missing Data or Outliers, Missing or Partial Dates	19
9.8	Treatment Misallocations.....	21
9.9	Summary Statistics.....	21
9.10	Data Listings	21
9.11	Graphical Displays.....	21
9.12	Hypothesis Testing.....	21
9.13	Multiplicity of Secondary Endpoints	21
10	EFFICACY ANALYSIS	21
10.1	Primary Estimand.....	21
10.2	Derivation of Primary Efficacy Endpoints	22
10.3	Primary Efficacy: MMRM (Mixed-Model Repeated Measures).....	23
	CCI	
10.6	Subgroup Analyses for Primary Efficacy Endpoint.....	26
10.6.1	Subgroup Analyses by Age, Gender, Race, Country, Region and Years Since Onset of Schizophrenia	26
10.6.2	Subgroup Analyses by Baseline Concomitant Medication According to CYP2D6 Major Substrate Status	27
10.6.3	Subgroup Analyses by Baseline Concomitant Antipsychotic Medications	27
10.7	Secondary Efficacy and Other Outcome Measures Analysis	28
10.7.1	Secondary Efficacy Analyses	28
10.7.2	Other Efficacy Analyses	28
10.8	SAFETY ANALYSIS.....	28
10.9	Adverse Events	28
10.10	Clinical Laboratory Assessments.....	30
10.11	ECGs	31
10.12	Vital Signs.....	33
10.13	Physical and Neurological Exams	33
10.14	C-SSRS	33
10.15	Simpson Angus Scale for Extrapyrarnidal Symptoms (SAS).....	34
10.16	Barnes Akathisia Scale (BAS).....	34
10.17	Abnormal Involuntary Movement Scale (AIMS).....	34
11	ADDITIONAL STATISTICAL ANALYSIS METHODS	35
11.1	Patient Disposition	35
11.2	Demographics and Baseline Characteristics	35
11.3	Medical History and Disease Characteristics	35

Statistical Analysis Plan: 18-AVP-786-207

11.4	Exposure and Treatment Compliance	36
11.5	Concomitant Medications	36
11.6	Analysis of Pharmacokinetics	37
12	ADDITIONAL STATISTICAL CONSIDERATIONS	37
12.1	Data Monitoring Committee (DMC)	37
13	REFERENCES	38
14	APPENDICES	39

LIST OF TABLES

Table 1	DMC Decisions Rule	10
Table 2	Study Day and Visit Window	19
Table 3	LAB PCS Criteria	31
Table 4	ECG PCS Criteria	32
Table 5	Vital Sign PCS Criteria	33

LIST OF FIGURES**LIST OF APPENDICES**

Appendix 1	Schedule of Assessments and Procedures	40
------------	--	----

Statistical Analysis Plan: 18-AVP-786-207

2 SIGNATURE PAGE

Study Title: A Multicenter, Randomized, Double-blind, Placebo-controlled, Parallel-arm Study to Assess the Efficacy, Safety, and Tolerability of AVP-786 (Deudextromethorphan Hydrobromide [d6-DM]/Quinidine Sulfate [Q]) for the Treatment of Negative Symptoms of Schizophrenia

Study Number: 18-AVP-786-207

Prepared by: _____ Date: _____
PPD
Statistician, Statistics
Pharmapace, Inc., a WuXi AppTec Company

Prepared by: _____ Date: _____
PPD
Associate Director, Biostatistics
Otsuka Pharmaceuticals Development & Commercialization, Inc.

Approved by: _____ Date: _____
PPD
Senior Director, Biostatistics
Data & Analytics
Otsuka Pharmaceuticals Development & Commercialization, Inc.

Approved by: _____ Date: _____
PPD
Director, Global Clinical Development
Otsuka Pharmaceuticals Development & Commercialization, Inc.

Statistical Analysis Plan: 18-AVP-786-207

3 LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

AE	adverse event
AIC	Akaike information criterion
AIMS	Abnormal Involuntary Movements Scale
ALT/SGPT	alanine aminotransferase/serum glutamic-pyruvic transaminase
AR1	first-order autoregressive
AST/SGOT	aspartate aminotransferase/serum glutamic-oxaloacetic transaminase
ATC	anatomical therapeutic chemical
AVP-786	deudextromethorphan hydrobromide and quinidine sulfate combination
BAS	Barnes Akathisia Scale
BID	twice-daily
BP	blood pressure
BUN	blood urea nitrogen
CDSS	Calgary Depression Scale for Schizophrenia
CFBL	change from baseline
CI	confidence interval
CK	creatinine kinase
CMH	Cochran-Mantel-Haenszel
CCI	
CRF	case report form
CS	compound symmetry
C-SSRS	Columbia Suicide Severity Rating Scale
CYP2D6	cytochrome P450 isoenzyme 2D6
d6-DM	deuterated (d6)-dextromethorphan (hydrobromide)
DBP	diastolic blood pressure
DM	dextromethorphan (hydrobromide)
DMC	Data Monitoring Committee
ECG	electrocardiogram
ET	early termination
GGT	gamma-glutamyl transferase
Ham-D	Hamilton Depression Scale
HBA1c	glycosylated hemoglobin
HR	heart rate
IA	interim analysis
IAP	interim analysis plan
ICH	International Conference on Harmonisation
IMP	investigational medicinal product
ITT	Intent-to-Treat
LDH	lactate dehydrogenase
LLOQ	lower limit of quantification

Statistical Analysis Plan: 18-AVP-786-207

LOCF	last observation carried forward
LOE	lack of efficacy
LS	least square
MAR	missing at random
MCAR	missing completely at random
MCMC	Monte Carlo Markov Chain
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram(s)
MI	multiple imputation
M.I.N.I.	Mini International Neuropsychiatric Interview
mITT	Modified Intent-to-Treat
MMRM	mixed-effects model for repeated measures
MNAR	missing not at random
NIMH	National Institute of Mental Health
NSA-16	16-Item Negative Symptom Assessment
OC	observed case
OTC	opportunity to complete
PANSS	Positive and Negative Syndrome Scale
PCS	potentially clinically significant
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	pharmacokinetic(s)
PMM	pattern mixture model
PT	preferred term
Q	quinidine (sulfate)
QAM	every morning
QPM	every evening
QRS	Q-R-S complex from an ECG tracing
QT	QT interval from an ECG tracing
QTcF	QT interval corrected for heart rate using the Fridericia's formula
RBC	red blood cell
SAE	serious adverse event
SAS	Simpson Angus Scale for Extrapyrarnidal Symptoms
SBP	systolic blood pressure
SNRI	serotonin-norepinephrine reuptake inhibitor
SOC	system organ class
SSRI	selective serotonin reuptake inhibitor
TEAE	treatment-emergent adverse event
UN	unstructured covariance
WBC	white blood cell

Statistical Analysis Plan: 18-AVP-786-207

WHODrug World Health Organization-Drug

Statistical Analysis Plan: 18-AVP-786-207

4 INTRODUCTION

This document provides a detailed description of the statistical methods and procedures to be implemented during the analysis of Otsuka Pharmaceuticals Development & Commercialization, Inc. Protocol 18-AVP-786-207. All amendments to the protocol and Addendum to protocol amendment are taken into consideration in developing this SAP.

5 STUDY OBJECTIVES

The objectives of this 12-week study are to evaluate the efficacy, safety, and tolerability of AVP-786, a combination of deuterated-dextromethorphan hydrobromide (d6-DM) and quinidine sulfate (Q), as a therapy compared with placebo in patients with negative symptoms of schizophrenia.

5.1 Primary Objectives

The primary objective of the study is to evaluate the efficacy, safety, and tolerability of AVP-786, as compared with placebo, for the treatment of negative symptoms of schizophrenia.

6 STUDY OVERVIEW

6.1 Study Design

This is a Phase 2/3, multicenter, randomized, double-blind, placebo-controlled study to assess the efficacy, safety, and tolerability of AVP-786 for the treatment of negative symptoms of schizophrenia.

The study consists of an up to 4-week Screening period, a 3-week double-blind placebo run-in period (Phase A), a 12-week double-blind treatment period (Phase B), and a 30-day follow-up period. After Screening, eligible patients will enter a 3-week double-blind placebo run-in period. After completion of the placebo run-in period, patients will be randomized in a 1:1 ratio stratified by placebo responder status and trial center to receive either AVP-786 or matching placebo capsules. Placebo run-in responders, are defined as all patients who demonstrate a 20% or greater reduction in the Positive and Negative Syndrome Scale (PANSS) Marder negative factors score during the run-in period.

After the end of the double-blind placebo run-in period, patients randomized to AVP-786 will begin a one-week titration period in which they will receive a dose of d6-DM 28 mg/Q 4.9 mg (AVP-786-28/4.9) every morning (qam) and a dose of placebo every evening (qpm) for the first 3 days, followed by AVP-786-28/4.9 twice daily (BID) for the next 4 days. Following this one-week titration period (Day 8 of the double-blind treatment period), patients will receive d6-DM 42.63 mg/Q 4.9 mg (AVP-786-42.63/4.9) administered orally BID for the remaining 11 weeks of treatment. Patients randomized to placebo group will receive the matching AVP-786 placebo capsule BID during the 12-week double blind treatment period beginning on Day 1 of the double-blind treatment period.

Statistical Analysis Plan: 18-AVP-786-207

This trial will be monitored by an independent Data Monitoring Committee (DMC). The DMC will periodically monitor safety based on a predetermined schedule. In addition, an interim analysis (IA) of efficacy data is planned during the course of the trial, and will be reviewed by the independent DMC. The sponsor will remain blinded to the observed results of the IA, and will only receive recommendations as per the interim analysis plan (IAP) and DMC Charter.

CCI



Statistical Analysis Plan: 18-AVP-786-207

CCI

6.3 Data Monitoring

An independent DMC is formed to independently monitor patients' safety and efficacy during the trial. The DMC will meet to review safety data periodically. The DMC will review the un-blinded efficacy data for the planned IA. The details of the DMC structure and its roles and responsibilities are documented in a Data Monitoring Committee Charter.

6.4 Sample Size Estimation and Power Justification

The initial sample size was calculated based on the standardized effect size of 0.40 in the change from baseline during the 12-week treatment period in the PANSS Marder negative factors score, to achieve 85% power at a 2-sided alpha level of 0.05. The resulting sample size is about 230 total patients (placebo run-in non-responders) randomized in a 1:1 ratio to AVP-786 or placebo. It is expected that approximately 15% of patients will be placebo run-in responders and excluded from the analysis. A total of 270 patients are expected to be enrolled at the start of the placebo run-in period to get 230 placebo run-in non-responders.

The statistical assumption of the standardized effect size of 0.40 was based on post-hoc analyses of data obtained from the completed phase 2 trial 15-AVP-786-202 that assessed the efficacy, safety and tolerability of AVP-786 as an adjunctive treatment for patients with residual schizophrenia. CCI

7 STUDY ENDPOINTS

7.1 Scales and Questionnaires

7.1.1 Positive and Negative Syndrome Scale (PANSS)

The PANSS is a 30-item clinical scale to assess positive and negative symptoms of schizophrenia, and will be performed at Screening (Day -28 to -1), Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 3 (Day 43), Visit 4 (Day 64), Visit 5 (Day 85), and Visit 6/Early Termination (Day 106).

Each item is rated on the following scale:

1 = Absent.

Statistical Analysis Plan: 18-AVP-786-207

- 2 = **Minimal**: questionable or subtle or suspected pathology, or may allude to the extreme end of the normal range.
- 3 = **Mild**: presence is clearly established but not pronounced, and interferes little in day-to-day functioning.
- 4 = **Moderate**: characterizes a symptom which, though representing a serious problem, either occurs only occasionally or intrudes on daily life only to a moderate extent.
- 5 = **Moderate to Severe**: indicates marked manifestations that distinct impact on one's functioning but are not all-consuming and usually can be contained at will.
- 6 = **Severe**: represents gross pathology that is present very frequently, proves highly disruptive to one's life, and often calls for direct supervision.
- 7 = **Extremely Severe**: refers to the most serious level of psychopathology, whereby the manifestations drastically interfere in most or all major life functions, typically necessitating close supervision and assistance in many areas.

