Official Title:	A Phase 3, multicenter, randomized, double-blind, placebo-controlled study to assess the efficacy, safety, and tolerability of AVP-786 (deuterated [d6]-dextromethorphan hydrobromide [d6-DM]/quinidine sulfate [Q]) for the treatment of agitation in patients with dementia of the Alzheimer's type
NCT Number:	NCT02442778
Document Date:	SAP Version 3.0: 26 April 2019

Statistical Analysis Plan

Sponsor:	Avanir Pharmaceuticals, Inc.
Protocol No:	15-AVP-786-302
Protocol Version No./ Date:	Amendment 4 28-Feb-2017
Title	A Phase 3, multicenter, randomized, double-blind, placebo-controlled study to assess the efficacy, safety, and tolerability of AVP-786 (deuterated [d6]-dextromethorphan hydrobromide [d6-DM]/quinidine sulfate [Q]) for the treatment of agitation in patients with dementia of the Alzheimer's type
CRF Version No./ Date:	6.0 18-Apr-2017
PRA Project ID:	AVA78601-786302
SAP Version No./ Date:	3.0 26-Apr-2019

Approvals

Prepared by	
PRA	
Biostatistician/ Title:	[redacted], MS / [redacted]
Signature/ Date:	
PRA	
Project Manager/ Title:	[redacted] / [redacted]
Signature/ Date:	
Sponsor	
Sponsor Name:	Avanir Pharmaceuticals, Inc.
Biostatistician / Title:	[redacted], MS / [redacted], Avanir [redacted]
Signature/ Date:	
CDL / Title:	[redacted], MD / [redacted], Avanir [redacted]
Signature/ Date:	
CSM / Title:	[redacted], Avanir [redacted]
Signature/ Date:	
Non-Clinical Dev Lead / Title:	[redacted], PhD / [redacted], Avanir [redacted]
Signature/ Date:	

Table of Contents

1.0 Introduction

This statistical analysis plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Avanir Pharmaceuticals, Inc. Protocol 15-AVP-786-302.

This SAP should be read in conjunction with the study protocol and case report form (CRF). This version of the plan has been developed using the Protocol Amendment 4 dated 28-Feb-2017 and CRF dated 18-Apr-2017. Any further changes to the protocol or CRF may necessitate updates to the SAP.

1.1 Changes from Protocol

Not applicable.

2.0 Study Objectives

The objectives of the study are to evaluate the efficacy, safety, and tolerability of AVP-786 compared to placebo for the treatment of agitation in patients with dementia of the Alzheimer's type.

3.0 Study Design

3.1 General Description

This is a Phase 3, multicenter, randomized, double-blind, placebo-controlled study with a treatment duration of 12 weeks. Table 1 provides an overview of the study duration:

TABLE 1. OVERVIEW OF STUDY DURATION

	Screening	Treatment Period	Follow-up
Duration in Days	Days -28 to -1	Days 1 to 85	30 days after last dose

Approximately 470 patients with a diagnosis of probable Alzheimer's disease (AD) and clinically meaningful moderate/severe agitation secondary to AD will be enrolled at approximately 75 centers in North America. There will be 8 scheduled clinic visits including a screening visit and 2 safety follow-up phone visits in the study. Clinic visits include Screening, Baseline (Day 1), and Weeks 1, 2, 3, 6, 9, and 12. Safety follow-up phone calls will occur on Week 4 and Week 10. An in-clinic Follow-up visit will occur 30 days after last dose of study medication for subjects who terminate early. Patients who do not roll over to the extension study (Study 15-AVP-786-303) will receive a safety follow-up call 30 days after the last dose of study medication. The schedule of events is provided in Table 2:

TABLE 2. SCHEDULE OF EVENTS AND VISITS

Procedure	Visit:	Screening[1]	Baseline	Visit 2[1]	Visit 2.1[1]	Visit 3[1,19]	Phone Call[1,2,19]	Visit 4[1,19]	Visit 5[1]	Phone Call[1,2]	Visit 6[1,4]/ET[3]	Follow-up Visit[3]
	Study Day:	Day -28 to -1	Day 1	Day 8	Day 15	Day 22	Day 29	Day 43	Day 64	Day 71	Day 85	
	End of Study Week:	Week -4 to -1		Week 1	Week 2	Week 3	Week 4	Week 6	Week 9	Week 10	Week 12	
Sign informed consent forms		X										
Medical history		X										
Review of eligibility[5]		X	X									
Randomization			X									
Physical and neurological examination		X						X			X	
Vital signs and weight		X	X[6]	X	X	X		X	X		X[6]	
ADCS-CGIC-Overall			X[7]					X			X	
CGIS-Agitation		X	X					X			X	
mADCS-CGIC-Agitation			X[8]					X			X	X
Risk assessment for falls (worksheet and TUG test)		X						X[9]			X[9]	
ECG		X[10]	X[11]			X		X	X		X	
AEs			X	X	X	X	X	X	X	X	X	X
Prior and concomitant medications, nondrug therapies, and nonpharmacological interventions for agitation		X	X	X	X	X	X	X	X	X	X	X
MMSE		X	X					X			X	
GMHR		X									X	
CMAI		X	X	X	X	X		X	X		X	X
NPI		X[12]	X	X[12]	X[12]	X		X	X		X	
CSDD		X						X			X	
ZBI			X					X			X	
DEMQOL[13]			X					X			X	
ADAS-cog[14]			X					X			X	
PGIC[15]								X			X	
RUD			X					X			X	
ESS[14]			X						X		X	
S-STS		X	X	X		X		X	X		X	

Procedure	Visit: / Study Day: / End of Study Week:	Screening[1] Day -28 to -1 Week -4 to -1	Baseline Day 1	Visit 2[1] Day 8 Week 1	Visit 2.1[1] Day 15 Week 2	Visit 3[1,19] Day 22 Week 3	Phone Call[1,2,19] Day 29 Week 4	Visit 4[1,19] Day 43 Week 6	Visit 5[1] Day 64 Week 9	Phone Call[1,2] Day 71 Week 10	Visit 6[1,4]/ET[3] Day 85 Week 12	Follow-up Visit[3]
Administer morning dose of study medication in clinic			X	X[16]	X[16]	X		X	X		X	X
Chemistry, hematology, and urinalysis		X[17]				X		X	X		X[17]	
Urine pregnancy test[18]		X	X					X			X	
PK blood sample								X			X	
[redacted]												
Dispense study drug and diary card			X			X		X	X			
Review and return unused study medication and diary card				X[16]	X[16]	X		X	X		X	

ADAS-cog = Alzheimer's Disease Assessment Scale-cognitive subscale; ADCS-CGIC-Overall = Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change for Overall Clinical Status; AE = adverse event; CGIS-Agitation = Clinical Global Impression of Severity of Illness scale for Agitation; CMAI = Cohen-Mansfield Agitation Inventory; CSDD = The Cornell Scale for Depression in Dementia; DEMQOL = Dementia Quality of Life scale; ECG = electrocardiogram; ET = early termination; GMHR = General Medical Health Rating; mADCS-CGIC-Agitation = modified Clinical Global Impression of Severity of Illness scale for Agitation; MMSE = Mini-Mental State Examination; NPI = Neuropsychiatric Inventory; PGIC = Patient Global Impression of Change rated by the caregiver; PK = pharmacokinetics; RUD = Resource Utilization in Dementia; S-STS = Sheehan Suicidality Tracking Scale; TUG = Timed Up and Go; ZBI = Zarit Burden Interview

Note: Whenever possible, each patient and caregiver should have the rating scales administered by the same raters throughout the study, for consistency of ratings. The following scales MUST be administered by the same rater at each visit. CMAI, NPI, mADCS-CGIC-Agitation, and CGIS-Agitation.

1 Study visits have a +/- 3-day window except Screening, Visit 2, and phone calls. Screening, Visit 2, and phone calls (excludes follow-up phone calls for ET patients) have a +3-day window. The screening period may be extended after discussion with and approval by the medical monitor.

2 Phone call should be made to patient/caregiver to collect adverse events and query on concomitant medication use

3 Early termination visit for patients who withdraw prior to study completion. Patients who terminate early from the study will receive daily phone calls for 5 consecutive days following the ET visit to query on their overall well-being and an in-clinic Follow-up visit 30 days after last dose of study medication for selected safety and efficacy assessments.

4 Patients who do not roll over to the extension study (Study 15-AVP-786-303) will receive a safety follow-up phone call 30 days after the last dose of study medication.

5 For each patient, a protocol eligibility form will be completed

6 Weight should be measured at the Baseline visit and Visit 6

7 The ADCS-CGIC-Overall baseline evaluation worksheet should be completed to record baseline information for assessing change at Visits 4 and 6

8 The mADCS-CGIC-Agitation baseline evaluation worksheet should be completed to record baseline information for assessing change at Visits 4 and 6

9 Only the TUG test should be performed for risk assessment of falls at Visits 4 and 6

10 ECG to be performed in triplicate at Screening visit

11 ECG to be performed pre-dose and post-dose

Procedure	Visit:	Screening[1]	Baseline	Visit 2[1]	Visit 2.1[1]	Visit 3[1,19]	Phone Call[1,2,19]	Visit 4[1,19]	Visit 5[1]	Phone Call[1,2]	Visit 6[1,4]/ET[3]	Follow-up Visit[3]
	Study Day:	Day -28 to -1	Day 1	Day 8	Day 15	Day 22	Day 29	Day 43	Day 64	Day 71	Day 85	
	End of Study Week:	Week -4 to -1		Week 1	Week 2	Week 3	Week 4	Week 6	Week 9	Week 10	Week 12	

12 Only the Agitation/Aggression domain of the NPI should be performed at the Screening visit, Visit 2, and Visit 2.1.

13 The proxy version is to be rated by the caregiver. The non-proxy version is to be rated only by patients with an MMSE score of ≥ 10 at baseline

14 ADAS-cog and ESS are to be rated only by patients with an MMSE score of ≥ 10 at baseline

15 PGIC is to be rated by the caregiver

16 The morning dose of study medication can be administered at home if the visit will occur within 2 hours of dosing; the time of dosing should be noted by the patient/caregiver. The blister card and diary card should be brought to the clinic and returned to the patient/caregiver after reviewing for compliance

17 Thyroid function tests (TSH, and reflex T3 and T4 if TSH is abnormal) should be performed at the Screening visit. Glycosylated hemoglobin (HbA1c) test should be performed at the Screening visit and Visit 6.

