Official Title: 52-week open-label extension study of pimavanserin for the treatment of agitation and aggression in subjects with Alzheimer's disease

NCT Number: NCT03118947

STATISTICAL ANALYSIS PLAN

Protocol No.:	ACP-103-033
Protocol Title:	52-Week Open-Label Extension Study of Pimavanserin for the Treatment of Agitation and Aggression in Subjects With Alzheimer's Disease
Protocol Version No. and Date	Amendment 3, 30NOV2017
Drug:	Pimavanserin
Sponsor:	ACADIA Pharmaceuticals Inc. [redacted]
SAP Version No. and Date	Version 1, 06 March 2019

SIGNATURE/APPROVAL PAGE

AUTHOR

[redacted]

14 MAR 2019

Date

Principal Biostatistician, Biometrics
ACADIA Pharmaceuticals Inc.

APPROVERS

[redacted]

14 MAR 2019

Date

Vice President, Biometrics
ACADIA Pharmaceuticals Inc.

TABLE OF CONTENTS

LIST OF TABLES

ABBREVIATIONS

AD	Alzheimer's disease
ADCS-ADL	Alzheimer's Disease Cooperative Study – Activities of Daily Living Inventory
AE	adverse event
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
bpm	beats per minute
CIs	confidence intervals
CMAI	Cohen-Mansfield Agitation Inventory
ECG	electrocardiogram
eCRF	electronic case report form
EOS	end-of-study
ET	early termination
GCAS	Global Clinician Assessment of Suicidality
IWRS	Interactive Web Randomization System
KSS	Karolinska Sleepiness Scale
MMSE	Mini-Mental State Examination
mADCS-CGIC	modified Alzheimer's Disease Cooperative Study - Clinical Global Impression of Change
MedDRA	Medical Dictionary for Regulatory Activities
msec	Milliseconds
NPI-C	Neuropsychiatric Inventory-Clinician Rating Scale
OL	open-label
PCI	potentially clinically important
QD	once daily
QTcB	QT interval corrected for heart rate using Bazett's formula
QTcF	QT interval corrected for heart rate using Fridericia's formula
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SOC	system organ class
TEAE	treatment-emergent adverse event
ZBI	Zarit Burden Interview

1 INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the summaries of efficacy and safety data. Specifications for tables, figures, and listings are contained in a separate document.

2 OBJECTIVES

The purpose of this study is to evaluate the safety and tolerability of pimavanserin treatment with up to 52 weeks of exposure (~64 weeks total for subjects who received pimavanserin in Study ACP-103-032), in subjects with probable Alzheimer's disease (AD).

2.1 Primary Objective and Endpoints

The primary objective of the study is to evaluate the safety and tolerability of pimavanserin after 52 weeks of treatment in subjects with probable AD who have symptoms of agitation and aggression.

2.1.1 Primary Endpoint

- Treatment-emergent adverse events (TEAEs)

2.1.2 Safety Endpoints

- Serious AEs (SAEs) and withdrawals due to AEs
- Global Clinician Assessment of Suicidality (GCAS)
- Clinically important changes from Baseline in vital sign measurements, weight, clinical laboratory assessments, physical examinations, and electrocardiograms (ECGs)

2.2 Exploratory Objectives and Endpoints

2.2.1 Exploratory Objectives

To evaluate the persistence of the effects of pimavanserin treatment on:

- Agitation and aggression
- Caregiver burden
- The clinician's global assessment of treatment benefit
- Other neuropsychiatric symptoms
- Cognition

- Functional status
- Sleep and daytime wakefulness

2.2.2 Exploratory Endpoints

- Change from Baseline in Cohen-Mansfield Agitation Inventory (CMAI) total score
- Change from Baseline in Zarit Burden Interview (ZBI) total score
- Modified Alzheimer's Disease Cooperative Study - Clinical Global Impression of Change (mADCS-CGIC) agitation score
- Change from Baseline in Neuropsychiatric Inventory-Clinician Rating Scale (NPI-C) combined agitation and aggression domain scores
- Change from Baseline in NPI-C total score
- Change from Baseline in NPI-C individual domain scores
- Change from Baseline in Mini-Mental State Examination (MMSE) score
- Change from Baseline in Karolinska Sleepiness Scale (KSS) score
- Change from Baseline in CMAI subscale scores
- Change from Baseline in Alzheimer's Disease Cooperative Study - Activities of Daily Living Inventory (ADCS-ADL) score

3 STUDY DESIGN

3.1 General Study Design

This study is an open-label extension study of the safety and tolerability of pimavanserin for the treatment of agitation and aggression in subjects with probable AD. This study will be conducted as a 52-week, open-label, flexible-dose extension of Study ACP-103-032. Subjects in Study ACP-103-033 must have completed ACP-103-032 and must meet eligibility criteria.