Therefore, the PANSS total score range is 30 to 210. Items can be grouped into the following subscales with the corresponding score ranges:

Subscale	Item Numbers	Score Range
Marder Negative Factors	N1 – N4, N6, G7, G16	7 – 49
Negative Subscale	N1 – N7	7 – 49
Positive Subscale	P1 – P7	7 – 49
General Psychopathology Subscale	G1 – G16	16 – 112
Prosocial Factors	G16, N2, N4, N7, P3, P6	6 – 42
Excitement Component	P4, P7, G4, G8, G14	5 – 35

The PANSS Marder negative factors score is the primary efficacy endpoint, which is comprised of the following 7 items of the 30-item PANSS: N1 – Blunted affect, N2 – Emotional withdrawal, N3 – Poor rapport, N4 – Passive/apathetic social withdrawal, N6 – Lack of spontaneity and flow of conversation, G7 – Motor retardation, G16 – Active social avoidance.

PANSS Positive subscale score is the secondary efficacy endpoint.

7.1.2 16-Item Negative Symptom Assessments (NSA-16)

The NSA-16 is considered a valid and reliable measure of the presence, severity, and range of negative symptoms associated with schizophrenia. The NSA-16 will be performed at Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 3 (Day 43), Visit 4 (Day 64), Visit 5 (Day 85), and Visit 6/Early Termination (Week Day 106).

The NSA-16 consists of 16 items, and uses a 5-factor model to describe negative symptoms:

Factor Domain	NSA-16 Item Numbers	Score Range
Communication	1, 2, 3, 4	4 – 24

Statistical Analysis Plan: 18-AVP-786-207

Emotion/Affect	5, 6, 7	3 – 18
Social Involvement	8, 9, 10	3 – 18
Motivation	11, 12, 13, 14	4 – 24
Retardation	15, 16	3 – 18
Total	1 – 16	16 – 96

Each item is rated on a 6-point Likert scale from 1 to 6, with detailed descriptions of anchors at each point on the scale. The severity scores can be interpreted as follows:

- 1 = The behavior being assessed is not reduced or absent as compared to a healthy young human.
- 2 = The behavior being assessed is minimally reduced, significance is questionable.
- 3 = The behavior being assessed is mildly reduced; it might only be noted as reduced by a trained rater, but he/she notes a definite reduction.
- 4 = The behavior being assessed is moderately reduced; the reduction should be obvious to an untrained rater.
- 5 = The behavior being rated is markedly reduced; this behavior is easily observable and definitely interferes with the patient's functioning.
- 6 = The behavior being assessed is severely reduced or entirely absent; it is glaring and markedly interferes with functioning.

For each factor domain, scores will be calculated only if all items in that domain are rated; otherwise, it will be regarded as missing.

Furthermore, each item admits a score of 9 if the item is not ratable. The NSA-16 total score is defined as the sum of all 16 items excluding any items receiving a score of 9. If more than 3 items are not ratable, the total score cannot be calculated and will be regarded as missing; otherwise, the total score will be calculated as 16 times the arithmetic mean of the scores of all ratable items, rounded to the nearest tenth. Therefore, the minimum score is 16 and the maximum score is 96, with a higher score indicating greater clinical severity of symptoms. Additionally, two global ratings are included: a global negative symptoms rating, which is scored on the basis of the overall impression of negative symptoms in the patient; and a global level of functioning, which represents the clinician's overall assessment of the patient's level of functioning, relative to a healthy young adult. Both are scored on the following scale:

- 1 = No evidence of this symptom/impairment
- 2 = Minimal evidence of this symptom/impairment
- 3 = Mild evidence of this symptom/impairment
- 4 = Moderate evidence of this symptom/impairment, apparent to the casual observer
- 5 = Marked evidence of this symptom/impairment, readily apparent to the casual observer
- 6 = Severe symptom/impairment, not only obvious but has marked impact on functioning
- 7 = Extremely severe symptom/impairment that is incapacitating for the patient

Missing global rating will not be imputed.

Statistical Analysis Plan: 18-AVP-786-207

The NSA-16 global negative symptom score is the secondary efficacy endpoint, while the NSA-16 total score is not an efficacy endpoint in the study.

7.1.3 Patient Global Impression – Severity (PGI-S)

The PGI-S is a 7-point (1–7), patient-rated scale used to assess the severity of the patients' illness as follows:

- 1 = Normal, not at all ill
- 2 = Borderline ill
- 3 = Mildly ill
- 4 = Moderately ill
- 5 = Markedly ill
- 6 = Severely ill
- 7 = Extremely ill

The PGI-S evaluation will be performed at Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 3 (Day 43), Visit 4 (Day 64), Visit 5 (Day 85), and Visit 6/Early Termination Visit (Day 106), and will be focused on the negative symptoms of schizophrenia. Missing PGI-S rating will not be imputed. The PGI-S is the secondary efficacy endpoint.

7.1.4 Patient Global Impression – Change (PGI-C)

The PGI-C is a 7-point (1–7), patient-rated scale used to assess treatment response as follows:

- 1 = Very much improved
- 2 = Much improved
- 3 = Minimally improved
- 4 = No Change
- 5 = Minimally worse
- 6 = Much worse
- 7 = Very much worse

The PGI-C evaluation will be performed at Baseline/Visit 2 (Day 22), Visit 3 (Day 43), Visit 4 (Day 64), Visit 5 (Day 85), and Visit 6/Early Termination Visit (Day 106), and will be focused on the negative symptoms of schizophrenia. Missing PGI-C rating will not be imputed. The PGI-C is the secondary efficacy endpoint.

7.1.5 Calgary Depression Scale for Schizophrenia (CDSS)

The CDSS is a 9-item scale derived from the Hamilton Depression Scale (Ham-D) that is designed to assess depression specifically in patients with schizophrenia. Unlike the Ham-D, the CDSS does not contain depressive symptoms that overlap with negative symptoms of

Statistical Analysis Plan: 18-AVP-786-207

schizophrenia, such as anhedonia and social withdrawal. The CDSS has shown excellent psychometric properties. Each item on the scale is scored as 0, Absent; 1, Mild; 2, Moderate; or 3, Severe. The CDSS score is obtained by adding each of the item scores. A score above 6 has an 82% specificity and 85% sensitivity for predicting the presence of a major depressive episode. Missing CDSS rating will not be imputed.

The CDSS evaluation will be performed at Screening (Day -28 to -1), Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 4 (Day 64), and Visit 6/Early Termination Visit (Day 106). The CDSS will be used as the other outcome measure for evaluation.

7.1.6 Columbia Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is a low-burden measure of the spectrum of suicidal ideation and behavior that was developed by Columbia University researchers for the National Institute of Mental Health (NIMH) Treatment of Adolescent Suicide Attempters Study to assess severity and track suicidal events through any treatment. It is a clinical interview providing a summary of both ideation and behavior that can be administered during any evaluation or risk assessment to identify the level and type of suicidality present. The C-SSRS can also be used during treatment to monitor for clinical worsening.

The C-SSRS evaluation will be performed at Screening (Day -28 to -1), Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 3 (Day 43), Visit 4 (Day 64), Visit 5 (Day 85), and Visit 6/Early Termination Visit (Day 106). The C-SSRS will be used to assess safety.

7.1.7 Abnormal Involuntary Movements Scale (AIMS)

The AIMS is a 12-item scale to assess dyskinesia. Items related to severity of orofacial, extremity, and trunk movements, global judgment about incapacitation, and patient awareness are rated using a 5-point scale (0 = None, 1 = Minimal, 2 = Mild, 3 = Moderate, 4 = Severe). Two items related to dental status are scored using “yes/no” responses. Missing AIMS response will not be imputed.

The AIMS evaluation will be performed at Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 4 (Day 64), and Visit 6/Early Termination Visit (Day 106). The AIMS will be used to assess safety.

7.1.8 Barnes Akathisia Scale (BAS)

The BAS is an assessment of the presence and frequency of akathisia (a movement disorder characterized by restlessness and a compulsion to move), the level of an individual’s subjective awareness and distress, and global severity. The objective item of the scale measures the clinician’s assessment of the patient’s restlessness. The subjective portion of the scale measures both the patient’s awareness of restlessness and the distress arising from such awareness. Each item is rated on a 4-point scale from 0 to 3, and the total score is the sum of the three item scores, thus yielding a total score ranging from 0 to 9. The Global Clinical Assessment of Akathisia is rated on a 5-point scale from 0 to 4. Higher scores indicate greater severity of symptoms. Missing BAS rating will not be imputed.

Statistical Analysis Plan: 18-AVP-786-207

The BAS evaluation will be performed at Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 4 (Day 64), and Visit 6/Early Termination Visit (Day 106). The BAS will be used to assess safety.

7.1.9 Simpson Angus Scale for Extrapyraxidal Symptoms (SAS)

The SAS is composed of 10 items and is used to assess pseudoparkinsonism. Grade of severity of each item is rated using a 5-point scale. SAS scores can range from 0 to 40. Signs assessed include gait, arm-dropping, shoulder-shaking, elbow rigidity, wrist rigidity, leg pendulousness, head rotation, glabella tap, tremor, and salivation. Missing SAS rating will not be imputed.

The SAS evaluation will be performed at Screening (Day -28 to -1), Visit 1 (Day 1), Baseline/Visit 2 (Day 22), Visit 4 (Day 64), and Visit 6/Early Termination Visit (Day 106). The SAS will be used to assess safety.

7.1.10 Mini International Neuropsychiatric Interview (M.I.N.I.) Version 7.0.2

The M.I.N.I. is an interview developed by psychiatrists and clinicians for psychiatric disorders and has an administration time of approximately 15 minutes.

The interview is divided into 15 categories, which are listed below:

- Major depressive episode
- Suicidality
- Manic and hypomanic episodes
- Panic disorder
- Agoraphobia
- Social phobia
- Obsessive-compulsive disorder
- Posttraumatic stress disorder
- Alcohol dependence/abuse
- Substance dependence/abuse (non-alcohol)
- Psychotic disorders and mood disorder with psychotic features
- Anorexia nervosa
- Bulimia nervosa
- Generalized anxiety disorder
- Antisocial personality disorder

Based on the answers, each category will be classified into a timeframe (e.g., current, past, etc.). Missing values will not be imputed. The M.I.N.I. is performed at Screening for the purpose of assessing whether the patient meets diagnostic criteria for schizophrenia type. Results will be presented in the data listings.

Statistical Analysis Plan: 18-AVP-786-207

7.2 Efficacy Endpoints

The primary efficacy endpoint is the change from baseline (Week 3) to Week 15 in PANSS Marder negative factors score.

Secondary endpoints are the change from baseline (Week 3) to Week 15 in the scores of the assessments listed below:

- NSA-16 Global Negative Symptom Score
- PGI-S
- PGI-C

Other efficacy endpoints to be analyzed:

- PANSS positive subscale
- CDSS

7.3 Safety Endpoints

Safety measurements will include: adverse events (AEs), serious adverse events (SAEs), physical and neurological examinations, vital signs, weight, pregnancy tests, clinical laboratory assessments, resting 12-lead electrocardiograms (ECGs), Columbia Suicide Severity Rating Scale (C-SSRS), Abnormal Involuntary Movements Scale (AIMS), Barnes Akathisia Scale (BAS), and Simpson Angus Scale for Extrapyrimalal Symptoms (SAS).

7.4 Pharmacokinetic, Genotype and Biomarker Parameters

Plasma concentrations of d6-DM, its metabolites (d3-DX and d3-3-MM), and Q will be measured from samples collected post-dose on Day 22 (Week 3), pre- and post-dose on Day 64 (Week 9), and pre-dose on Day 106 (Week 15). These samples will be collected per instructions provided by the sponsor. Descriptive summary statistics will be provided.

Blood sample for CYP2D6 enzyme genotype will be taken at Visit 1 (Day 1), and used to classify patients as either poor, intermediate, extensive, or ultra-rapid metabolizers of d6-DM. Efficacy analyses using the PANSS Marder negative factors score will compare the different groups.