18 Urine pregnancy test to be performed for females of child bearing potential only

19 A one-time downward dose adjustment is allowed after Visit 3 (Week 3) up to and including Visit 4 (Week 6), ie Day 23 to Day 43. Patient will need to return to the clinic for an unscheduled visit for safety assessments

3.2 Treatment Assignments

Following screening procedures for assessment of inclusion and exclusion criteria, eligible patients will be randomized into the study at Baseline to receive AVP-786-28/4.9 (deudextromethorphan hydrobromide/ quinidine sulfate; also referred to as AVP-786 28 mg), AVP-786-42.63/4.9 (also referred to as AVP-786 42.63 mg), or placebo. Prior to protocol amendment 4, patients were randomized in a [redacted] ratio to receive AVP-786-28/4.9:Placebo. After protocol amendment 4 when approximately 100 patients had been enrolled, patients are randomized in a [redacted] ratio to receive AVP-786-28/4.9:AVP-786-42.63/4.9:Placebo. Randomization is stratified by the Neuropsychiatric Inventory (NPI) Agitation/Aggression (AA) domain score (≤ 6 vs. > 6), risk assessment for falls (normal/mild vs. moderate/severe), and concomitant use of antipsychotic medications (yes or no). Study medication (active or placebo) will be administered orally twice daily (BID) with approximately a 12-hour separation between doses throughout the treatment period. Dosing schedules for each treatment group are provided below:

- Patients randomized to receive AVP-786-28/4.9 will start with AVP-786 18 mg once a day in the morning and placebo in the evening for the first 7 days of the study. From Day 8, patients will receive AVP-786 18 mg BID for 14 days. From Day 22, patients will receive AVP-786-28/4.9 BID for the remaining 9 weeks of the study. If deemed necessary by the investigator, a one-time downward dose adjustment to AVP-786 18 mg will be allowed after Visit 3 up to and including Visit 4 (i.e., Day 23 to Day 43). The patients will remain on the lower dose of study medication for the remaining study duration.
- Patients randomized to receive AVP-786-42.63/4.9 will start with AVP-786-28/4.9 once a day in the morning and placebo in the evening for the first 7 days of the study. From Day 8, patients will receive AVP-786-28/4.9 BID for 14 days. From Day 22, patients will receive AVP-786-42.63/4.9 BID for the remaining 9 weeks of the study. If deemed necessary by the investigator, a one-time downward dose adjustment to AVP-786-28/4.9 will be allowed after Visit 3 up to and including Visit 4 (i.e., Day 23 to Day 43). The patients will remain on the lower dose of study medication for the remaining study duration.
- Patients randomized to receive placebo will be dosed with placebo throughout the study duration.

3.3 Sample Size Considerations

Power calculation was performed assuming a normal distribution for the primary efficacy endpoint. [redacted]

The study protocol was amended to include another dose arm (AVP-786-42.63/4.9) with 140 patients in the treatment arm, when approximately 100 patients have already been enrolled (50 AVP-786-28/4.9 patients and 50 placebo patients) in the study. To maintain an approximate [redacted] chance for an incoming patient to receive active drug, the sample size for the placebo treatment group was amended to 190. Overall, the total sample size of 470 patients will include 140 patients each for AVP-786-28/4.9 and AVP-786-42.63/4.9 groups and 190 patients in the placebo group. This sample size will provide approximately [redacted] power to detect a treatment effect size of [redacted] between AVP-786-42.63/4.9 and the concurrent placebo treatment group.

3.4 Randomization

Eligible patients will be randomized to receive AVP-786-28/4.9, AVP-786-42.63/4.9, or matching placebo capsules on Day 1 (Baseline) in a double-blind manner according to a randomization scheme generated by Avanir or its representative and managed within an Interactive Web Response System (IWRS). Prior to protocol amendment 4, patients were randomized in a [redacted] ratio to receive AVP-786-28/4.9:Placebo.

After protocol amendment 4 when approximately 100 patients had been enrolled, patients are randomized in a [redacted] ratio to receive AVP-786-28/4.9:AVP-786-42.63/4.9:Placebo. The randomization are stratified by NPI AA domain score (≤ 6 vs. > 6), risk assessment for falls (normal/mild vs. moderate/severe), and concomitant use of antipsychotic medications (yes vs. no).

4.0 Study Endpoints and Covariates

4.1 Scales and Questionnaires

4.1.1 Cohen-Mansfield Agitation Inventory (CMAI)

The CMAI will be used as the primary efficacy measure in this study. The CMAI is used to assess the frequency of manifestations of agitated behaviors in elderly persons. The responses are based on the 2 weeks preceding the assessment of the CMAI and will be assessed at Screening, Baseline, Week 1, Week 2, Week 3, Week 6, Week 9, Week 12, and Follow-up visit (for ET patients).

The CMAI consists of 29 agitated behaviors that are further categorized into 3 distinct agitation factors (dimensions, Rabinowitz 2005), also known as CMAI factors of agitation. [redacted]

[redacted]

Each of the [redacted] items is rated on a 7-point scale of frequency consisting of the following:

CMAI Item Scores

- 1 = Never
- 2 = Less than once a week but still occurring
- 3 = Once or twice a week
- 4 = Several times a week
- 5 = Once or twice a day
- 6 = Several times a day
- 7 = Several times an hour

The CMAI total score is calculated as the sum of ratings for all 29 items and range from 29 to 203. The CMAI total score will be unevaluable if less than 24 of the 29 items have recorded responses. If 24 to 28 items are recorded, then the total score will be the mean of the recorded items multiplied by 29 rounded to the first decimal place. [redacted].

[redacted]

[redacted]

4.1.2 Neuropsychiatric Inventory (NPI)

The NPI is a validated clinical instrument for evaluating psychopathology in a variety of disease settings, including dementia. The NPI will be administered to the patient's caregiver at Baseline, Week 3, Week 6, Week 9, and Week 12. Additionally, the Agitation/Aggression domain of the NPI will be administered to the patient's caregiver at Screening, Week 1, and Week 2.

The NPI is a retrospective caregiver-informant interview covering 12 neuropsychiatric symptom domains:

- Delusions
- Hallucinations
- Agitation/Aggression
- Depression/Dysphoria
- Anxiety
- Elation/Euphoria
- Apathy/Indifference
- Disinhibition
- Irritability/Lability
- Aberrant Motor Behavior
- Sleep and Nighttime Behavioral Disorders
- Appetite/Eating Disorders

The scripted NPI interview includes a compound screening question for each symptom domain, followed by a list of interrogatives about domain-specific behaviors that is administered when a positive response to a screening question is elicited. Neuropsychiatric manifestations within a domain are collectively rated by the caregiver in terms of frequency and severity, yielding a domain score. Domain scores range from 1 to 12 and are calculated using the following formula:

domain score = frequency x severity

where, frequency and severity are rated according to the following scales:

Frequency
- 1 = Rarely – less than once per week
- 2 = Sometimes – about once per week
- 3 = Often – several times per week but less than every day
- 4 = Very often – once or more per day

Severity
- 1 = Mild – produces little distress in the patient
- 2 = Moderate – more disturbing to the patient but can be redirected by the caregiver
- 3 = Severe – very disturbing to the patient and difficult to redirect

Frequency and severity rating scales have defined anchor points to enhance the reliability of caregiver responses. Caregiver distress is also rated for each positive neuropsychiatric symptom domain using the following anchored scores:

Distress
- 0 = Not at all
- 1 = Minimally (almost no change in work routine)
- 2 = Mildly (some change in work routine but little time rebudgeting required)
- 3 = Moderately (disrupts work routine, requires time rebudgeting)
- 4 = Severely (disruptive, upsetting to staff and other residents, major time infringement)

- 5 = Very Severely or Extremely (very disruptive, major source of distress for staff and other residents, requires time usually devoted to other residents or activities)

The NPI nursing-home (NPI-NH) version will be used for patients from in-patient or assisted living facilities. The questions in the NPI-NH are rephrased for professional caregivers who may not know the patient prior to the onset of illness however, the overall instrument domains and scoring is identical to the NPI except for the caregiver distress section which is replaced with occupational disruptiveness in the NPI-NH version.

A total NPI score is calculated by adding the scores of all 12 domain scores together. The possible total scores are from 0 to 144. The NPI Total Score will be unevaluable if less than 10 of the 12 items are recorded. If 10 or 11 of the 12 items are available, then the total score is the mean of the available scores times 12. All imputed scores are rounded to the first decimal place. The distress score is not included in the total NPI score.

The total caregiver distress score is generated by adding together the scores of all 12 domains of the NPI distress questions. Scores range from 0 to 60 with a higher score indicating more distress. Total score calculation with missing items follows the same algorithm as the NPI total score.