Study ACP-103-033 subjects must be consented prior to the final procedures being performed for Study ACP-103-032 at Week 12. Procedures from the Week 12 visit of Study ACP-103-032 will be carried over to the ACP-103-033 study to be included as Baseline information and this visit will be considered the Baseline Visit (Visit 1) of the ACP-103-033 study.

All subjects will receive once daily (QD) doses of pimavanserin (active study drug) over 52 weeks of treatment, starting with 20 mg a day. At the Week 2 visit, the pimavanserin 20 mg dose may be increased to 34 mg a day based on the Investigator's assessment of clinical response. The Investigator may then adjust the dose from 20 mg to 34 mg or from

34 mg to 20 mg at any visit based on clinical response. A dose change at a time other than at a scheduled clinic visit will require an unscheduled visit to dispense the appropriate dose of study drug. Unscheduled clinic visits may occur as needed.

The duration of participation for individual subjects will be up to approximately 56 weeks: each subject will participate in a 52-week open-label treatment period followed by an approximately 30-day safety follow-up period. The end of the clinical trial will be when the last subject completes the last scheduled assessment (i.e., 30-day follow-up).

3.2 Schedule of Assessments

The schedule of events and assessments for the study is presented in Table 1.

Table 1 Schedule of Assessments

	Treatment Period											Follow-Up
Visit Number	**Baseline[b] 1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**(EOS/ET) 10**	**Unscheduled**	**11**
Visit Week[a]	**0**	**Week 2**	**Week 4**	**Week 8**	**Week 12**	**Week 20**	**Week 28**	**Week 36**	**Week 44**	**Week 52**		**Week 56**
Allowable visit window (# days)		±3	±3	±3	±3	±7	±7	±7	±7	+7		+7
Informed consent[b]	X											
Inclusion/exclusion criteria	X											
Physical examination	X				X					X		
Vital signs	X	X	X	X	X	X	X	X	X	X	X	
Weight	X	X	X	X	X	X	X	X	X	X		
ECG[c]	X		X		X		X			X		
Clinical laboratory tests	X		X		X		X			X		
Pregnancy test[d]	X		X		X	X	X	X	X	X		
CMAI	X	X	X	X	X	X	X	X	X	X	X[h]	
ZBI	X	X	X	X	X	X	X	X	X	X		
NPI-C (all domains)	X				X		X			X		
NPI-C (agitation and aggression domains only)		X	X	X								
MMSE	X		X	X	X	X	X	X	X	X		
KSS	X				X		X			X		
mADCS-CGIC	X	X	X	X	X	X	X	X	X	X	X[h]	
ADCS-ADL	X				X		X			X		
Assessment of concomitant medications/procedures	X	X	X	X	X	X	X	X	X	X	X	X
Assessment of AEs[e]	X	X	X	X	X	X	X	X	X	X	X	X
GCAS	X	X	X	X	X	X	X	X	X	X	X	X
Dispense study drug[f]	X[g]	X	X	X	X	X	X	X	X		X[h]	
Study drug accountability		X	X	X	X	X	X	X	X	X	X[h]	

Abbreviations: ADCS-ADL=Alzheimer's Disease Cooperative Study–Activities of Daily Living; AE(s)=adverse event(s); CMAI=Cohen-Mansfield Agitation Inventory; ECG=electrocardiogram; EOS/ET=end-of-study/early termination; GCAS= Global Clinician Assessment of Suicidality; KSS=Karolinska Sleepiness Scale; mADCS-CGIC=modified Alzheimer's Disease Cooperative Study – Clinical Global Impression of Change; MMSE=Mini-Mental State Examination; NPI-C=Neuropsychiatric Inventory-Clinician rating scale; ZBI=Zarit Burden Interview

[a] Study visits are designated by weeks and have a ±3-day window (Visits 2 through 5), or a ±7-day window (Visit 6 through 9), or a +7-day window (Visits 10 and 11) calculated from the Baseline Visit.