8 ANALYSIS POPULATIONS

A summary table containing the number of patients in each of the populations described below will be provided. Placebo run-in responders, defined as all patients who demonstrate a 20% or greater reduction in the PANSS Marder negative factors score from Visit 1 (Day 1) to Baseline\Visit 2 (Week 3), will be excluded from all efficacy analyses, but will be included in all safety analyses.

Statistical Analysis Plan: 18-AVP-786-207

8.1 Safety Population

Safety population is defined as all patients who are randomized in Phase B and take at least one dose of study drug. All safety analyses will be based on the safety population.

8.2 Modified Intent-to-Treat (mITT) Population

mITT population is defined as all patients who are placebo run-in non-responders, randomized in Phase B and in the Safety Population with Phase B Baseline (Visit 2, Week 3) and at least one post Baseline PANSS measurement. All efficacy analyses will be based on the mITT population.

9 OVERALL STATISTICAL CONSIDERATIONS

9.1 Placebo Run-in Period and Study Treatment Period

Placebo run-in period is from Day 1 to Day 22 (Week 3) and study treatment period is from Day 22 (Week 3) to Day 106 (Week 15). The total duration of study drug treatment is 12 weeks.

9.2 Study Baseline

Because there is a 3 week placebo run-in period, baseline is generally defined as the end of the placebo run-in period (Day 22, Week 3), and to be the last assessment prior to the first dose of study drug.

9.3 End of Treatment

End of treatment is defined as the last value for a given patient during the study treatment period, whenever it occurred.

9.4 Change from Baseline

Change from baseline (CFBL) will always be calculated as (post-baseline – baseline). It will be calculated for patients with both a baseline and post-baseline value as applicable.

If a baseline value has not been recorded for a parameter, then CFBL will not be calculated for that parameter, and the patient will be excluded from CFBL analysis.

Percent CFBL, is the CFBL divided by the baseline value multiplied by 100%. Patients with a value of 0 at baseline cannot have percent CFBL calculated.

9.5 Visit Windows

Study visit windows will be used to map visits using study day intervals. This visit window convention applies to tables and listings for all efficacy (PANSS Marder negative factors score, NSA-16 Global Negative Symptom Score, PGI-S, PGI-C, PANSS positive subscale, CDSS) and safety scales (C-SSRS, AIMS, BAS, SAS). This derived study window variable

Statistical Analysis Plan: 18-AVP-786-207

will be named as WEEK and will be footnoted. In listings it will be listed along with the case report form (CRF) study visit.

Table below (Table 2) shows classifications for study day intervals in the Phase B double-blind treatment period. The variable “target day” is defined using the number of days since the start of Phase B double-blind dosing. The first day of Phase B double-blind dosing is defined as “Day 1”. If more than one observation falls within a particular study day interval, then the last observation within that interval is used. Evaluations occurring more than 3 days after the last double-blind dosing date will not be mapped into study visit windows and will be excluded from the analysis.

Table 2 Study Day and Visit Window

Visit	Week	Target Day ^a	Study Day Interval ^a
	4	7	2-10
	5	14	11-17
3	6	21	18-24
	7	28	25-31
	8	35	32-38
4	9	42	39-45
	10	49	46-52
	11	56	53-59
5	12	63	60-66
	13	70	67-73
	14	77	74-80
6	15	84	81-87 ^b

^a Relative to the first day of double-blind study medication in the double-blind treatment period.

^b Evaluations occurring more than 3 days after the last dosing date of double-blind study medication in the double-blind treatment period will be excluded from the efficacy analyses.

9.6 Handling of Multiple Observations

For multiple post-baseline values associated with the same scheduled visit, mean values will be computed first and will be used as the value associated with that visit. However, when summarizing the laboratory parameter shift tables, the worst value of the multiple observations will be used when applicable.

All results, scheduled or unscheduled, will be listed.

9.7 Handling of Missing Data or Outliers, Missing or Partial Dates

Missing data will be handled differently depending on the parameter and analysis. Note that analyses done on ‘observed cases’ will not follow any imputation rules below. See below for considerations:

- Missing baseline values will not be imputed in any situation.

Statistical Analysis Plan: 18-AVP-786-207

- For any safety assessments of patients without a Week 15 safety endpoint assessment, the last non-missing post-baseline observation will be used as the end of treatment value.
- Missing post-baseline values for by-visit safety data will be summarized using the Visit Windows from [Section 9.5](#). If a value is not available within a given window, no imputation will be done.
- Missing data for AE relationship will be imputed as “Related.”

In general, no imputation of missing data will be done except for those described below:

Values below the lower limit of quantification of an assay (<LLOQ) will be assigned LLOQ/2 for the purposes of analysis; however, for safety laboratory test values that is recorded as “> x”, “>= x”, “< x” or “<= x”, a value of x will be taken for summary; for safety laboratory test values that is recorded as “None Detected”, they will be treated as missing. In listings, these data will be presented as recorded with the sign.

Partial or incomplete start/end dates of the adverse event (AE) or concomitant medications will be imputed to help determine treatment-emergent adverse events and prior or concomitant medications.

For incomplete start dates, the following imputation rules will be implemented:

- Missing day - Impute as the 1st of the month unless month and year are the same as month and year of first dose of study drug then impute using the day of first dose date.
- Missing day and month – impute as 1st January unless year is the same as first dose date then impute using the day and month of first dose date.
- Completely missing – impute using the first dose date unless the end date suggests it could have started prior to this in which case impute as 1st January of the same year as the end date.

When imputing a start date, ensure that the new imputed date is sensible (i.e. is prior or equal to the end date of the AE or medication).

For incomplete end dates following imputation rules will be implemented:

- Missing day - Impute as the last day of the month unless month and year are the same as the month and year of the end of study, then impute as the end of study date.
- Missing day and month – Impute as 31st December unless year is the same as the year of the end of study date then impute as the end of study date.
- Completely Missing – need to look at whether the AE/medication is still ongoing before imputing a date and when it started in relation to study drug. If the ongoing flag is not missing, then assume that AE is still present/medication is still being taken (i.e. do not impute a date). If the AE/medication has stopped and start date is prior to first dose date then impute as the 1st dose date, if it started on or after first dose date then impute as the end of study date.

Statistical Analysis Plan: 18-AVP-786-207

9.8 Treatment Misallocations

Efficacy data will be summarized “as randomized,” meaning the patient will be summarized under the treatment they were randomized to, regardless of what treatment was actually received.

Safety data will be summarized “as treated,” meaning the patient will be summarized based on the treatment that was actually received.

9.9 Summary Statistics

Quantitative displays will be summarized using descriptive statistics. The mean, standard deviation, median, minimum, and maximum will be provided. Decimal precision will be based on the mean value. The median contains the same number of decimal places as the mean, the standard deviation contains one more decimal place, and the minimum and maximum contain one less decimal place. The mean will typically have one more decimal place than the raw values but if necessary, decimal precision for the mean will be provided in the table specifications.

9.10 Data Listings

Unless otherwise specified, data listings will be provided on observed values. Efficacy listings for the primary and secondary endpoints will include analysis flags and/or imputations to show which values were used in the summary tables. In addition, the analysis visit will be displayed if it is important for the corresponding summary. Listings will be provided to serve as support for all summary tables or figures.

9.11 Graphical Displays

Supporting figures may be used for some efficacy, pharmacokinetic (PK) or safety analyses in addition to the summary tables. Details regarding the content, layout, and structure of figures will be provided in the table specifications.

9.12 Hypothesis Testing

A significance level of 0.001 (2-sided) is allocated to the IA. The corresponding final significance level is 0.049 (2-sided) using Bonferroni method.

9.13 Multiplicity of Secondary Endpoints

Not applicable, the key secondary endpoint is not identified in protocol.

10 EFFICACY ANALYSIS

10.1 Primary Estimand

The primary estimand defining the treatment effect of interest in the trial uses the hypothetical strategy specified in the draft ICH E9 (R1) Addendum. The estimand, or target

Statistical Analysis Plan: 18-AVP-786-207

of estimation, following the hypothetical strategy is the pharmacological effect seen, had no withdrawals occurred. This hypothetical estimand is justifiable in this case, since the focus is on the pharmacological effect of the drug additional to non-specific effects. Subjects who withdraw from a symptomatic investigational medicinal product (IMP) treatment either could have lost their treatment effect, had the subjects not taken any other symptomatic medication after withdrawal, or could have their treatment effect been masked, had the subjects taken other symptomatic medication after withdrawal. This means that any observations taken after subjects stop study medication will most likely not contribute relevant information about the pharmacological effect of the drug. Due to this strategy, the last collected efficacy assessment after premature trial discontinuation will be done only once at the early termination (ET) Visit. Every effort will be made to complete all the ET evaluations prior to administering any additional medications for the treatment of negative symptoms of schizophrenia or other prohibited medications. In the case of terminal or lost to follow-up events, no ET evaluations would be expected, and only scheduled assessments performed before such an event has occurred.

In this hypothetical strategy, the event of withdrawing study medication is considered missing at random (MAR), and the primary endpoint of the trial will use the Mixed Model Repeated Measurements (MMRM) method. All data collected during the trial treatment period will be used for statistical analysis. Under the MAR assumption, MMRM provides an unbiased estimate of treatment effect for the treatment period.

The primary estimand for this study is defined by the following components:

- a. Target Population: mITT
- b. Endpoint: Change from baseline (Week 3) to Week 15 in the PANSS Marder negative factors score
- c. Intercurrent Events: Premature Treatment Discontinuation
- d. Measure of Intervention effect: Difference in endpoint means between AVP-786 and Placebo arms had no subject discontinued the treatment

10.2 Derivation of Primary Efficacy Endpoints

PANSS are assessed at Screening, Visit 1, baseline (Week 3), Week 6, 9, 12, and Week 15/ET Visit. Observed PANSS scores including those from unscheduled visits and early discontinuation, are assigned analysis visit based on the visit window in [Table 2](#) using assessment's study days. If more than one observation falls within a particular study day interval, then the last observation within that interval will be used. After applying visit window, there is only one value (non-missing or missing) for each analysis visit. These data are called observed data and change from baseline (Week 3) in PANSS Marder negative factors score (observed data) for efficacy analyses are derived by subtracting corresponding baseline value from the value at each visit. Other efficacy endpoints derivation follows the same algorithm.

Statistical Analysis Plan: 18-AVP-786-207

10.3 Primary Efficacy: MMRM (Mixed-Model Repeated Measures)

The primary efficacy endpoint, change from baseline (Week 3) to Week 15 in PANSS Marder negative factors score, will be analyzed using likelihood-based mixed-model repeated measures (MMRM) analysis on the observed data for the mITT population. The primary statistical comparison of interest is AVP-786 vs. placebo. The null hypotheses of the comparison is that there is no difference between the AVP-786 and placebo treatment groups in change from baseline (Week 3) to Week 15 in PANSS Marder negative factors score. The MMRM analysis will include fixed effects for treatment, trial center, visit, treatment-by-visit interaction, and baseline-by-visit interaction. An unstructured covariance (UN) model will be used. The methods from Lu and Mehrotra (2010) will be used to resolve the convergence issues by Fisher scoring algorithm and the factor analytic structure before moving on to a structured covariance matrix. Baseline-by-visit interaction is added to the MMRM to allow for more appropriate model fitting by accounting for different coefficients of baseline at each post-baseline visit and thus avoid potential convergence issue due to insufficient model specification. Small trial centers will be pooled according to region and practice for the analysis before database unblinding. The primary comparison between the AVP-786 and the placebo at Week 15 will be estimated as the difference between Least Squares (LS) means from the interaction term of treatment by visit week utilizing the computing software SAS procedure PROC MIXED.