4.1.3 Clinical Global Impression of Severity Illness-Agitation (CGIS-Agitation)

The CGIS-Agitation is an observer-rated scale that measures illness severity and will be assessed at Screening, Baseline, Week 6, and Week 12.

The clinician is asked to rate the patient relative to their past experience with other patients with the same diagnosis, with or without collateral information. The CGIS-Agitation is a 7-point scale and will be assessed for severity of agitation. A value of 0 is given to patients who are not assessed. The 7-point scale for CGIS-Agitation is the following:

Severity of Agitation Scores

- 0 = Not assessed
- 1 = Normal, not at all ill
- 2 = Borderline ill
- 3 = Mildly ill
- 4 = Moderately ill
- 5 = Markedly ill
- 6 = Severely ill
- 7 = Among the most extremely ill patient

4.1.4 Alzheimer's Disease Cooperative Study – Clinical Global Impression of Change Rating (ADCS-CGIC-Overall)

The ADCS version of the Clinical Global CGIC, referred to as ADCS-CGIC-Overall, provides a reliable means to assess change from a baseline level of global function within the timeframe of a clinical trial. The ADCS-CGIC-Overall focuses on clinicians' observations of change in the patient's cognitive, functional, and behavioral performance since the beginning of the trial. Once the baseline level of severity is established, the change score at the follow-up visits is based on information gathered from the patient and caregiver interviews by an independent evaluator. The response for the ADCS-CGIC-Overall will consist of one of the following:

Overall Clinical Status Responses

- 1 = Marked Improvement
- 2 = Moderate Improvement
- 3 = Minimal Improvement
- 4 = No Change
- 5 = Minimal Worsening

- 6 = Moderate Worsening
- 7 = Marked Worsening

The baseline ADCS-CGIC-Overall evaluation will be conducted at the Baseline visit. The ADCS-CGIC-Overall will be assessed at Week 6 and Week 12 for change in Overall Clinical Status. At Week 6, change from the Baseline visit will be assessed. At Week 12, change from Week 6 and change from the Baseline visit will be assessed.

4.1.5 Modified Alzheimer's Disease Cooperative Study – Clinical Global Impression of Change Rating (mADCS-CGIC-Agitation)

The mADCS-CGIC-Agitation is a modified version of the ADCS-CGIC containing additional questions related to agitation and an assessment of the Clinician's Impression of Change focused specifically on agitation. The response for the mADCS-CGIC-Agitation will consist of one of the following:

Responses
- 1 = Marked Improvement
- 2 = Moderate Improvement
- 3 = Minimal Improvement
- 4 = No Change
- 5 = Minimal Worsening
- 6 = Moderate Worsening
- 7 = Marked Worsening

The baseline mADCS-CGIC-Agitation evaluation will be conducted at the Baseline visit. The mADCS-CGIC-Agitation will be assessed at Week 6, Week 12, and Follow-up visit (for ET patients) for change in agitation. At Week 6, change from the Baseline visit will be assessed. At Week 12, change from Week 6 and change from the Baseline visit will be assessed. At the Follow-up visit (for ET patients), change from Week 12 will be assessed.

4.1.6 Zarit Burden Interview (ZBI)

The ZBI is a 22-item scale used to assess the impact of patient's disabilities on the caregiver's life and will be assessed at Baseline, Week 6, and Week 12. It is designed to reflect the burden experienced by caregivers of dementia patients and can either be completed by the caregiver or administered as an interview. Each question is rated on the following 5-point scale:

Zarit Burden Interview Scores for Individual Questions
- 0 = Never
- 1 = Rarely
- 2 = Sometimes
- 3 = Quite Frequently
- 4 = Nearly Always

The ZBI is scored by summing the responses of the individual questions and ranges from 0 to 88. Higher scores indicate greater caregiver distress. The ZBI score will be unevaluable if 4 or more of the 22 items are missing. If 19 to 21 of the 22 items are available, then the total score is the mean of the available scores times 22. All imputed scores are rounded to the first decimal place.

4.1.7 Patient Global Impression of Change (PGIC)

The PGIC is a 7-point scale used to assess treatment response and will be assessed and rated by the patient's caregiver at Week 6 and Week 12 and will focus on the patient's agitation. The response values for the PGIC are provided below:

Patient Global Impression of Change Scale
- 1 = Very much improved

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

4.1.8 Dementia Quality of Life (DEMQOL)

The DEMQOL is a scale used to evaluate health related quality of life in patients with dementia and their caregivers. There are 2 versions of the DEMQOL, a 29-item version (rated by the patient) and a 32-item version (DEMQOL-proxy, rated by the caregiver). The following provides the response categories for the individual DEMQOL questions and the overall quality of life question:

DEMQOL Responses for Individual Questions
- 1 = A lot
- 2 = Quite a bit
- 3 = A little
- 4 = Not at all

Overall Quality of Life Question
- 1 = Very good
- 2 = Good
- 3 = Fair
- 4 = Poor

The DEMQOL-proxy (and DEMQOL for patients with a Mini Mental State Examination (MMSE) score ≥ 10 at the Baseline visit) will be assessed at Baseline, Week 6, and Week 12.

Scores for DEMQOL Item 1, 3, 5, 6, 10, 29, DEMQOL-proxy item 1, 4, 6, 8, 11, 32, need to be reversed for analysis, i.e., 4=A lot, 3=Quite a bit, 2=A little, 1=Not at all. Total score is derived by sum of all item scores with item 29 in DEMQOL and item 32 in DEMQOL-proxy being excluded. Missing item handling for DEMQOL, 1) count the number of missing items for questions 1 to 28, 2) if the number < 15, set the missing item values to the mean of the non-missing item 1 to 28, then add them up; if the number of missing >= 15, then the total is missing. Missing item handling for DEMQOL-proxy, 1) count the number of missing items for questions 1 to 31, 2) if the number < 16, set the missing item values to the mean of the non-missing item 1 to 31, then add them up; if the number of missing >= 16, then the total is missing. In addition, DEMQOL Item 29 and DEMQOL-proxy Item 32 will be summarized separately.

4.1.9 Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-cog)

The ADAS was designed to evaluate the cognitive and non-cognitive behavioral dysfunction characteristics of patients with AD and will be assessed at Baseline, Week 6, and Week 12 for patients with an MMSE score ≥ 10 at the Baseline visit. The cognitive subscale (ADAS-cog) consists of 11 subsets related to memory, praxis, and language. Scoring of the individual subsets is specific to each particular subset. The ADAS-cog is scored by summing the 11 subset scores and ranges in scale from 0 to 70. Scoring for each component of the ADAS-cog is as follows:

Component	Measurement	Scoring
Word Recall Task	Mean number of words not recalled on 3 trials	Maximum score = 10
Naming Task	Number of objects named incorrectly	0 (0 – 2 items) to 5 (15-17 items)
Commands	Number of correct and incorrect commands	0 (all commands correct) to 5 (all commands incorrect)
Constructional Praxis	Number of drawings correct	0 (all drawings correct) to 5 (no figures drawn, no recognizable attempt at drawing any side/section of any figure)
Ideational Praxis	Number of components performed correctly	0 (all components performed correctly) to 5 (failure to perform 5 components)

Orientation	One point given for each incorrect response	Maximum score = 8
Word Recognition	Mean number of words not recognized on 3 trials	Maximum error score = 12
Remembering Test Instructions	Ability to remember the requirements of the Word Recognition Task	0 (patient never needs extra reminders) to 5 (severe – must be reminded 7 or more times)
Spoken Language Ability	Quality of speech	0 (no instances of difficulty understanding the patient) to 5 (severe – 1 or 2 word utterance; fluent but empty speech; mute)
Word-finding Difficulty	Ability in expressive language	0 (no evidence) to 5 (severe – nearly total loss of content words; speech sound empty; 1 or 2 word utterances)
Comprehension of Spoken Language	Ability to understand what is being said to them	0 (patient understands) to 5 (severe – patient rarely responds to questions appropriately; not due to poverty of speech)

4.1.10 Cornell Scale for Depression in Dementia (CSDD)

The CSDD was developed to assess signs and symptoms of major depression in patients with dementia and will be assessed at Screening, Week 6, and Week 12. Information is elicited through one interview with the caregiver and one interview with the patient. The CSDD has 19 items and each item is rated for severity on the following scale:

Severity Scale for CSDD Items

- 0 = Absent
- 1 = Mild or intermittent
- 2 = Severe

The CSDD score is calculated by summing the non-missing scores from each item score. Items which could not be evaluated are assigned value 0 (Psych Congress Network: https://www.psychcongress.com/saundras-corner/scales-screeners/depression/cornell-scale-depression-dementia-csdd). Scores above 10 indicate probable major depression. Scores above 18 indicate definite major depression. Scores below 6 as a rule are associated with absence of significant depressive symptoms.

4.1.11 Resource Utilization Dementia (RUD)

The RUD is used to calculate healthcare costs associated with dementia and will be assessed at Baseline, Week 6, and Week 12.

The RUD evaluates dementia patients' utilization of formal and informal healthcare resources, including hospitalizations and doctor visits, living assistance, and time spent by nonprofessional caregivers. The RUD is administered as a semi-structured interview with the patient's primary caregiver. One section focuses on caregiver impact (loss of work and leisure time incurred by the caregiver) and the other section focuses on the patient's use of healthcare resources. The total healthcare costs associated with the patient's dementia is estimated by multiplying the number of units used (e.g., hours of caregiver time, visits to doctors, nights in accommodation) by the corresponding unit price vector.