CCI



The SAS code for the PROC MIXED procedure to carry out the above MMRM analysis with an unstructured variance covariance structure is illustrated as follows:

Statistical Analysis Plan: 18-AVP-786-207

```
proc mixed;  
  class treatment center visit subjid;  
  model change=treatment center visit treatment*visit baseline*visit /s cl ddfm=kr;  
  repeated visit /type=un subject=subjid r rcorr;  
  lsmeans treatment*visit / pdiff cl alpha=0.05 slice=visit;  
  ods output diffs=diffs lsmeans=lsmeans;  
run;
```

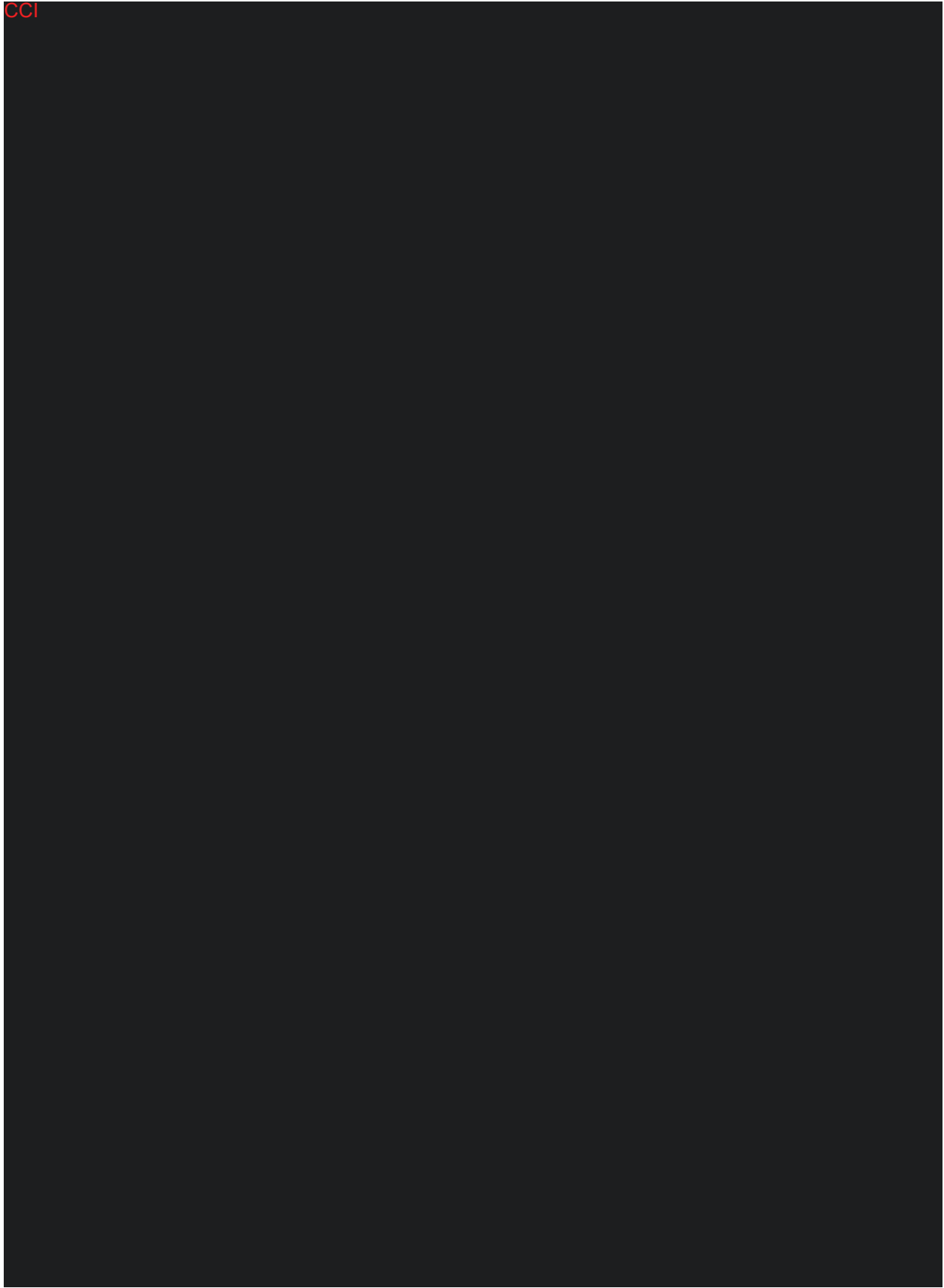
Baseline is the PANSS Marder negative factors score at baseline/Visit 2 (Week 3).

Small centers will be defined as centers that do not have at least one evaluable subject (mITT Population) in each treatment arm in the double-blind treatment period. All small centers will be pooled within region to form “pseudo centers” for the purpose of analysis according to the following algorithm. Small centers will be ordered from the largest to the smallest based on the number of evaluable subjects. The process will start by pooling the largest of the small centers with the smallest of the small centers until a non-small center is formed. This process will be repeated using the centers left out of the previous pass. In case of ties in center size, the center with the smallest center code will be selected. If any centers are left out at the end of this process, they will be pooled with the smallest pseudo centers, or if no pseudo centers exist, they will be pooled with the smallest non-small center. The pooling of the small centers will be performed within each region (North America, Eastern Europe, and Western Europe). In case of gross violations of the MMRM model assumptions, additional supportive analyses will be provided using the generalized version of the Cochran-Mantel-Haenszel (CMH) procedure (van Elteren test), controlling for trial center (Davis and Chung, 1995).

CCI

Statistical Analysis Plan: 18-AVP-786-207

CCI



Statistical Analysis Plan: 18-AVP-786-207

CCI



10.6 Subgroup Analyses for Primary Efficacy Endpoint

Due to potential small sample sizes for some subgroups, all subgroup analyses will be performed by the same MMRM as in the primary efficacy analysis but without trial center as a class effect. In addition, interaction of treatment-by-subgroup will be assessed at Week 15 by including subgroup, treatment-by-subgroup to the SAS PROC MIXED programming codes. Results of the treatment-by-subgroup interaction analyses will be discussed in the study report.

All estimated treatment effect by subgroup will be derived from separate MMRM, including only patients in the corresponding subgroup. The subgroup analysis MMRM will be similar to the one applied to the primary endpoint analysis without trial center as a class effect.

10.6.1 Subgroup Analyses by Age, Gender, Race, Country, Region and Years Since Onset of Schizophrenia

Subgroup analyses of change from baseline in PANSS Marder negative factors score will be performed by the following factors:

- Gender
- Race (White and All Other Races)
- Age group (<45; ≥45, or median of the mITT Population)
- Country (MMRM will not be performed if the sample size is too small for some countries. Summary statistics will be provided.)
- Region (US and Non-US)
- Years Since Onset of Schizophrenia (<20 and ≥20)

Statistical Analysis Plan: 18-AVP-786-207

10.6.2 Subgroup Analyses by Baseline Concomitant Medication According to CYP2D6 Major Substrate Status

A baseline concomitant medication is defined as a medication with a start date on or before first dose day and an end date on or after first dose day (or the medication is ongoing). Baseline concomitant medications are considered to be stable throughout the study. All baseline concomitant medications will be checked against the list of psychotropic medications substrate status with respect to cytochrome P450 isoenzyme 2D6 (CYP2D6). From a broader list of P450 CYP2D6 major substrates listed below, psychotropic medications (antidepressants, antipsychotics) that are major substrates of CYP2D6 have been selected and will be analyzed for any potential impact on efficacy and safety. These drugs are: aripiprazole, risperidone, duloxetine, fluoxetine, fluvoxamine, mirtazapine, paroxetine, and venlafaxine.

The PANSS Marder negative factors score will be analyzed for those who take baseline concomitant psychotropic medications that are CYP2D6 major substrates and for those who take baseline concomitant psychotropic medications that are not CYP2D6 major substrate.

In addition, concomitant use of beta blockers at baseline will be analyzed for treatment-emergent adverse events (TEAEs) (cardiovascular, falls, etc.). Analysis will be done for patients taking concomitant beta blockers that are and are not major substrates of CYP2D6. Beta blockers that are major CYP2D6 substrates include: carvedilol, metoprolol, propranolol and timolol. All other beta blockers are considered to be not major CYP2D6 substrates. For this analysis, TEAEs by system organ class (SOC) and preferred term (PT) will be assessed for those who take concomitant beta blockers that are CYP2D6 major substrates and for those who took beta blockers that were not CYP2D6 major substrates.

Drugs that are susceptible CYP2D6 major substrates include: amitriptyline, aripiprazole, atomoxetine, captopril, carvedilol, chlorpromazine, chlorpheniramine, clomipramine, codeine, desipramine, dextromethorphan, doxepin, duloxetine, flecainide, fluoxetine, fluphenazine, fluvoxamine, haloperidol, iloperidone, maprotiline, mequitazine, methamphetamine, metoprolol, mexiletine, mirtazapine, nefazodone, norcodeine, nortriptyline, paroxetine, perphenazine, procainamide, promethazine, propranolol, risperidone, tamoxifen, thioridazine, timolol, tolterodine, tramadol, trimipramine, venlafaxine, zuclopenthixol.

10.6.3 Subgroup Analyses by Baseline Concomitant Antipsychotic Medications

Baseline concomitant antipsychotic medications will be summarized. The antipsychotic medications will be categorized as Benzodiazepine, serotonin and norepinephrine reuptake inhibitor (SNRI) and selective serotonin reuptake inhibitor (SSRI). The treatment effect on PANSS Marder negative factors score for those patients who take different antipsychotic medications will be evaluated at Week 15.

Statistical Analysis Plan: 18-AVP-786-207

10.7 Secondary Efficacy and Other Outcome Measures Analysis

Secondary and other efficacy analyses are listed below. All these efficacy endpoints will be evaluated at a nominal 0.05 level (2-sided) without adjusting for multiplicity.

10.7.1 Secondary Efficacy Analyses

- 1) NSA-16 global negative symptom score
- 2) PGI-S
- 3) PGI-C

Endpoints (1) and (2) will be evaluated using the same MMRM model described in the primary analysis.

Endpoint (3) will be evaluated by the Cochran-Mantel-Haenszel (CMH) Row Mean Score Differ Test controlling, in last observation carried forward (LOCF) analysis, for study center. Observed case (OC) analysis will also be conducted for endpoint (3), but will not control for study center.

10.7.2 Other Efficacy Analyses

- 4) PANSS positive subscale
- 5) CDSS
- 6) Change from baseline (Week 3) in PANSS Marder negative factors score for every trial week visit in Phase B other than the Week 15 visit
- 7) PANSS Marder negative factors score, response defined as 30% improvement from baseline (Week 3) by visit

Endpoints (4) through (6) will be evaluated using the same MMRM model described in the primary analysis.

Variable (7) will be evaluated by the Cochran-Mantel-Haenszel (CMH) General Association Test controlling, in LOCF analysis, for study center. OC analysis will also be conducted for the same variable, but will not control for study center.

10.8 SAFETY ANALYSIS

Safety will be assessed through the analysis of AEs, serious AEs (SAEs), physical and neurological examinations, vital signs, clinical laboratory assessments, resting 12-lead ECGs, the C-SSRS, the AIMS, the BAS, and the SAS. All Safety analyses will be completed on the Safety Population.

10.9 Adverse Events

Adverse event tables (except the AE overview table) will only include summaries of TEAEs. Treatment-emergent adverse events on Safety population are defined as AEs which first occur, or worsen, after the first dose of study medication in double-blind treatment period and within 30 days after the permanent discontinuation of the study medication. Treatment-emergent adverse events on Run-in population (all patients who are in double-blind placebo

Statistical Analysis Plan: 18-AVP-786-207

run-in period and take at least one dose of study drug) are defined as AEs which first occur, or worsen, after the first dose of study medication in double-blind placebo run-in period and before the first dose of study medication in double-blind treatment period which starts at Visit 2 (Week 3, Day 22).

An overview table containing the number and percent of the following will be included:

- Number of total AEs, TEAEs, and deaths
- Incidence of patients with at least one TEAE, drug-related TEAE, serious TEAE, drug-related Serious TEAE, non-serious TEAEs, and death
- Incidence of patients who discontinued treatment due to TEAE, drug-related TEAE, Serious TEAE, and drug-related Serious TEAE
- Incidence of deaths and deaths due to drug-related TEAE

TEAEs will be summarized by SOC and PT, descending frequency of PT, by maximum severity, by dose at onset, and by age group (< 45 , ≥ 45). Summaries of PTs will also be done for those with at least 5% incidence in any treatment or study stage. In addition, summaries of TEAE for patients taking beta blockers at baseline will be summarized as described in [Section 10.6.1](#).

Adverse events leading to treatment discontinuation will be summarized by SOC and PT. Drug-related AEs will be summarized by SOC and PT and in descending frequency of PT.

Time to onset for common TEAEs (as defined below) will be summarized descriptively for each of these events. In addition, summary stats for the duration and percentage of total study days will be provided for each AE. The number and percent of patients with recurrences will also be given.

Serious TEAEs will be summarized by SOC and PT, descending frequency of PT, and will include a summary of drug-related events. Non-serious TEAEs will also be summarized by SOC and PT. Data will be listed for the AE that required hospitalizations or day clinics.

Below are the rules to follow for AE summaries:

- A drug-related AE is defined as an AE with an assigned relationship of “possibly related,” “related,” or missing.
- When assessing severity, the TEAE with the worst severity will be chosen for a patient. AEs with missing severity will be excluded from summaries of AE by severity.
- For dose at onset summaries, patients are counted only once within a dose, but can be counted in multiple dose groups if the same AE occurred while on different doses.
- A common TEAE is defined as a TEAE with an incidence of $\geq 3\%$ in the All AVP-786 treatment group AND ≥ 2 times the incidence of the Placebo group.
- Time to onset will be calculated in days as (AE start date – first dose date).

Statistical Analysis Plan: 18-AVP-786-207

- Duration of AE is generally defined as (AE end date – AE start date + 1). Duration will be calculated for placebo and AVP-786 separately.
- For a given patient, percentage of total study days is defined as total duration (as defined above) divided by (last dose date – first dose date + 1) $\times 100\%$.
- Recurrence is defined as a new report of the same TEAE with a new AE start date within a given treatment.

AEs will be coded using MedDRA version 22.0.

10.10 Clinical Laboratory Assessments

Clinical labs will be reported for hematology, chemistry and urinalysis. Labs are collected at Screening, Week 0, Week 3, Week 9 and Week 15, and will be summarized by visit and for the Placebo and AVP-786 groups as described below.

The following parameters will be summarized descriptively through CFBL and percent CFBL for numeric values. If a parameter is categorical, it will be listed only.

- Chemistry: calcium, magnesium, phosphorus, glucose, sodium, potassium, chloride, carbon dioxide, blood urea nitrogen (BUN), serum creatinine, uric acid, albumin, total bilirubin, alkaline phosphatase, lactate dehydrogenase (LDH), aspartate aminotransferase/serum glutamic oxaloacetic transaminase (AST/SGOT), alanine aminotransferase/serum glutamic pyruvic transaminase (ALT/SGPT), creatine kinase (CK), gamma-glutamyl transferase (GGT), triglycerides, total protein, total cholesterol, and glycosylated hemoglobin (HbA1c, at Screening and Visit 6 or ET only).
- Hematology: red blood cell (RBC) count, hemoglobin, hematocrit, white blood cell (WBC) count, neutrophils, bands, lymphocytes, monocytes, eosinophils, basophils, platelet count, and morphology
- Urinalysis: pH, specific gravity, protein, glucose, ketones, blood, leucocyte esterase, nitrates, and microscopic appearance

Out-of-range values will be assessed through shift tables. Each lab value will be assessed as low, normal or high based on the normal ranges provided by the central lab. Frequencies of each combination of shifts will be provided by treatment group.

Potentially clinically significant (PCS) tables will also be used to summarize out-of-range values. The PCS values are found in [Table 3](#). The number and percent of patients meeting the criteria below will be summarized by treatment group. Summaries will be given for any time post-baseline. The denominator for the percentages will be the number of patients who had a post-baseline assessment for each parameter.

Statistical Analysis Plan: 18-AVP-786-207

Table 3 LAB PCS Criteria

Laboratory Parameter	Unit	Low PCS Criteria	High PCS Criteria	Laboratory Parameter	Unit	Low PCS Criteria	High PCS Criteria
Chemistry							
Albumin	g/L	≤ 26	≥ 60	GGT	U/L	None	≥ 60
Alkaline Phosphatase	U/L	None	≥ 3×ULN	Glucose (Random, Fasting)	mg/dL	≤ 39	≥ 401
ALT (SGPT)	U/L	None	≥ 3×ULN	LDH	U/L	None	≥ 3×ULN
AST (SGOT)	U/L	None	≥ 3×ULN	Magnesium	mmol/L	< 0.37	> 1.23
Bilirubin	μmol/L	None	≥ 1.5×ULN	Phosphate	mmol/L	≤ 0.4522	> 3.88
BUN	mmol/L	None	≥ 10.71	Potassium	mmol/L	≤ 3.0	≥ 5.5
Calcium	mmol/L	≤ 1.75	≥ 3.0	Protein	g/L	≤ 50	≥ 100
Carbon Dioxide	mmol/L	≤ 9	> 40	Sodium	mmol/L	≤ 130	≥ 155
Chloride	mmol/L	≤ 85	≥ 120	Triglycerides	mmol/L	None	> 3.39
Cholesterol	mmol/L	None	≥ 7.77	Uric acid (Male)	μmol/L	None	≥ 624.54
Creatine Kinase	U/L	None	≥ 3×ULN	Uric acid (Female)	μmol/L	None	≥ 505.58
Creatinine	μmol/L	None	> 132.6	HbA1c	%	None	> 6.5%
Hematology							
Hemoglobin	g/L	< 100	> 180	Monocytes	×10 ⁹ /L	None	> 1
Hematocrit	proportion of 1.0	< 0.3	> 0.5	Monocytes/Leukocytes	%	None	≥ 15
Basophils	×10 ⁹ /L	None	> 0.3	Neutrophils/Leukocytes	%	≤ 15	None
Eosinophils/Leukocytes	%	None	≥ 10	Leukocytes	×10 ⁹ /L	≤ 2.8	≥ 16
Lymphocytes	×10 ⁹ /L	≤ 0.5	> 4	Erythrocytes	×10 ¹² /L	≤ 2.5	≥ 7.0
Lymphocytes/Leukocytes	%	≤ 10	≥ 60	Platelet Count	×10 ⁹ /L	≤ 100	≥ 700

Over the course of the study, there may be some lab tests performed that were not mentioned in the protocol. These tests will not be summarized but will be included in the listings and flagged as non-protocol tests.

10.11 ECGs

Electrocardiograms will be assessed by a central reader and will be recorded at the following visits:

- Screening (triplicate readings)
- Visit 1 (Day 1) (pre-dose)

Statistical Analysis Plan: 18-AVP-786-207

- Visit 1 (Day 1) (1-2 hours post-dose)
- Visit 2/Week 3 (Day 22) (pre-dose)
- Visit 2/Week 3 (Day 22) (1-2 hours post-dose)
- Visit 3/Week 6 (Day 43) (pre-dose)
- Visit 3/Week 6 (Day 43) (1-2 hours post-dose)
- Visit 4/Week 9 (Day 64) (pre-dose)
- Visit 5/Week 12 (Day 85) (pre-dose)
- Visit 6/Early Termination (Day 106) (pre-dose)

The following quantitative parameters will be reported by the central reader: heart rate (HR), PR interval, QRS duration, QT interval (uncorrected), and QT interval with Fridericia's correction (QTcF). Change from baseline and percent change from baseline by visit will be calculated for each parameter and summarized by the treatment groups. The pre-dose baseline ECG parameter values will be the pre-dose measurements at Visit 2 and the post-dose baseline values will be the post-dose measurements at Visit 2. The pre- and post-dose ECG values change from baseline at each visit will be calculated by subtracting pre-dose and post-dose baseline values correspondingly.

In addition, change from pre- to post-dose will be summarized by visits.

PR interval and QTcF will be further investigated through PCS tables, for which the criteria are found in [Table 4](#) below. The number and percent of patients meeting the criteria below will be summarized by treatment group. Summaries will be given for both overall (i.e., any time post-baseline) and by visit. For QTcF, males and females will be assessed separately. Patients will be included in all categories for which they qualify. For criteria on the "Actual" values, the denominator for the percentages is the number of patients who had a post-baseline assessment for each parameter. For criteria on the change, the denominator is the number of patients who had a baseline and post-baseline assessment.

Table 4 ECG PCS Criteria

ECG Parameter	Sex	Actual or Change	PCS Criteria
PR Interval (msec)	Both	Actual	> 200 to ≤ 220, > 220 to ≤ 250, > 250
QTcF (msec)	Males	Actual	> 450 to ≤ 480, > 480 to ≤ 500, > 500
	Females	Actual	> 460 to ≤ 485, > 485 to ≤ 500, > 500
	Both	Change from baseline (increase)	≥ 30, ≥ 60

ECG overall interpretations will be summarized by the number and percent that were normal or abnormal. The interpretations by the cardiologist (i.e., central ECG, iCardiac) will be used

Statistical Analysis Plan: 18-AVP-786-207

for these summaries. For females, iCardiac uses QTcF > 460 msec to ≤ 485 msec for PCS criteria and this will be flagged in corresponding tables and listings. The listings will provide all interpretations and corresponding details.

10.12 Vital Signs

Vital signs will be assessed at all visits. Orthostatic blood pressure is taken at Screening and sitting/semi-recumbent blood pressure is taken at all subsequent visits. All sitting/semi-recumbent measurements are recorded twice, so the mean of the 2 measurements will be taken and used for all summaries mentioned below. The following parameters recorded in the sitting/semi-recumbent positions will be summarized: systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate. Weight, height, respiratory rate, and body temperature will also be summarized. These parameters will be summarized through CFBL and percent CFBL in similar fashion as the ECG parameters. For heart rate, SBP and DBP, box plots by visit will be presented.

Vital signs will also be assessed through PCS criteria, which are given in [Table 5](#). Patients will be counted if they meet the criteria below at any time post-baseline. The definition of baseline is consistent with those for ECG parameters. The denominators are the number of patients with both a baseline and post-baseline assessment.

Table 5 Vital Sign PCS Criteria

Vital Sign Parameter	High Values	Low Values
SBP (mmHg)	> 180 AND ≥ 20 increase from baseline	≤ 90 AND ≥ 20 decrease from baseline
DBP (mmHg)	≥ 105 AND ≥ 15 increase from baseline	≤ 50 AND ≥ 15 decrease from baseline
HR (bpm)	≥ 120 AND ≥ 15 increase from baseline	≤ 50 AND ≥ 15 decrease from baseline
SBP and HR	SBP ≥ 10 increase from baseline AND HR ≥ 5 increase from baseline	Not Applicable
DBP and HR	DBP ≥ 5 increase from baseline AND HR ≥ 5 increase from baseline	Not Applicable

All vital sign parameters will be included in the listings.

10.13 Physical and Neurological Exams

Physical and neurological exams are assessed at Screening and at the final clinic visit, which will be listed.

10.14 C-SSRS

Scoring for the C-SSRS will be analyzed using the following indicators:

- Ideation severity: each of the 5 questions regarding type of ideation (yes/no)
- Intensity of most severe ideation: a sum of the 5 intensity items
- Suicidal behavior types: each of the 4 suicidal behavior types (yes/no)

Statistical Analysis Plan: 18-AVP-786-207

- Suicidal behavior: suicidal behavior question
- Actual lethality: actual lethality question (0-5 scale) on most lethal attempt
- Potential lethality: potential lethality question (0-2 scale) on most lethal attempt

The individual questions listed above that have yes/no responses will be summarized by treatment group and visit using number and percent.

Intensity of most or severe ideation will be summarized descriptive by treatment group and visit.

Items that contain an ordinal response will be summarized both through descriptive statistics and number and percent.

All C-SSRS data, including individual text descriptions (included as part of some questions), will be included in the listings.

10.15 Simpson Angus Scale for Extrapyramidal Symptoms (SAS)

The individual items, as well as the total score, will be summarized. Descriptive statistics will be provided by treatment group and visit for males and females separately, as well as both sexes overall. For each item, a shift table of the number and percent of patients in each score (0 to 4) from Baseline to end of study will be summarized overall and for males and females separately. A shift table for total score from Baseline to end of study will also be summarized overall and for males and females separately as well. The same MMRM model as described in the primary efficacy analysis will be applied for the change from baseline (Week 3) in SAS total score at Week 15.