4.1.12 General Medical Health Rating (GMHR)

The GMHR is a global clinical rating designed to quantify the severity of general co-morbidity in a patient with dementia and will be assessed at Screening and Week 12.

The GMHR is rated according to the following scale:

GMHR Rating Scale
- 1 = Poor
- 2 = Fair
- 3 = Good
- 4 = Excellent to very good

4.2 Efficacy Endpoints

4.2.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the change from Baseline to Week 12 in the CMAI total score.

4.2.2 Secondary Efficacy Endpoints

Secondary efficacy endpoints include change from Baseline to Week 12 for the following efficacy measures:

- mADCS-CGIC-Agitation
- NPI - Agitation/Aggression domain score
- NPI – Agitation/Aggression domain Caregiver Distress score
- NPI – Aberrant Motor Behavior domain
- ZBI
- NPI – Irritability/Lability domain
- Total NPI
- CGIS-Agitation
- ADCS-CGIC-Overall
- PGIC
- DEMQOL
- CSDD
- ADAS-cog
- RUD
- GMHR (change from Baseline to Week 12)

4.3 Safety Endpoints

Safety and tolerability measurements of AVP-786 will include: adverse events (AEs), physical and neurological examinations, vital signs, clinical laboratory assessments, resting 12-lead electrocardiograms (ECGs), Sheehan Suicidality Tracking Scale (S-STS), MMSE, Time Up and Go (TUG) test, and the Epworth Sleepiness Scale (ESS).

4.4 Pharmacokinetic and Genotype Data

Blood samples will be collected between 0 to 3 hours after the morning dose of study medication on Week 6 and Week 12 for analysis of plasma levels of d6-dextromethorphan (DM), d6-DM metabolites, and quinidine (Q).

[redacted]

5.0 Analysis Populations

Three analysis populations will be used, modified intent-to-treat (mITT), intent-to-treat (ITT), and Safety.

5.1 mITT Population

The mITT population will include all patients randomized in the study who had at least one post-baseline efficacy assessment. The mITT population will be used for all efficacy analyses health outcome analyses. Patients will be included in the treatment group to which they were randomized regardless of treatment received.

5.2 ITT Population

The ITT population includes all randomized patients in the study. The ITT population will be used for sensitivity analyses. Patients will be included in the treatment group to which they were randomized regardless of treatment received.

5.3 Safety Population

The Safety population includes all patients who received at least one dose of study medication. The Safety population will be used for all analyses of safety data. Patients will be included in the treatment group based on the actual treatment received.

6.0 Definitions

Age

The following SAS® code will be used to calculate patient age (years):

Age = floor ([intck{‘month’, birth date, screen date} - {day(informed consent date) < day(birth date)}] / 12), where intck is a SAS® function counting integer days.

Baseline

Baseline is generally defined as the last assessment prior to the first dose of study drug.

Change from Baseline

Change from baseline (CFB) will be calculated as (post-baseline – baseline). CFB 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 CFB will not be calculated for that parameter. Patients with missing CFB values will be excluded from analyses in which CFB is the endpoint.

Percent CFB, when needed, will be calculated by dividing CFB by the baseline value multiplied by 100. Patients with a value of 0 at baseline will not have a percent CFB calculated.

Concomitant medications
Concomitant medications are defined as any medications taken on or after the date of first dose of study drug.

Discontinuation of study
A patient will be considered discontinued from the study when a Study Exit CRF page is completed indicating primary reason for discontinuation.

End of Treatment (EOT)

EOT is defined as the last value for a given patient, whenever it occurred.

Enrolled patient
An enrolled patient is one with a record in the database that is not a screen failure.

Prior medication
Prior medications are defined as any medications with start and stop dates prior to the date of first dose of study drug. Medications are defined as prior or concomitant, but not both (see also definition of concomitant medication in this section).

Protocol deviation
Potential planned or unplanned protocol deviations noted during clinical monitoring will be documented by category (i.e., inclusion criteria, exclusion criteria, study drug, safety assessment, efficacy assessment, visit window, informed consent, prohibited medication, and other). All deviations will be reviewed, categorized, and finalized prior to database lock.

Relationship to study treatment
AEs related to study treatment will be defined as those entered as possibly related or related. Unlikely related or not related entries will be considered not related to study treatment. If relationship is missing, then it will be considered as related to study treatment.

Study day
Study day is defined relative to the date of the first dose of study drug. For assessments that occur after this visit date, study day is calculated as (assessment date – study drug first dose date + 1). For assessments that occur prior to study drug first dose date, study day will be calculated as (assessment date – study drug first dose date); there is no Study Day 0.

Treatment-emergent adverse event
An AE will be considered a treatment-emergent AE (TEAE) if it begins or worsens on or after the first dose date and before the last dose date + 30 days.

7.0 Interim Analyses

There is no Interim analysis planned for this study.

8.0 Data Handling and Review

8.1 Visit Windows

Data at scheduled visits will be assigned to analysis visits as defined in the Visit Window tables below. Visit Windows will be used to classify observed data including unscheduled and early termination visits using study days. If 2 or more visits occur within the same analysis window, then the values closest to the target day will be used. If the assessment is the same distance from the target day, then the latest one will be used.

TABLE 3-1. VISIT WINDOWS FOR ASSESSMENTS ON-TREATMENT AT WEEKS 1, 2, 3, 6, 9, AND 12

Week/Visit	Target Day	Study Days
Week 1/Visit 2	8	2, 11
Week 2/Visit 2.1	15	12, 18
Week 3/Visit 3	22	19, 32
Week 6/Visit 4	43	33, 53
Week 9/Visit 5	64	54, 74
Week 12/Visit 6	85	75, 95

Note: for CMAI, NPI AA.

TABLE 3-2. VISIT WINDOWS FOR ASSESSMENTS ON-TREATMENT AT WEEKS 2, 3, 6, 9, AND 12

Week/Visit	Target Day	Study Days
Week 2/Visit 2.1	15	2, 18
Week 3/Visit 3	22	19, 32
Week 6/Visit 4	43	33, 53
Week 9/Visit 5	64	54, 74
Week 12/Visit 6	85	75, 95

Note: for ECG

TABLE 3-3. VISIT WINDOWS FOR ASSESSMENTS ON-TREATMENT AT WEEKS 3, 6, 9, AND 12

Week/Visit	Target Day	Study Days
Week 3/Visit 3	22	2, 32
Week 6/Visit 4	43	33, 53
Week 9/Visit 5	64	54, 74
Week 12/Visit 6	85	75, 95

Note: for NPI other than NPI-AA.

TABLE 3-4. VISIT WINDOWS FOR ASSESSMENTS ON-TREATMENT AT WEEKS 6 AND 12

Week/Visit	Target Day	Study Days
Week 6/Visit 4	43	2, 64
Week 12/Visit 6	85	65, 95

Note: for mADCS-CGIC-Agitation, ZBI, CGIS-Agitation, ADCS-CGIC-Overall, PGIC, DEMQOL, CSDD, ADAS-cog, RUD.

TABLE 3-5. VISIT WINDOWS FOR ASSESSMENTS ON-TREATMENT AT WEEK 12 ONLY

Week/Visit	Target Day	Study Days
Week 12/Visit 6	85	2, 95

Note: for GMHR.

8.2 Missing Data Conventions

Missing data will be handled differently depending on the parameter and analysis. Analyses that are done on 'observed cases' will not follow any imputation rules below. Imputation rules are as follows:

- Missing baseline values will not be imputed in any situation.
- A last observation carried forward (LOCF) and a worst observation carried forward (WOCF) + LOCF approach will be used for efficacy analyses, where specified. The table below provides imputation rules for LOCF and WOCF.

Missing Imputation for LOCF and WOCF		
Missing Value	LOCF	WOCF + LOCF [1]

Missing at scheduled visits	Impute with last non-missing value.	Missing due to lack of efficacy – Impute with worst value after Day 1. If there are no values post Day 1, then impute with Day 1. Missing due to other reasons - Impute with last non-missing value.
[1] Only applicable for the primary efficacy endpoints.		

- Any safety assessments for patients without a Week 12 value will use the last non-missing post-baseline observation as EOT.
- Missing post-baseline values for by-visit efficacy and safety data will be summarized using the Visit Windows from Section 8.1. If a value is not available within a given window, then no imputation will be done.
- Missing data for AE relationship will be imputed as “Related”.
- Rules for partial dates are provided in Appendices 2 and 3. These rules will apply to adverse events and medications.

8.3 Treatment Misallocations

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

Safety data will be summarized, “as treated”, according to the treatment the patient actually received.

8.4 Data Handling and Transfer

Data will be entered by investigational sites into a clinical database built with Bioclinica and exported as SAS® version 9.4 or higher datasets (SAS Institute, Inc., Cary, NC). Converted datasets are created using SAS® and following Clinical Data Interchange Standards Consortium (CDISC) Standard Data Tabulation Model conventions (v3.1.3 implementation guide v1.3). Derived analysis datasets are generated using SAS® and following standard CDISC Analysis Dataset Model conventions (implementation guide v1.0). Data analyses including summary tables, figures, and listings (TFLs) are produced using SAS®.

Medical history and AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 18.1 to assign a system organ class (SOC) and preferred term (PT) to each AE. AEs severity will be graded on a 3-point scale and reported in detail as indicated on the eCRF.

Prior and concomitant medications are coded to preferred drug names using the World Health Organization Drug Dictionary Enhanced (WHODRUG DDE, 2015SEP01).

8.5 Data Screening

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

Review of a pre-freeze TFL run (Dry Run#1) on clean patients and a post-freeze TFL run (Dry Run#2) on the frozen database allow for further data screening prior to database lock (DBL). The post-freeze TFLs will be discussed with AVANIR in a data review meeting to identify any final data issues and seek corrections prior to DBL. The PRA statistician and AVANIR must approve database lock.

8.6 Data Safety Monitoring Board (DSMB)

This study is monitored on an ongoing basis by an external data safety monitoring board (DSMB). Safety analyses will be provided to the DSMB approximately every 3 months. Select efficacy results may be included, although no p-values will be provided. Details of the DSMB can be found in the DSMB charter and DSMB SAP.

9.0 Overall Statistical Considerations

9.1 Summary Statistics

All analyses will use SAS® version 9.4 or higher. Summary tables will be organized by treatment group. All available data will be used in the analyses, and important data will be included in data listings, sorted by treatment group, patient, and by visit within patient.

Unless otherwise noted, categorical data will be presented using counts and percentages, with the number of patients in the analysis population by treatment group as the denominator for percentages. Percentages are rounded to one decimal place except 100% will be displayed without any decimal places and percentages will not be displayed for zero counts. Continuous data, unless otherwise noted, will be summarized using the number of observations (n), mean, standard deviation (SD), median, minimum, and maximum. The mean will be rounded to 1 decimal place greater than the precision of the original value, up to a maximum of 3 decimal places. The SD will be rounded to 2 decimal places greater than the precision of the original value, up to a maximum of 3 decimal places.

Supporting figures will be provided for some efficacy or safety analyses in addition to the summary tables. Efficacy listings will include analysis flags and/or imputed values to show which values were used in the summary tables. In addition, analysis visits will be displayed in the listing if it is important for the corresponding summary.

9.2 Hypothesis Testing and Multiplicity of Endpoints

The primary efficacy endpoint is the change from Baseline to Week 12 in the CMAI total score. For the primary efficacy analysis, the null hypothesis is that there is no treatment effect between AVP-786-42.63/4.9 and placebo, and between AVP-786-28/4.9 and placebo ($H_{01}\ and\ H_{02}$) during the study versus the alternative hypothesis that there is a treatment effect in at least one of the AVP-786 treatment groups ($H_{a1}\ or\ H_{a2}$).

AVP-786-28/4.9 (Low Dose) vs. placebo:

$$H_{01}: \mu_L - \mu_{pbo} = 0$$
$$H_{a1}: \mu_L - \mu_{pbo} \neq 0$$

AVP-786-42.63/4.9 (High Dose) vs. placebo:

$$H_{02}: \mu_H - \mu_{pbo} = 0$$
$$H_{a2}: \mu_H - \mu_{pbo} \neq 0$$

Average AVP-786-28/4.9 and AVP-786-42.63/4.9 vs. placebo:

$$H_0: (\mu_L + \mu_H)/2 - \mu_{pbo} = 0$$
$$H_a: (\mu_L + \mu_H)/2 - \mu_{pbo} \neq 0$$

The mITT efficacy family-wise error rate (FWE) for multiple comparisons:

1) Family 1 - the primary efficacy endpoint (CMAI total score) comparisons: AVP-786-28/4.9 vs. Placebo, and AVP-786-42.63/4.9 vs. Placebo

2) Family 2 - the key secondary efficacy endpoint (mADCS-CGIC-Agitation) comparisons: AVP-786-28/4.9 vs. Placebo, and AVP-786-42.63/4.9 vs. Placebo

is controlled by testing the difference between the average treatment effect of AVP-786-28/4.9 and AVP-786-42.63/4.9 vs. placebo ($global\ test\ H_0$) first, at significant level 2-sided α=0.05. If this global test is significant, and the estimated treatment difference is in the predicted direction favoring the active treatment, comparisons for each treatment group (AVP-786-28/4.9 and AVP-786-42.63/4.9) versus placebo (H_{01}, H_{02}). will be performed at 2-sided α=0.05 level. For the primary efficacy endpoint comparisons (Family 1), a treatment group comparison is significant at 2-sided α=0.05 FWE level if both the global test and the test of that group comparison are significant at 2-sided α=0.05 (Kong et al, 2005). If both the global test (H_0) and the comparison for each AVP-786 group vs placebo ($H_{01}\ and\ H_{02}$) are all significant at 2-sided α=0.05, then the same procedure will be repeated for the key secondary efficacy endpoint comparisons (Family 2), at 2-sided α=0.05.

For the remaining secondary efficacy endpoints and subgroup analyses which are supportive analyses, treatment comparisons will be performed at nominal 2-sided α=0.05 level without adjusting for multiplicity.

Sample SAS code to carry out the global test (H_0): average AVP-786-28/4.9 and AVP-786-42.63/4.9 vs. placebo:

[redacted]

10.0 Efficacy Statistical Analysis Methods

10.1 Consideration of Primary Estimands and Sensitivity Analysis

The purpose of this study is to assess the expected drug effect in a future population that results from patients initiating AVP-786-28/4.9, AVP-786-42.63/4.9 versus Placebo. The estimand of primary interest is defined as follows:

- Population: All patients who meet the inclusion/exclusion criteria, and who have at least 1 postbaseline efficacy assessment (mITT).
- Variable: Change from baseline to Week 12 in the CMAI total score.
- Intercurrent event(s): The "treatment policy" strategy will be followed, whereby the value for the variable of interest will be used regardless of adherence to randomized treatment and/or initiation of any concomitant medications which may include rescue medication (Lorazepam) and disallowed medications.
- Population-level summary: The between-treatment difference in the mean changes from baseline to Week 12 in the CMAI total score.

All data collected during the study treatment period will be used for statistical analysis. For the primary efficacy analysis, the treatment effect will be estimated using a likelihood-based linear MMRM on observed data. Under the MAR assumption, MMRM provides an unbiased estimate of treatment effect for the treatment period. Sensitivity analyses of the primary efficacy endpoint will be performed with missing values imputed by multiple imputation under MNAR (Section 10.4.2). In addition, missing value imputed

by LOCF and WOCF+LOCF will also be performed. These sensitivity results in totality will be used to substantiate the credibility of the results.

10.2 Derivation of Primary Efficacy Endpoints

CMAI was assessed at Baseline, and Weeks 1, 2, 3, 6, 9, and 12. Observed CMAI total scores including those from unscheduled visits and early discontinuation, are assigned analysis visit based on visit window Table 3-1 using assessment's study days. If 2 or more visits occur within the same analysis window, then the values closest to the target day will be used. If the assessment is the same distance from the target day, then the latest one will be used. After applying visit window, there is only one non-missing or missing value for each analysis visit. These data are called observed data and change from baseline in CMAI total score (observed data) for efficacy analyses are derived by subtracting corresponding baseline value at each visit. Other efficacy endpoints derivation follows the same algorithm.

The original protocol was started on May 2015. On Feb. 2017, a high dose treatment group AVP-786-42.63/4.9 was added by protocol amendment 4 when approximately 100 patients had already been enrolled. To account for potential period effect, a cohort factor is included in the efficacy analysis model with cohort 1 representing the period prior to the protocol amendment 4 being implemented and cohort 2 representing the period afterwards.

10.3 Primary Efficacy: Mixed Model Repeated Measures (MMRM)

The primary efficacy endpoint (change from Baseline to Week 12 in the CMAI total score) will be analyzed using a likelihood-based mixed model repeated measures (MMRM) analysis. This model will be run on the mITT population, using observed data.

The MMRM model will include terms for treatment, cohort, visit, treatment-by-visit interaction, baseline CMAI total score, baseline-by-visit interaction, baseline NPI AA (≤ 6 vs. > 6), risk assessment for falls (normal/mild vs. moderate/severe), concomitant use of antipsychotic medications (yes vs. no).

Unstructured covariance matrix will be used. If there are convergence problems, then the use of first-order autoregressive (AR1) and/or the compound symmetry (CS) covariance matrices will be considered and the first covariance structure converging to the best fit will be used as the primary analysis.

Model estimates (treatment difference and its 95% CI) will be reported in addition to the p-value. Multiple comparison family wise error rate (FWE) control is specified in Section 9.2.

[redacted]

This MNAR sensitivity analysis is to departure from MAR assumption by progressively increasing the

10.5 Secondary Efficacy Endpoint Analyses

Continuous secondary efficacy endpoints described in Section 4.2.2 will be analyzed using the statistical methodologies described in the following sections.

10.5.1 Secondary Efficacy Endpoint MMRM Analysis

The MMRM described in Section 10.3 will be used for analysis of the secondary efficacy endpoints (except RUD and GMHR) for the mITT population. Note that the raw score at post baseline visits for mADCS-CGIC-Agitation, ADCS-CGIC-Overall, and PGIC measure change from corresponding baseline. Analyses for these endpoints, baseline CMAI total score will be used as the covariate. For secondary efficacy endpoints, except key secondary mADCS-CGIC-Agitation, treatment comparisons will be performed at nominal 2-sided α=0.05 level without adjusting for multiplicity as specified in Section 9.2.