10.16 Barnes Akathisia Scale (BAS)

The objective assessment, subjective awareness, subjective distress, and global clinical assessment, as well as the total score, will be summarized. Descriptive statistics will be provided by treatment group and visit for males and females separately, as well as both sexes overall. Shift tables of the number and percent of patients of the total score and the global clinical assessment will be presented for Baseline to end of study, for males and females separately as well as both sexes overall. The same MMRM model as described in the primary efficacy analysis will be applied for the change from baseline (Week 3) in BAS total score at Week 15.

10.17 Abnormal Involuntary Movement Scale (AIMS)

The following table represents the scores to be calculated and summarized:

Subscale/Score	Item Numbers	Score Range
Orofacial Movements	1, 2, 3, 4	0 – 16
Extremity Movements	5, 6	0 – 8
Trunk Movements	7	0 – 4
Global Judgment	8, 9, 10	0 – 12

Statistical Analysis Plan: 18-AVP-786-207

Dental Status	11, 12	Yes/No
Observed Movements	1 – 7	0 – 28
Overall Severity	8	0 – 4

Descriptive statistics will be provided by treatment group and visit for males and females separately, as well as both sexes overall. Shift tables of the number and percent of patients will be presented for Baseline to end of study for the overall severity, for males and females separately as well as both sexes overall. The same MMRM model as described in the primary efficacy analysis will be applied for the change from baseline (Week 3) in AIMS total score at Week 15.

11 ADDITIONAL STATISTICAL ANALYSIS METHODS

11.1 Patient Disposition

Counts of patient enrollment will provide the number of patients screened along with the reason for screen failures. A summary of randomized treatment group by responder status will be provided.

11.2 Demographics and Baseline Characteristics

Demographics will be summarized by treatment group for the mITT and randomized populations. The following characteristics will be summarized.

- Sex (Male, Female)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Age
- Age group (< 45, ≥ 45)
- Height
- Weight
- Body mass index

A summary of baseline characteristics will include baseline values of all primary and secondary endpoints, in addition to height, weight, and body mass index. These will be summarized by randomized treatment group.

For categorical parameters, the denominators for the percentages are the number of patients who had the parameter assessed.

11.3 Medical History and Disease Characteristics

Medical history will be summarized by SOC and PT based on randomization and will be presented in the listings. Summary tables will be provided for the frequency and percentage of patients.

Statistical Analysis Plan: 18-AVP-786-207

Additionally, descriptive statistics for the number of years since the onset of schizophrenia will be presented. In both cases, the number of years will be counted relative to the patient's randomization date.

11.4 Exposure and Treatment Compliance

Duration of exposure will be summarized quantitatively using the number of days on study medication for each patient, displayed by treatment group. Duration will be calculated as (last dose date – first dose date + 1).

Overall treatment compliance will be calculated as a percentage using the total number of capsules that were dispensed and returned. Patients will be grouped into categories of < 80%, 80% to 120%, > 120%. Counts will be summed over the visits for each patient to calculate an overall compliance value.

Compliance will be summarized as described above and through descriptive statistics. The number of doses taken and number of doses should have taken are necessary to calculate compliance.

$$\text{Compliance} = (\text{doses taken} / \text{doses should have taken}) \times 100\%$$

where doses taken = dispensed amount - returned amount; doses should have taken = $2 \times (\text{last dose day} - \text{first dose day}) + 1$, which are based on the assumption that patients take 2 capsules per day except for the last day, in which they only take 1.

The number of capsules taken for kits that were not returned will be imputed to be the number of capsules that were dispensed.

11.5 Concomitant Medications

The number and percent of concomitant medications will be provided by the treatment groups used for the safety analyses.

Medications will be classified as:

- Pre-treatment medications: medications taken and stopped prior to the first dose date of study drug.
- Concomitant medications: medications taken after receiving the first dose of the study drug. These belong to the following categories:
 - Prior concomitant medications: medications started prior to the first dose of treatment and continued past the first dose date.
 - New concomitant medications: medications started during or after the first dose date of the study drug.

Concomitant medications will be coded using the World Health Organization-Drug (WHODrug) Dictionary Version 01 March 2019 (Global, B3 format). Summaries will be provided by anatomical therapeutic chemical classification (ATC) and preferred term.

Statistical Analysis Plan: 18-AVP-786-207

11.6 Analysis of Pharmacokinetics

Plasma concentrations of d6-DM, its metabolites d3-DX and d3-3-MM, and Q will be measured from samples collected post-dose on Day 22 (Week 3), pre- and post-dose on Day 64 (Week 9), and pre-dose on Day 106 (Week 15). These concentrations will be summarized descriptively overall and by CYP2D6 metabolizer group for those taking AVP-786.

A blood draw to assess concentrations of antipsychotics will be performed at Screening, Week 3, Week 9, and Week 15, and results from the plasma concentrations will be summarized descriptively.

12 ADDITIONAL STATISTICAL CONSIDERATIONS

12.1 Data Monitoring Committee (DMC)

DMC role and responsibilities as well as statistical methodology will be described in a separate IAP and DMC Charter.

Statistical Analysis Plan: 18-AVP-786-207

13 REFERENCES

1. Chen YHJ, DeMets DL, Lan KKG. Increasing the sample size when the unblinded interim results are promising. *Stat Med* 2004; 23: 1023-38.
2. Davis CS, Chung Y. Randomization model methods for evaluating treatment efficacy in multicenter clinical trials. *Biometrics*. 1995; 51: 1163-1174.
3. Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. *Applied Statistics*. 1994; 43: 49-93.
4. Little RJA. Modeling the drop-out mechanism in repeated measures studies. *J Am Stat Assoc*. 1995; 90: 1112-21.
5. Lu, K, & Mehrotra, DV. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. *Statistics in Medicine*. 2009; 29: n. pag.

Statistical Analysis Plan: 18-AVP-786-207

14 APPENDICES

Appendix 1 Schedule of Assessments and Procedures

			Study Treatment Period											
Procedure	Visit:	Screening	Visit 1	Phone ²	Visit 2 ¹	Phone ²	Visit 3 ¹	Phone ²	Visit 4 ¹	Phone ²	Visit 5 ¹	Phone ²	Visit 6 ¹ /ET ¹¹	Post Exit Phone ²
	Study Day:	Day -28 to -1	Day 1	Day 8	Day 22	Day 29	Day 43	Day 50	Day 64	Day 71	Day 85	Day 92	Day 106	Day 136
	End of Study Week:	Week -4 to -1	Week 0	Week 1	Week 3	Week 4	Week 6	Week 7	Week 9	Week 10	Week 12	Week 13	Week 15	Week 19
Obtain Informed Consent		X												
CTSDatabase Entry*		X											X	
Medical History/Psychiatric History		X												
M.I.N.I. Exam		X												
Review Inclusion and Exclusion Criteria		X	X											
Document Patient's Informant Information		X												
NSA-16 and NSA Global			X		X		X		X		X		X	
PANSS		X	X		X		X		X		X		X	
PGI-S			X		X		X		X		X		X	
PGI-C					X		X		X		X		X	
Enrollment/Randomization			X											
AiCure (Patient Training)			X											
Physical and Neurological Examinations		X											X	
Resting 12-lead ECG ³		X	X		X		X		X		X		X	
Chemistry, Hematology, and Urinalysis ⁴		X	X		X				X				X	

Statistical Analysis Plan: 18-AVP-786-207

Table 5: Schedule of Evaluations and Visits (Continued)

			Study Treatment Period											
Procedure	Visit:	Screening	Visit 1	Phone ²	Visit 2 ¹	Phone ²	Visit 3 ¹	Phone ²	Visit 4 ¹	Phone ²	Visit 5 ¹	Phone ²	Visit 6 ¹ /ET ¹¹	Post Exit Phone ²
	Study Day:	Day -28 to -1	Day 1	Day 8	Day 22	Day 29	Day 43	Day 50	Day 64	Day 71	Day 85	Day 92	Day 106	Day 136
	End of Study Week:	Week -4 to -1	Week 0	Week 1	Week 3	Week 4	Week 6	Week 7	Week 9	Week 10	Week 12	Week 13	Week 15	Week 19
Pregnancy Test ⁵		X	X		X		X		X		X		X	
Urine Drug Screen		X												
Hepatitis B, C; HIV antibody		X												
Lab -CYP2D6 ⁶			X											
Pharmacokinetic sampling ⁷					X				X				X	
Plasma Antipsychotic Levels		X			X				X				X	
Review of Adverse Events			X	X	X	X	X	X	X	X	X	X	X	X
Review Concomitant Medications		X	X	X	X	X	X	X	X	X	X	X	X	X
Record Vital Signs/Weight ⁸		X	X		X		X		X		X		X	
CDSS		X	X		X				X				X	
Extrapyramidal Symptoms Assessment Scales (AIMS, BAS, SAS)		X ⁹	X		X				X				X	
C-SSRS		X	X		X		X		X		X		X	
Dispense Study Drug			X		X		X		X		X			

Statistical Analysis Plan: 18-AVP-786-207

Table 5: Schedule of Evaluations and Visits (Continued)

			Study Treatment Period											
Procedure	Visit:	Screening	Visit 1	Phone ²	Visit 2 ¹	Phone ²	Visit 3 ¹	Phone ²	Visit 4 ¹	Phone ²	Visit 5 ¹	Phone ²	Visit 6 ¹ /ET ¹¹	Post Exit Phone ²
	Study Day:	Day -28 to -1	Day 1	Day 8	Day 22	Day 29	Day 43	Day 50	Day 64	Day 71	Day 85	Day 92	Day 106	Day 136
	End of Study Week:	Week -4 to -1	Week 0	Week 1	Week 3	Week 4	Week 6	Week 7	Week 9	Week 10	Week 12	Week 13	Week 15	Week 19
Dose in Clinic (first dose of the day)			X		X		X		X		X		X	
Review Returned Unused Study Drug				X ¹⁰	X	X ¹⁰	X	X ¹⁰	X	X ¹⁰	X	X ¹⁰	X	

*US patients/sites only

¹ Visit 2 has a ± 3 day window; Visits 3-6 have a ± 6 day window.² Telephone calls will have a +3 day window.³ Triplicate ECGs at Screening only. Electrocardiogram will be performed pre-dose and 1-2 hours (± 15 minutes) post-dose at Visit 1, Visit 2, and Visit 3; ECGs at all other visits will be performed pre-dose only (Visits 4, 5, and 6).⁴ Fasting glucose and lipids will be measured at Screening, Visits 2, 4 and Visit 6/ ET visit for patients who withdraw prior to study completion. HbA1c will be measured at Screening and Visit 6 / ET. Thyroid function tests will be measured at Screening only.⁵ Urinary (beta-hCG) test will be performed for all females of childbearing potential (serum beta-hCG at Screening only).⁶ Blood samples for CYP2D6 genotyping will be taken at Visit 1.⁷ Plasma concentrations of d6-DM, its metabolites (d3-DX and d3-3-MM), and Q will be measured from samples collected post-dose on Day 22 (Week 3), pre- and post-dose on Day 64 (Week 9), and pre-dose on Day 106 (Week 15). Samples collected after efficacy assessments are completed.⁸ Orthostatic BP and HR measurements will be performed (in duplicate) at all clinic visit. Respiratory rate and body temperature will be measured at Screening, Day 1 (Visit 1), and Day 106 (Visit 6)/ET. Weight and height will be measured at Screening, Day 1 (Visit 1), and Day 106 (Visit 6).⁹ Only the SAS will be measured at Screening (AIMS and BAS not measured at Screening).¹⁰ Patient will be asked if they have taken their medications as directed during telephone calls at Days 8, 29, 50, 71, and 92.¹¹ Patients who withdraw prior to study completion (ET) will be asked to return to the clinic to complete the Visit 6 assessments, and will receive a follow-up phone call 7 days after the last dose of study medication (for adverse events and concomitant medications information). These patients will also receive the safety follow-up phone call 30 days after the last dose of study medication.