10.5.2 Response Analysis

10.5.3 Resource Utilization in Dementia (RUD) and General Medical Health Rating (GMHR) Analysis

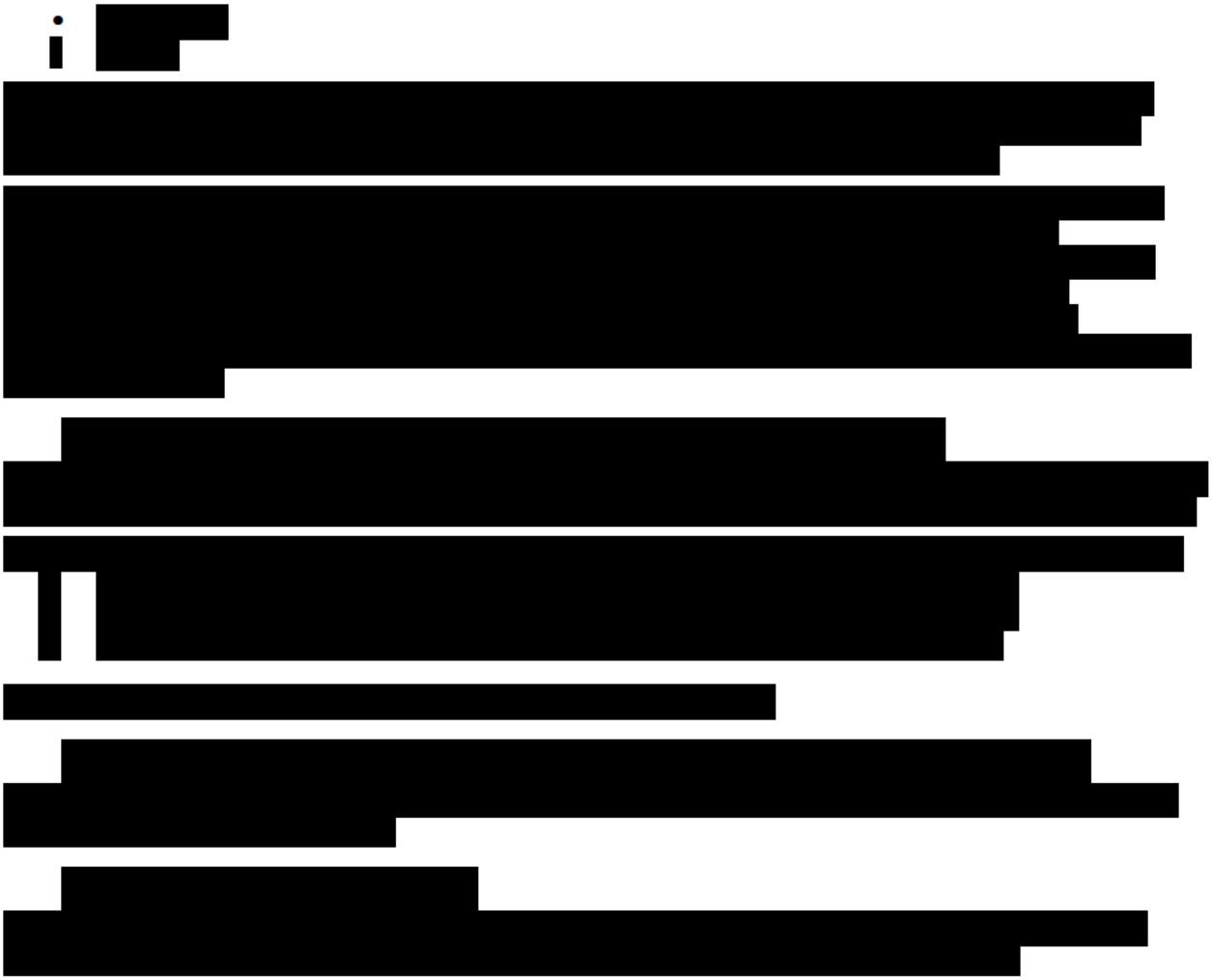
Descriptive analyses of the following RUD variables will be provided at baseline, week 6 and 12:

- Primary caregiver demographic and other characteristics: age, sex, relationship to subject, number of children, cohabitation (Yes/No), number of additional caregivers, level of contribution,
- Subject living accommodation characteristics: living accommodation; living arrangements, number of subjects with temporary arrangements and type, average nights in temporary accommodations
- Caregiving time
    - Average time spent assisting the subject each day for items 2a) 3a) 4a)
    - Average number of days per month for items 2b) 3b) 4b)
- Aspects of caregiver burden
    - Sleeping hours of the caregiver: average daily sleep time, and the change from baseline on this outcome
    - Caregiver responsibilities affect their work: work (Yes/No) and reasons for No; average paid hours, average hours paid to care for subject, average hours cut down due to caring for subject, average number of days missed work due to caregiving, average number of part of a day missed due to caregiving
    - Among the subset of subjects whose caregiver has reported they work for pay at baseline, the following will be analyzed:

        - Number of days of work missed completely
        - Number of days of work missed partially or completely
- Subject temporary institutionalization; institutionalisation (identified as the following living accommodations: "dementia-specific residential accommodation" and "long-term institutional care"):
    - Proportion of subjects reporting at least one institutionalisation
- Subject healthcare resource utilization
    - Number of hospitalizations and type of ward, number of emergency visits, number of health services and type of healthcare professional, and type and duration of services
    - Overall amount of utilization expressed in the appropriate unit (visits/contacts, days) including number of days receiving care
    - Percentage of subjects with at least one utilisation
    - Amount of utilization expressed in the appropriate unit (visits/contacts, days) within subjects with at least one utilisation

Descriptive analyses of the GMHR variable will be provided at baseline (Screening visit) and Week 12.

[redacted]



11.0 Safety Statistical Analysis Methods

Safety will be assessed through the analysis of AEs, clinical laboratory assessments, ECGs, vital signs, physical and neurological examinations, S-STS, MMSE, TUG test, and ESS. All safety analyses will be completed on the Safety population.

Unless otherwise specified, safety analyses will be displayed by treatment groups. As noted in Section 6, baseline for safety assessments is the last non-missing value prior to taking study drug at baseline.

11.1 Adverse Events

AE tables (except the AE Overview table) will only include summaries of TEAEs. Treatment-emergent adverse events are defined as AEs which first occur, or worsen, after the first dose of study medication and within 30 days after the permanent discontinuation of the study medication (i.e., first dose date ≤ AE start date ≤ last dose date + 30 days).

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

- Number of total AEs, TEAEs and deaths

- Incidence of patients with at least one TEAE, drug-related TEAE, non-serious, serious adverse events (SAEs), drug-related serious adverse event, and death
- Incidence of patients who discontinued due to TEAE, drug-related TEAE, SAE and drug-related SAE
- Incidence of deaths and deaths due to drug-related TEAE

TEAEs will be summarized by system organ class (SOC) and preferred term (PT), descending frequency of PT, and by maximum severity. Summaries of PTs will also be done for those occurring in at least 5% of patients in any treatment group.

AEs leading to 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 n (%) of patients with recurrences will also be given.

SAEs will be summarized by SOC and PT and will include a summary of drug-related events.

Below are the rules to follow for AE summaries:

- If a patient has multiple AEs within the same SOC or PT, the patient will only be counted once within a level of MedDRA.
- A drug-related AE is defined as an AE with an assigned relationship of “possibly related,” “related,” or missing.
- When assessing severity, if a patient has 2 or more TEAEs, the TEAE with the worst severity will be chosen. AEs with missing severity will be excluded from summaries of AE by severity.
- A common TEAE is defined as a TEAE with an incidence of ≥3% in either AVP-786 treatment group AND ≥2 times the incidence of the placebo treatment group.
- Time to onset will be calculated in days as (AE start date – first dose date).
- Duration of AE is generally defined as (AE end date – AE start date + 1). Below are some additional considerations for AE duration:
    - If the patient has an AE that has not ended when the patient ends the study, AE end date is defined as the last dose date of study drug.
    - If the same AE occurs more than once, total duration will be the sum of the individual AE durations.
- 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) x 100.
- Recurrence is defined as a new report of the same TEAE with a new AE start date.

AEs will be coded using MedDRA version 18.1.

11.2 Clinical Laboratory Assessments

Clinical labs will be reported for hematology, chemistry and urinalysis. Labs are collected at Screening, Week 3, Week 6, Week 9, and Week 12. The following parameters will be summarized descriptively through change from baseline and percent change from baseline by treatment and visit based on the Safety Population.

- 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])

- 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)

- Thyroid function tests: (thyroid stimulating hormone [TSH], and reflex T3 and T4 if TSH is abnormal). Thyroid function tests are collected at the Screening visit only and will only be provided in a data listing.

Out-of-range values will be assessed through shift tables of baseline to EOT. 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.

The denominator for the percentages will be the total number of patients in the treatment group with a baseline and post-baseline value. Lab tests without normal ranges will be excluded from the shift tables.

Potentially clinically significant (PCS) tables will also be used to summarize out-of-range values. PCS values are found in Table 4. The number and percentage 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.