This page is a manifestation of an electronically captured signature

SIGNATURE PAGE

Document Name: AVP-786 207 SAP

Document Number: 1000219277

Signed by	Meaning of Signature	Server Date (dd-MMM- yyyy hh:min) - UTC timezone
PPD	Clinical Approval	27-Jun-2023 18:17:44
PPD	Biostatistics Approval	26-Jun-2023 23:26:41



DMC CLOSED SESSION EXECUTIVE SUMMARY

Data Monitoring Committee (DMC) Meeting #01 (Interim Analysis) Executive Summary for CLOSED Session

Study Title and Version: A Multicenter, Randomized, Double-blind, Placebo-controlled, Parallel-arm Study to Assess the Efficacy, Safety, and Tolerability of AVP-786 (Deudextromethorphan Hydrobromide [d6-DM]/Quinidine sulfate [Q]) for the Treatment of Negative Symptoms of Schizophrenia

Protocol version 3.0 dated 09February2018

Protocol Number: 18-AVP-786-207

Sponsor: Otsuka Pharmaceutical Development & Commercialization, Inc.
508 Carnegie Ctr.
Princeton, NJ 08540

Meeting Number: 01 (IA)

Session Type CLOSED

Meeting Place: Teleconference

Date of Meeting: 29MAR2023

Data Snapshot Date: 03MAR2023

Date of Document Distribution: 21MAR2023

CONFIDENTIALITY STATEMENT

This document contains confidential information. Nothing herein may be disclosed without written consent from Agios Pharmaceuticals. In the event of an actual or suspected breach of this obligation, Agios Pharmaceuticals should be notified promptly.

CONFIDENTIAL

Page 1 of 13
ST-AD-018 version 04

Until the information herein is released by Otsuka to the public domain, the contents of this document are Otsuka confidential information and should not be duplicated or re-distributed without prior written consent of Otsuka.



DMC CLOSED SESSION EXECUTIVE SUMMARY

Meeting Invitees

DMC Members

PPD

DMC Chairperson

PPD

DMC Member

PPD

DMC Member

PPD

Labcorp DMC Support Team

PPD

DMC Support Statistician
Executive Director Biostatistics
Labcorp Drug Development

IDMC Support Statistician
Senior Biostatistician
Labcorp Drug Development

CONFIDENTIAL

Page 2 of 13
ST-AD-018 version 04

Until the information herein is released by Otsuka to the public domain, the contents of this document are Otsuka confidential information and should not be duplicated or re-distributed without prior written consent of Otsuka.



DMC CLOSED SESSION EXECUTIVE SUMMARY

Table of Contents

1	INTRODUCTION	4
1.1	GENERAL STUDY DESIGN	4
2	SUBJECT DISPOSITION	5
3	SUBJECT CHARACTERISTICS.....	6
4	DISEASE CHARACTERISTICS.....	8
5	ADVERSE EVENTS	9
5.1	TREATMENT EMERGENT ADVERSE EVENTS (TEAE)	9
5.2	TREATMENT EMERGENT ADVERSE EVENTS (TEAE) AS RELATED TO STUDY TREATMENT	9
5.3	SERIOUS TREATMENT EMERGENT ADVERSE EVENTS (TEAE).....	9
5.4	SERIOUS TREATMENT EMERGENT ADVERSE EVENTS (TEAE) AS RELATED TO STUDY TREATMENT.....	10
5.5	TREATMENT EMERGENT ADVERSE EVENTS (TEAE) LEADING TO DISCONTINUATION OF STUDY TREATMENT	10
5.6	TEAES LEADING TO DEATH	10
6	EFFICACY ANALYSIS.....	10

CONFIDENTIAL

Page 3 of 13
ST-AD-018 version 04

Until the information herein is released by Otsuka to the public domain, the contents of this document are Otsuka confidential information and should not be duplicated or re-distributed without prior written consent of Otsuka.



DMC CLOSED SESSION EXECUTIVE SUMMARY

1 INTRODUCTION

This document will provide an overview of the unblinded tables, listings, and figures (TLFs) that were distributed to the independent Data Monitoring Committee (DMC) members. This summary is meant to aid in the review of the TLFs not to provide any conclusions. Details about the study can be found in the associated protocol.

1.1 General Study Design

18-AVP-786-207 is a Phase 2/3, multicenter, randomized, double-blind, placebo-controlled study to assess the efficacy, safety, and tolerability of AVP-786 in patients with negative symptoms of schizophrenia. The study consists of up to a 4-week Screening period, a 3-week double-blind placebo run-in period (Phase A), a 12-week double-blind treatment period (Phase B), and a 30-day follow-up period. Up to 370 patients will be enrolled at approximately 40 study centers in the United States (US).

The Screening period will be 28 days (4 weeks) and will begin when consent has been obtained. An extension of up to 14 additional days (42-day maximum total) can be requested from the Medical Monitor, if needed, to meet eligibility requirements.

After Screening, eligible patients will enter a 3-week, double-blind, placebo run-in period, in which all patients will receive placebo capsules twice daily (BID). Both investigators and patients are blinded to the placebo run-in period.

Placebo run-in responders, defined as all patients who demonstrate a 20% or greater reduction in the PANSS Marder negative factors score from Visit 1 (Day 1) to Visit 2 (Day 22), will be eligible to enter the double-blind treatment period (described below), but will be excluded from all efficacy analyses.

After completion of the Phase A, patients will be randomized in a 1:1 ratio, stratified by placebo responder status and trial center, to receive either AVP-786 or matching placebo capsules during Phase B. Patients randomized to AVP-786 in Phase B will begin a one-week titration period in which they will receive a dose of d6-DM 28 mg/Q 4.9 mg (AVP-786-28/4.9) every morning (qam) and a dose of placebo every evening (qpm) for the first 3 days, followed by AVP-786-28/4.9 BID for the next 4 days. Following this one-week titration period, patients will receive d6-DM 42.63 mg/Q 4.9 mg (AVP-786-42.63/4.9) administered orally BID for the remaining 11 weeks of the double-blind treatment period (Phase B).

The 30-day follow-up period will begin following the last visit (Visit 6 [Week 15] or Early Termination [ET]). The patient will be contacted by phone at Week 19 to collect adverse event (AE) and concomitant medication information for the 30 days following the last dose of study medication.

CCI

CONFIDENTIAL

Page 4 of 13
ST-AD-018 version 04

Until the information herein is released by Otsuka to the public domain, the contents of this document are Otsuka confidential information and should not be duplicated or re-distributed without prior written consent of Otsuka.



DMC CLOSED SESSION EXECUTIVE SUMMARY

CCI

2 SUBJECT DISPOSITION

Review TLFs: Table 14.1 – 1 Summary of Analysis Populations (Randomized), Table 14.1 – 2.1 Overall Patient Disposition (All Screened Patients)

For this DMC, data was extracted on 03 March 2023. As of the extraction date, a total of 453 subjects screened and 438 (96.7%) subjects completed screening. Of these subjects 124 subjects were randomized and treated (65 subjects in the AVP-786 treatment group and 59 subjects in the Placebo treatment group). The enrollment goal of this study is to randomize approximately 370 subjects. Therefore, the current study enrollment is 33.5% of the enrollment goal. (14.1-2.1)

Of the 124 subjects randomized, 113 (91.1%) subjects are Placebo Non-responders (60 (92.3%) in AVP-786 and 53 (89.8%) in Placebo), 10 (8.1%) are Placebo Responders (4 (6.2%) in AVP-786 and 6 (10.2%) in Placebo). (14.1-2.1) Modified Intent to Treat (mITT) population include 113 placebo non-responders (60 subjects in AVP-786 and 53 subjects in placebo).

CONFIDENTIAL

Page 5 of 13
ST-AD-018 version 04

Until the information herein is released by Otsuka to the public domain, the contents of this document are Otsuka confidential information and should not be duplicated or re-distributed without prior written consent of Otsuka.



DMC CLOSED SESSION EXECUTIVE SUMMARY

Subject **PPD** was randomized to the AVP-786 treatment arm, but had no response/non-response indicator, hence the subject was excluded from the mITT set.

Of the 124 subjects randomized, 104 (83.9%) subjects Completed Study (55 (84.6%) subjects in AVP-786 and 49 (83.1%) in Placebo), 15 (12.1%) subjects Discontinued from Study (8 (12.3%) subjects in AVP-786 and 7 (11.9%) subjects in Placebo). The reasons for Discontinuation from Study – Withdrawal by Subject (7 (5.6%) subjects), Adverse Event (2 (1.6%) subjects), Non-Compliance with Study Drug (2 (1.6%) subjects), Protocol Deviation (2 (1.6%) subjects), Lack of Efficacy (1 (0.8%) subject), and Physician Decision (1 (0.8%) subject). (14.1-2.1)

3 SUBJECT CHARACTERISTICS

Review TLFs: Table 14.1 – 3.1 Demographics (mITT), Table 14.1 – 3.2 Demographics (Randomized)

Of the 113 subjects in the modified Intent-to-Treat (mITT) Population, 72 (63.7%) subjects are Male, 41 (36.3%) subjects are Female. The mean age (SD) is 44.6 (9.69); the mean height (SD) is 172.58 (8.958) cm; the mean weight (SD) is 94.04 (23.413) kg; the mean BMI (SD) is 31.63 (7.815) kg/m². Ninety-five (95, 84.1%) subjects are younger than 55, 18 (15.9%) subjects are 55 years old or older. The Race distribution is as follows: 25 (22.1%) subjects are White, 57 (50.4%) subjects are Black or African American, 1 (0.9%) subject is Asian, 1 (0.9%) subject is Other, and for 29 (25.7%) subjects the race is Missing. The Ethnicity distribution is as follows: 12 (10.6%) subjects are Hispanic or Latino, 72 (63.7%) subjects are Not Hispanic or Latino, and for 29 (25.7%) subjects the Ethnicity is missing. (14.1-3.1)

Table 1a. Demographics for mITT Population

	AVP-786 (N=60)	Placebo (N=53)	Total (N=113)
Sex			
Male	41 (68.3%)	31 (58.5%)	72 (63.7%)
Female	19 (31.7%)	22 (41.5%)	41 (36.3%)
Race			
White	14 (23.3%)	11 (20.8%)	25 (22.1%)
Black or African American	28 (46.7%)	29 (54.7%)	57 (50.4%)
Asian	1 (1.7%)	0	1 (0.9%)
Other	0	1 (1.9%)	1 (0.9%)
Missing	17 (28.3%)	12 (22.6%)	29 (25.7%)
Ethnicity			
Hispanic or Latino	6 (10.0%)	6 (11.3%)	12 (10.6%)
Not Hispanic or Latino	37 (61.7%)	35 (66.0%)	72 (63.7%)
Missing	17 (28.3%)	12 (22.6%)	29 (25.7%)
Age (years)			
n	60	53	113
Mean (SD)	45.2 (9.64)	43.8 (9.78)	44.6 (9.69)
Median	47.5	46.0	47.0

CONFIDENTIAL



DMC CLOSED SESSION EXECUTIVE SUMMARY

Min, Max	22, 60	23, 60	22, 60
Age Group (years)			
<55	48 (80.0%)	47 (88.7%)	95 (84.1%)
>=55	12 (20.0%)	6 (11.3%)	18 (15.9%)
Height (cm)			
n	57	52	109
Mean (SD)	172.32 (8.423)	172.87 (9.585)	172.58 (8.958)
Median	172.00	173.75	173.00
Min, Max	149.8, 188.0	154.0, 195.7	149.8, 195.7
Weight (kg)			
n	57	52	109
Mean (SD)	91.44 (23.015)	96.90 (23.734)	94.04 (23.413)
Median	89.40	94.35	92.60
Min, Max	52.2, 147.4	57.8, 157.6	52.2, 157.6
Body Mass Index (kg/m ²)			
n	57	52	109
Mean (SD)	30.81 (7.534)	32.53 (8.089)	31.63 (7.815)
Median	29.51	31.90	30.12
Min, Max	19.2, 46.8	19.7, 55.2	19.2, 55.2

(14.1-3.1)

Of the 124 subjects randomized, 80 (64.5%) subjects are Male, 44 (35.5%) subjects are Female. The mean age (SD) is 44.6 (9.65) years; the mean height (SD) is 172.74 (8.999) cm; the mean weight (SD) is 94.52 (23.189) kg ; the mean BMI (SD) is 31.72 (7.681) kg/m². Hundred-three (103, 83.1%) subjects are younger than 55, 21 (16.9%) subjects are 55 years old or older. The Race distribution is as follows: 25 (20.2%) subjects are White, 67 (54.0%) subjects are Black or African American, 1 (0.8%) subject is Asian, 1 (0.8%) subject is Other, and for 30 (24.2%) subjects race is Missing. The Ethnicity distribution is as follows: 12 (9.7%) subjects are Hispanic or Latino, 82 (66.1%) subjects are Not Hispanic or Latino, and for 30 (24.2%) subjects the Ethnicity is Missing. (14.1-3.2)

Table 1b. Demographics for Randomized subjects

	AVP-786 (N=65)	Placebo (N=59)	Total (N=124)
Sex			
Male	45 (69.2%)	35 (59.3%)	80 (64.5%)
Female	20 (30.8%)	24 (40.7%)	44 (35.5%)
Race			
White	14 (21.5%)	11 (18.6%)	25 (20.2%)
Black or African American	33 (50.8%)	34 (57.6%)	67 (54.0%)
Asian	1 (1.5%)	0	1 (0.8%)
Other	0	1 (1.7%)	1 (0.8%)
Missing	17 (26.2%)	13 (22.0%)	30 (24.2%)
Ethnicity			

CONFIDENTIAL



DMC CLOSED SESSION EXECUTIVE SUMMARY

Hispanic or Latino	6 (9.2%)	6 (10.2%)	12 (9.7%)
Not Hispanic or Latino	42 (64.6%)	40 (67.8%)	82 (66.1%)
Missing	17 (27.2%)	13 (22.0%)	30 (24.2%)
Age (years)			
n	65	59	124
Mean (SD)	44.8 (9.54)	44.4 (9.84)	44.6 (9.65)
Median	46.0	47.0	46.5
Min, Max	22, 60	23, 60	22, 60
Age Group (years)			
<55	53 (81.5%)	50 (84.7%)	103 (83.1%)
>=55	12 (18.5%)	9 (15.3%)	21 (16.9%)
Height (cm)			
n	62	57	119
Mean (SD)	172.19 (8.133)	173.34 (9.894)	172.74 (8.999)
Median	171.75	174.00	173.00
Min, Max	149.8, 188.0	154.0, 195.7	149.8, 195.7
Weight (kg)			
n	62	57	119
Mean (SD)	92.09 (22.732)	97.17 (23.590)	94.52 (23.189)
Median	90.40	94.40	92.80
Min, Max	52.2, 147.4	57.8, 157.6	52.2, 157.6
Body Mass Index (kg/m ²)			
n	62	57	119
Mean (SD)	31.08 (7.529)	32.41 (7.851)	31.72 (7.681)
Median	29.93	31.64	30.80
Min, Max	19.2, 46.8	19.7, 55.2	19.2, 55.2

(14.1-3.2)

4 DISEASE CHARACTERISTICS

Review TLFs: Table 14.1 – 4 Baseline Disease Characteristics (Randomized)

All randomized subjects reported onset of Schizophrenia.

Table 2. Number of Years Since the Onset of Schizophrenia (All Randomized Subjects)

	AVP-786 N=65	Placebo N=59	Total N=124
Mean (SD)	20.4 (10.29)	19.7 (9.34)	20.1 (9.82)
Median	20.0	21.0	20.5
Min, Max	4, 43	1, 41	1, 43

CONFIDENTIAL



DMC CLOSED SESSION EXECUTIVE SUMMARY

5 ADVERSE EVENTS

5.1 Treatment Emergent Adverse Events (TEAE)

Review TLFs: Table 14.3 – 1.1 Overview of Adverse Events, Table 14.3 – 1.4.1 Summary of TEAEs by Preferred Term

A total of 95 Adverse Events were reported (60 in AVP-786 and 35 in Placebo). Of these Events, 67 are considered to be Treatment-Emergent (47 in AVP-786 and 20 in Placebo).

Of the 124 subjects in the safety population, 39 (31.5%) subjects had at least one TEAE (25 (38.5%) subjects in AVP-786, and 14 (23.7%) subjects in Placebo).

The most common TEAEs reported (occurrence rate in any treatment group >1.5%) are shown in the following Table.

Table 3. Most Common Treatment Emergent Adverse Events (TEAE)

	AVP-786 N=65	Placebo N=59	Total N=124
At Least One TEAE	25 (38.5%)	14 (23.7%)	39 (31.5%)
Headache	2 (3.1%)	2 (3.4%)	4 (3.2%)
Arthralgia	2 (3.1%)	1 (1.7%)	3 (2.4%)
Dizziness	2 (3.1%)	1 (1.7%)	3 (2.4%)
Back pain	0	2 (3.4%)	2 (1.6%)
Diabetes mellitus	1 (1.5%)	1 (1.7%)	2 (1.6%)
Fall	2 (3.1%)	0	2 (1.6%)
GGT increased	2 (3.1%)	0	2 (1.6%)
Hypertension	2 (3.1%)	0	2 (1.6%)
Nasopharyngitis	2 (3.1%)	0	2 (1.6%)
Weight Increased	2 (3.1%)	0	2 (1.6%)

5.2 Treatment Emergent Adverse Events (TEAE) as Related to Study Treatment

Review TLFs: Table 14.3 – 1.1 Overview of Adverse Events

Six (6, 4.8%) subjects experienced at least one Drug-related TEAE (4 (6.2%) in AVP-786 and 2 (3.4%) in Placebo).

5.3 Serious Treatment Emergent Adverse Events (TEAE)

Review TLFs: Table 14.3 – 1.1 Overview of Adverse Events

Two (2, 1.6%) subjects experienced at least one Serious TEAE (2 (3.1%) subjects in AVP-786).

CONFIDENTIAL



DMC CLOSED SESSION EXECUTIVE SUMMARY

5.4 Serious Treatment Emergent Adverse Events (TEAE) as Related to Study Treatment

Review TLFs: Table 14.3 – 1.1 Overview of Adverse Events

No subject experienced a Drug-related Serious TEAE.

5.5 Treatment Emergent Adverse Events (TEAE) Leading to Discontinuation of Study Treatment

Review TLFs: Table 14.3 – 1.1 Overview of Adverse Events

Two (2, 1.6%) subjects experienced at least one TEAE that led to discontinuation of study treatment (2 (3.1%) in AVP-786). One (1, 1.5%) subject experienced a Drug-related TEAE that led to the discontinuation of study treatment (one (1) subject in AVP-786).

5.6 TEAEs Leading to Death

Review TLFs: Table 14.3 – 1.1 Overview of Adverse Events

No subject experience a TEAE that lead to death.

6 EFFICACY ANALYSIS

1. Primary Efficacy Analysis - PANSS Marder Negative Factors Score

Review TLFs: Table 14.2-1.1.1.1 PANSS Marder Negative Factors Score: Change from Baseline MMRM, Observed Date; CCI

Based on the results from Table 14.2 – 1.1.1 PANSS Marder Negative Factors Score: Change from Baseline MMRM, Observed Data. One hundred thirteen (113) subjects have baseline PANSS Marder Negative Scores, 96 subjects had Week 15 scores available (51 in AVP-786, and 45 in Placebo arms).

Seventeen (17) subjects did not have the Visit 6 (Week 15) and therefore are excluded from this analysis. Table 4 contains the list of these 17 subjects that did not have a PANSS Marder Negative Factors Score at Visit 6 (Week 15).



DMC CLOSED SESSION EXECUTIVE SUMMARY

Table 4. List of Subjects Who Did not have a PANSS Marder Negative Factors Score at Visit 6 (Week 15)

Subject ID	PANSS Baseline Value	Last visit	PANSS Last Visit Value	Last value date
PPD	30	Visit 3 (Week 6)	29	PPD
	21	Visit 3 (Week 6)	22	
	24	Visit 5 (Week 12)	25	
	24	Visit 4 (Week 9)	19	
	29	Visit 4 (Week 9)	29	
	26	Visit 5 (Week 12)	23	
	17	Visit 4 (Week 9)	17	
	27	Visit 4 (Week 9)	24	
	24	Visit 4 (Week 9)	27	
	29	Visit 4 (Week 9)	24	
	24	Visit 3 (Week 6)	27	
	25	Visit 4 (Week 9)	25	
	22	Visit 6 (Week 15)	21	
	25	Baseline	25	
	25	Visit 3 (Week 6)	25	
	19	Visit 4 (Week 9)	13	
	23	Baseline	23	

Change from baseline to Visit 6 (Week 15) in Least Square Means (LS) was -2.3 (SE 0.47) in AVP-786 arm and -2.9 (SE 0.50) in placebo arm . The treatment difference was 0.6 with a 95% CI of -0.71 to 1.88 and a p-value of 0.371.

Table 5. PANSS Negative Factors Score- Change from Baseline and Treatment Difference

Visit/ Week	Parameter	AVP-786		Placebo	
		N	Mean (SD) or LS Mean (SE)	N	Mean (SD) or LS Mean (SE)
Baseline		60	25.2 (4.08)	53	25.1 (3.58)
Visit 3/ Week 6	Change from Baseline	58	-1.0 (2.86)	53	-1.0 (2.56)
	LS Change from Baseline		-1.1 (0.39)		-1.5 (0.42)
	Difference (95% CI)		0.4 (-0.60 – 1.41)		
	P-value		0.494		
Visit 4/ Week 9	Change from Baseline	57	-1.4 (3.25)	50	-1.9 (3.03)

CONFIDENTIAL



DMC CLOSED SESSION EXECUTIVE SUMMARY

	LS Change from Baseline		-1.4 (0.43)		-2.3 (0.46)
	Difference (95% CI)		0.9 (-0.29 – 2.08)		
	P-value		0.135		
Visit 5/ Week 12	Change from Baseline	52	-1.7 (3.21)	46	-2.1 (3.56)
	LS Change from Baseline		-1.8 (0.46)		-2.6 (0.49)
	Difference (95% CI)		0.8 (-0.51 – 2.01)		
	P-value		0.241		
Week 6/ Week 15	Change from Baseline	51	-2.2 (3.24)	45	-2.4 (3.45)
	LS Change from Baseline		-2.3 (0.47)		-2.9 (0.50)
	Difference (95% CI)		0.6 (-0.71 – 1.88)		
	P-value		0.371		

CCI

2. Sensitivity Analyses – PANSS Marder Negative Factors Score

Review TLFs:

Table 14.2 – 1.3.1 PANSS Marder Negative Factors Score: Sensitivity Analysis of MNAR using Pattern Mixture Model with Multiple Imputation (Delta Method) [Assume All Dropouts as MNAR] (mITT),

Table 14.2 – 1.3.2 PANSS Marder Negative Factors Score: Sensitivity Analysis of MNAR using Pattern Mixture Model with Multiple Imputation (Delta Method) [Assume All Dropouts Due to AE or LOE or Withdrew Consent as MNAR] (mITT),

Table 14.2 – 1.3.3 PANSS Marder Negative Factors Score: Sensitivity Analysis of MNAR using Pattern Mixture Model with Multiple Imputation (Delta Method) [Assume All Dropouts due to AE or LOE as MNAR] (mITT),

CCI

CONFIDENTIAL



DMC CLOSED SESSION EXECUTIVE SUMMARY

CCI

Secondary Analysis – NSA-16 Global Negative Symptom Score

Table 14.2 – 2.1 NSA-16 Global Negative Symptom Score: Change from Baseline MMRM, Observed Data (mITT)

NSA-16 results are available in the table 14.2 – 2.1. Treatment difference based on NSA-16 Global Negative Symptom Score at Visit 6 (Week 15) was 0.1 with a 95% CI of -0.20 to 0.36 and a p-value of 0.556.

CONFIDENTIAL

Page 13 of 13
ST-AD-018 version 04

Until the information herein is released by Otsuka to the public domain, the contents of this document are Otsuka confidential information and should not be duplicated or re-distributed without prior written consent of Otsuka.



This page is a manifestation of an electronically captured signature

SIGNATURE PAGE

Document Name: 18-AVP-786-207 DMC IA Executive Summary

Document Number: 1000266958

Signed by	Meaning of Signature	Server Date (dd-MMM- yyyy hh:min) - UTC timezone
PPD	Biostatistics Approval	21-Dec-2023 22:34:52