TABLE 4. 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	≥3X ULN		Glucose	mmol/L	≤2.775	≥11.1
ALT (SGPT)	U/L	None	≥3X ULN		LDH	U/L	None	≥3X ULN
AST (SGOT)	U/L	None	≥3X ULN		Magnesium	mmol/L	<0.37	>1.23
Bilirubin	umol/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

Laboratory Parameter	Unit	Low PCS Criteria	High PCS Criteria		Laboratory Parameter	Unit	Low PCS Criteria	High PCS Criteria
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		Urate (Male)	umol/L	None	≥624.54
Creatine Kinase	U/L	None	≥3 ULN		Urate (Female)	umol/L	None	≥505.58
Creatinine	umol/L	None	>132.6					
Hematology								
Hemoglobin	g/L	<100	>180		Monocytes	x10^9/L	None	>1
Hematocrit	proportion of 1.0	<0.3	>0.5		Monocytes/ Leukocytes	%	None	≥15
Basophils	x10^9/L	None	>0.3		Neutrophils/ Leukocytes	%	≤15	None
Eosinophils/Leukocytes	%	None	≥10		Leukocytes	x10^9/L	≤2.8	≥16
Lymphocytes	x10^9/L	≤0.5	>4		Erythrocytes	x10^12/L	≤2.5	≥7.0
Lymphocytes/ Leukocytes	%	≤10	≥60		Platelet Count	x10^9/L	≤100	≥700

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

11.3 ECGs

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

- Screening
- Baseline (pre-dose)
- Baseline (2-3 hours post-dose)
- Week 2
- Week 3
- Week 6 (pre-dose)
- Week 9
- Week 12

The following quantitative parameters will be reported by the central reader: heart rate, PR interval, QRS duration, QT interval (uncorrected), and QT interval due to Fridericia's correction (QTcF). Change from baseline and percent change from baseline will be calculated for each parameter and summarized by the treatment groups.

In addition, since ECGs are recorded pre- and post-dose at Baseline, change from pre- to post-dose will be summarized at these visits.

PR interval and QTcF will be further investigated through PCS tables, for which the criteria are found in Table 5 below. The number and percentage 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 5. 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	>470 to ≤485, >485 to ≤500, >500
	Both	Change from baseline (increase)	>30, >60

ECG overall interpretations will be summarized by the number and percentage that were normal or abnormal. The interpretations by the cardiologist (i.e., central ECG) will be used for these summaries. The listings will provide all interpretations and corresponding details.

11.4 Vital Signs

Vital signs will be assessed at all visits (except the Follow-up Visit for ET patients). Starting with Protocol Amendment #3, orthostatic blood pressure and heart rate measurements will be performed at all clinic visits. Supine blood pressure and heart rate will be measured twice. The average of the measurements will be presented. Additionally, a single measurement of blood pressure and heart rate while standing will be collected. The following parameters will be summarized: systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate, respiratory rate, and temperature. These parameters will be summarized through change from baseline and percent change from baseline in similar fashion as the ECG parameters.

Vital signs will also be assessed through PCS criteria, which are given in Table 6. 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 6. 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
Heart rate (bpm)	≥120 AND ≥15 increase from baseline	≤50 AND ≥15 decrease from baseline
SBP and heart rate	SBP ≥10 increase from baseline AND heart rate ≥5 increase from baseline	
DBP and heart rate	DBP ≥5 increase from baseline AND heart rate ≥5 increase from baseline	

Orthostatic changes in blood pressure and heart rate from supine to standing will also be summarized. Additionally, patients meeting orthostatic hypotension or postural tachycardia PCS criteria at any time post-baseline will be summarized according to the criteria in Table 7.

TABLE 7. ORTHOSTATIC HYPOTENSION AND POSTURAL TACHYCARDIA PCS CRITERIA

Category	PCS Criterion
Orthostatic hypotension	≥ 20 mmHg decrease in SBP or ≥ 10 mmHg decrease in DBP from supine to standing
Postural tachycardia	≥ 30 bpm increase in heart rate from supine to standing or a standing heart rate ≥ 120 bpm

11.5 Physical and Neurological Examinations

Physical and neurological examinations will be performed at Screening, Week 6, and Week 12 and include assessments of the head, eyes, ears, nose, throat, lymph nodes, skin extremities, respiratory, gastrointestinal, musculoskeletal, cardiovascular, and nervous systems. Physical and neurological examination data will be provided in a data listing.

11.6 Sheehan Suicidality Tracking Scale (S-STS)

The S-STS is a prospective scale that assesses treatment-emergent suicidal thoughts and behaviors. Each item of the S-STS is scored on a 5-point Likert scale as shown below:

S-STS Item Scale

- 0 = Not at all
- 1 = A little
- 2 = Moderate
- 3 = Very
- 4 = Extremely

The S-STS can be analyzed as individual item scores, suicidal ideation subscale score, suicidal behavior subscale score, and a total score. For the screening visit, the timeframe for the items will be 'in the past 6 months' and for all other visits it will be 'since last visit'.

The S-STS will be assessed at all clinic visits. Scores, CFB in scores, and percent CFB in scores will be summarized descriptively by visit and treatment group for each subscale and total score. S-STS suicidal ideation subscale score, suicidal behavior subscale score, and a total score will be calculated as follow:

Parameter	Description of Derivation	Programming Algorithm	Missing Data Handling
Suicidal Ideation	Sum of: Questions 2 - 11; Questions are on a 0-4 scale	Sum of: SSTS102, SSTS103, SSTS104, SSTS105, SSTS106, SSTS107, SSTS108, SSTS109, SSTS110, SSTS111	if any missing for Questions 2 - 11, then score will be missing
Suicidal Behavior	If Question 1b=Yes, then use the following algorithm: Sum of: Questions 1a, highest score of (Q12 or any row of Q16), highest score of (Q14 or any row of Q15), 17, and 20 If Question 1a is not present,	If SSTS101B=Yes (looks like this is only present in data if =Yes, then do: Sum of: SSTS101A, max(SSTS112 or SSTSQ16C or SSTSQ16D), max(SSTS114 or SSTSQ15C or SSTSQ15D), SSTS17=Yes add 100, SSTS20=Yes add 4 If SSTS101B not present or No, then do:	if Questions 12 or 14 are missing, then score will be missing

	then use the following algorithm: Sum of: highest score of (Q12 or any row of Q16), highest score of (Q14 or any row of Q15), 17, and 20 Notes: Questions 15, 16, 17, and 20 are optional and only included in the QS datasets if populated with data. Scores should still be calculated if these Q15, Q16, Q17, Q20 are missing.	Sum of: max(SSTS112 or SSTSQ16C or SSTSQ16D), max(SSTS114 or SSTSQ15C or SSTSQ15D), SSTS17=Yes add 100, SSTS20=Yes add 4	
Total Scale Score	If Question 1b=Yes, then use the following algorithm: Sum of: Questions 2 - 11, Questions 1a, highest score of (Q12 or any row of Q16), highest score of (Q14 or any row of Q15), 17, and 20 If Question 1a is not present, then use the following algorithm: Sum of: Questions 2 - 11, highest score of (Q12 or any row of Q16), highest score of (Q14 or any row of Q15), 17, and 20 Notes: Questions 15, 16, 17, and 20 are optional and only included in the QS datasets if populated with data. Scores should still be calculated if these Q15, Q16, Q17, Q20 are missing.	If SSTS101B=Yes (looks like this is only present in data if =Yes, then do: Sum of: SSTS102, SSTS103, SSTS104, SSTS105, SSTS106, SSTS107, SSTS108, SSTS109, SSTS110, SSTS111, SSTS101A, max(SSTS112 or SSTSQ16C or SSTSQ16D), max(SSTS114 or SSTSQ15C or SSTSQ15D), SSTS17=Yes add 100, SSTS20=Yes add 4 If SSTS101B not present or No, then do: Sum of: SSTS102, SSTS103, SSTS104, SSTS105, SSTS106, SSTS107, SSTS108, SSTS109, SSTS110, SSTS111, max(SSTS112 or SSTSQ16C or SSTSQ16D), max(SSTS114 or SSTSQ15C or SSTSQ15D), SSTS17=Yes add 100, SSTS20=Yes add 4	if any missing for Questions 2 - 12, 14, then score will be missing

11.7 Mini Mental State Examination (MMSE)

The MMSE is a brief 30-point questionnaire test that is used to screen for cognitive impairment and severity of cognitive impairment. The MMSE scale comprises 11 questions or simple tasks concerning orientation, memory, attention, and language to evaluate a patient's cognitive state and are scored according to the following ranges:

Item	Score Range
• Orientation to Time	• 0 to 5
• Orientation to Place	• 0 to 5
• Registration	• 0 to 3
• Attention and Calculation	• 0 to 5
• Recall	• 0 to 3

• Naming	• 0 to 2
• Repetition	• 0 to 1
• Comprehension	• 0 to 3
• Reading	• 0 to 1
• Writing	• 0 to 1
• Drawing	• 0 to 1

The total score is calculated by summing all of the item scores and ranges from 0 to 30. Higher scores indicate milder cognitive impairment.

The MMSE will be assessed at Screening, Baseline, Week 6, and Week 12. The MMSE total score, CFB in total score, and percent CFB in total score will be summarized descriptively by visit and treatment group.

11.8 Timed Up and Go (TUG) Test

The TUG test measures the time (in seconds) taken for an individual to stand up from a standard armchair, walk 3 meters, turn, walk back to the chair and sit down. It is a commonly used scale for measuring functional mobility and risk of falls.

The TUG test will be performed at Screening, Week 6, and Week 12. The TUG time, CFB in TUG time, and percent CFB in TUG time will be summarized descriptively by visit and treatment group.

11.9 Epworth Sleepiness Scale (ESS)

The ESS is an 8-item questionnaire that is used to measure sleepiness by rating the probability of falling asleep on 8 different situations that most people engage in during the day. The questions are rated on a 4-point scale (0 to 3) where 0 = would never doze, 1 = slight chance of dozing, 2 = moderate chance of dozing, and 3 = high chance of dozing. A total score of 0 to 9 is considered to be normal.

The ESS will be assessed at Baseline, Week 6, and Week 12. The total score, CFB in total score, and percent CFB in total will be summarized descriptively by visit and treatment group.

12.0 Additional Summaries

12.1 Patient Disposition

Patient enrollment will provide the number of patients screened along with the reason for screen failures. A summary of randomized treatment group will be provided. Patient status will be summarized by the following:

- Randomized
- Took study medication
- Discontinued
- Completed study

Primary reasons for discontinuation will be provided based on the number of patients in the treatment group. The number of patients in the mITT and Safety populations will be provided.

12.2 Protocol Deviations

Protocol deviations will be reported by category for each treatment group and will be listed and summarized.

12.3 Demographic and Baseline Characteristics

Demographics will be summarized by randomized treatment group and overall for the mITT and Safety 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 (<65, ≥65)
- Weight

In addition, below information will be summarized by treatment group unless specified otherwise.

- Baseline disease characteristics: Safety Population, caregiver relationship to patient, patient living arrangement, number of patients enrolled prior to protocol amendment 4 and after.
- Randomization Stratification factors: mITT population, Baseline NPI AA domain score (≤ 6 and > 6), Risk for falls (normal/mild and moderate/severe), Use of concomitant antipsychotic medications (yes/no).
- Baseline efficacy assessment: mITT population, for CMAI, NPI, CGIS-Agitation

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

12.4 Medical History

Medical history will be summarized for the mITT and Safety populations by SOC and PT and will be provided in data listings.

12.5 Exposure

Duration of exposure will be summarized quantitatively for the Safety population 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).

An additional summary of exposure will be provided by dose taken for patients who took AVP-786 at any time during the study. The number of days at the specified dose level will be summarized. Doses summarized will be 18 mg QD, 18 mg BID, 28 mg QD, 28 mg BID, and 42.63 mg BID.

12.6 Compliance

Overall treatment compliance will be calculated as a percentage using the total number of capsules that were taken. 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 for the Safety population as described above and using descriptive statistics by treatment group within each dose regimen: Placebo, 18 mg QD, 18 mg BID, and 28 mg QD, 28 mg BID, and 42.63 mg BID. The number of doses taken and number of doses should have taken are necessary to calculate compliance. In the first 3 weeks the actual number of doses taken by QD or BID may not be distinguishable. In which case a combined compliance will be reported. The 3 steps for calculating compliance are shown here:

1. Doses taken = actual amount taken

2. Doses should have taken: 2 * (last dose date – first dose date) + 1

3. Compliance = (doses taken / doses should have taken) * 100

The formula is based on the assumption that patients take 2 capsules per day except for the last day, in which they only take 1.

12.7 Prior and Concomitant Medications, Non-drug Therapies and Interventions

The number and percentage of prior and concomitant medications will be provided by the treatment group used for the safety analyses. Summaries will be provided by anatomical therapeutic chemical (ATC) and preferred term for the Safety population. Non-drug therapies and nonpharmacological interventions will be only provided in data listings.

In addition, below information will be summarized;

- At baseline, number and % patients who used medications for AD (donepezil, donamem, rivastigmine, galantamine, memantine), medications/medication-classification for agitation secondary to AD (atypical antipsychotics, antidepressants, butyrophenones, buspirone).
- Cumulative number and % patients who used short term rescue medication lorazepam by visit.

12.8 Pharmacokinetics and Pharmacodynamics

Plasma concentrations for d6-DM, d6-DM metabolites, and Q will be collected at Week 6 and Week 12. Plasma concentration will be summarized by visit and treatment group for the Safety population overall [redacted]. Additionally, PK/PD correlations will be provided.

13.0 Validation

PRA's goal is to ensure that each TFL delivery is submitted to the highest level of quality. Our quality control procedures will be documented separately in the study specific quality control plan.

14.0 References

[redacted]

2. Lan Kong, et al, Performance of some multiple testing procedures to compare three doses of a test drug and placebo. Pharmaceutical Statistics. 2005;4:25-35.

Appendix 1 Glossary of Abbreviations

Glossary of Abbreviations:	
AA	Agitation/Aggression
AD	Alzheimer's disease
ADAS-cog	Alzheimer's Disease Assessment Scale – Cognitive Subscale
ADCS	Alzheimer's Disease Cooperative Study
ADCS-CGIC-Overall	Alzheimer's Disease Cooperative Study – Clinical Global Impression of Change Rating for Overall Clinical Status
AE	Adverse event
ALT/SGPT	Alanine aminotransferase/serum glutamic-pyruvic transaminase
ANCOVA	Analysis of covariance
AR1	First-order autoregressive
AST/SGOT	Aspartate aminotransferase/serum glutamic-oxaloacetic transaminase
ATC	Anatomic therapeutic classification
BID	Twice daily
BP	Blood pressure
BUN	Blood urea nitrogen
CDISC	Clinical Data Interchange Standards Consortium
CFB	Change from baseline
CGIC	Clinical Global Impression of Change
CGIS-Agitation	Clinical Global Impression of Severity of Illness-Agitation
CK	Creatinine kinase
CRF	Case report form
CS	Compound symmetry
CSDD	Cornell Scale for Depression in Dementia
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
[redacted]	[redacted]
D6-DM	Deuterated (d6)-dextromethorphan hydrobromide
DBP	Diastolic blood pressure
DEMQOL	Dementia Quality of Life
DM	Dextromethorphan hydrobromide
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram

EQ-5D-5L	EuroQol 5-Dimension 5-Level
EOT	End of treatment
ESS	Epworth Sleepiness Scale
GGT	Gamma-glutamyl transferase
GMHR	General Medical Health Rating
ITT	Intent-to-Treat
IWRS	Interactive Web Response System
LDH	Lactate dehydrogenase
LOCF	Last observation carried forward
LS	Least squares
mADCS-CGIC-Agitation	Modified Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change Rating for Agitation
MAR	Missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
mITT	Modified Intent-to-Treat
MMRM	Mixed effects model repeated measures
MMSE	Mini Mental State Examination
MNAR	Missing not at random
NPI	Neuropsychiatric Inventory
NPI-NH	Neuropsychiatric Inventory Nursing-Home version
OLS	Ordinary least squares
PCS	Potentially clinically significant
PGIC	Patient Global Impression of Change
pH	Potential hydrogen
PR	The P-R interval from an ECG tracing
PT	Preferred term
Q	Quinidine sulfate
QD	Once daily
QRS	The Q-R-S complex from an ECG tracing
QT	QT interval from an ECG tracing
QTc	QT interval corrected for heart rate
QTcB	QT interval corrected for heart rate using the Bazett's formula

QTcF	QT interval corrected for heart rate using the Fridericia's formula
RBC	Red blood cell
RUD	Resource Utilization in Dementia
SAE	Serious adverse event
SAP	Statistical analysis plan
SBP	Systolic blood pressure
SD	Standard deviation
SNRI	Serotonin-norepinephrine reuptake inhibitor
SSRI	Selective serotonin reuptake inhibitor
SOC	System organ class
SPCD	Sequential parallel comparison design
S-STS	Sheehan Suicidality Tracking Scale
T3	Triiodothyronine
T4	Thyroxine
TEAE	Treatment-emergent adverse event
TFLs	Tables, figures, and listings
TSH	Thyroid stimulating hormone
TUG	Timed Up and Go
VAS	Visual analog scale
WBC	White blood cells
WHODRUG DDE	World Health Organization Drug Dictionary Enhanced
WOCF	Worst observation carried forward
ZBI	Zarit Burden Interview

Appendix 2 Adverse Event Start/Stop Date Imputation

Imputation Rules for Partial Dates (D = day, M = month, Y = year)

Parameter	Missing	Additional Conditions	Imputation
Start date for AEs	D	M and Y same as M and Y of first dose of study drug	Date of first dose of study drug
		M and/or Y not same as date of first dose of study drug	First day of month
	D and M	Y same as Y of first dose of study drug	Date of first dose of study drug
		Y prior to Y of first dose of study drug but same as Y of screening date	Date of screening date
	D, M, Y	None - date completely missing	Date of first dose of study drug
Stop date for AEs	D	M and Y same as M and Y of last dose of study drug	Date of last dose of study drug
		M and/or Y not same as date of last dose of study drug	Use last day of month
	D and M	Y same as Y of last dose of study drug	Date of last dose of study drug
		Y not same as Y of last dose of study drug	Use Dec 31
	D, M, Y	None - date completely missing	No imputation, but assume ongoing

Note: In all cases, if an estimated start date is after a complete stop date, use the first day of the stop date month.
Similarly, if the estimated stop date is before a complete or imputed start date, use the last day of the start day month.

Appendix 3 Prior and Concomitant Medication Start/Stop Date Imputation

Imputation Rules for Partial Dates (D = day, M = month, Y = year)

Parameter	Missing	Additional Conditions	Imputation
Start date for con meds	D only	M and Y same as M and Y of first dose of study drug	Date of first dose of study drug
		M and/or Y not same as date of first dose of study drug	First day of month
	M and D	Y same as Y of first dose of study drug	Date of first dose of study drug
		Y not same as Y of first dose of study drug	Use Jan 01 of Y
	M, D, and Y	None - date completely missing	Day prior to date of first dose of study drug
Stop date for con meds	D only	M and Y same as M and Y of last dose of study drug	Date of last dose of study drug
		M and/or Y not same as date of last dose of study drug	Last day of month
	M and D	Y same as Y of last dose of study drug	Date of last dose of study drug
		Y not same as Y of last dose of study drug	Use Dec 31 of Y
	M, D, and Y	None - date completely missing and NOT ongoing	Date of last dose of study drug

Note: In all cases, if an estimated start date is after a complete stop date, use the first day of the stop date month.
Similarly, if the estimated stop date is before a complete or imputed start date, use the last day of the start day month.