[b] Study ACP-103-033 subjects **must be** consented prior to the final procedures being performed for Study APC-103-032 at Week 12. Procedures from the Week 12 visit of Study ACP-103-032 will be carried over as baseline information, if applicable.

[c] A single 12-lead ECG can be performed any time before blood sampling or at least 30 minutes after blood sampling during clinic visits.

[d] A urine pregnancy test should be performed for female study subjects of childbearing potential.

[e] Any untoward medical occurrence that occurs after the completion of procedures at the Week 12 visit in double-blind Study ACP 103-032 should be recorded as an AE, even if dosing for ACP-103-033 has not begun.

[f] Study drug will be dispensed to the subject at either scheduled or unscheduled visits.

[g] Study drug will be dispensed to the subject to take home at the Baseline visit. The subject and their study partner/caregiver will be provided instructions to take the first dose of study drug on the following day.

[h] To be completed at unscheduled visits where there is a dose change.

3.3 Randomization

Not applicable.

3.4 Blinding

Not applicable.

3.5 Determination of Sample Size

The planned sample size for this study is not based on statistical power but will depend on the number of subjects who complete Study ACP-103-032 and who then transition into this open-label extension study.

Up to approximately 111 male and female subjects with probable AD who have symptoms of agitation and aggression and who have completed Study ACP-103-032 and meet the eligibility criteria for ACP 103-033 will enroll in this study.

4 ANALYSIS SETS

The Safety Analysis Set will include all subjects who received at least 1 dose of open-label study drug. The Safety Analysis Set will be used for all analyses.

5 DATA HANDLING CONVENTIONS

All data collected in the study will be listed.

5.1 General Data Reporting Conventions

For continuous variables, the following summary statistics will be provided: number of subjects, mean, standard error of the mean (SE), standard deviation (SD), minimum, maximum and median. Unless specified otherwise, means, medians, and confidence intervals (CIs) will be presented to one more decimal place than the raw data, and the SDs deviations and SEs will be presented to 2 more decimal places than the raw data. Unless specified otherwise, all confidence intervals will be 2-sided 95% confidence intervals.

For categorical variables, summaries will include the number and percentage of subjects in each category, using the number of subjects with non-missing values as the denominator for the percentages (unless otherwise specified). Categories with zero counts will not have zero percentages displayed. Percentages will be presented with 1 decimal place.

When converting number of days to months, it will be calculated by dividing the number of days by 365.25 and then multiplying by 12. When converting number of days to years, the number of days will be divided by 365.25.

No hypothesis testing is planned. Descriptive summaries of all safety and efficacy endpoints will be provided. All safety and efficacy measures will be summarized for the Safety Analysis Set. Summaries by treatment group according to the original treatment in double-blind Study ACP-103-032 (placebo, pimavanserin 20 mg, and pimavanserin 34 mg) will be provided. All references to double-blind treatment group refer to the initial treatment group in Study ACP-103-032.

For each continuous measure in safety and efficacy analyses, change from Baseline results will be presented in two ways:

1. Using the Baseline from Study ACP-103-033 and reporting the changes across Study ACP-103-033 timepoints;
2. Using the Baseline from Study ACP-103-032 and reporting the changes across the timepoints of both the double-blind Study ACP-103-032 and the open-label Study ACP-103-033.

5.2 Derived Variables

In general, assessment total scores and subscale scores will be derived within the analysis datasets. In the event that total scores and/or subscale scores are also collected on the electronic case report form (eCRF), the derived values will be used for all summaries. Both the raw and derived scores will be presented in listings.

All Baseline assessments that are referred to in the following sections will be carried over from the Week 12 visit of double-blind Study ACP-103-032, where applicable.

5.2.1 Cohen-Mansfield Agitation Inventory Long-form (CMAI)

The CMAI is assessed at Baseline, and Weeks 2, 4, 8, 12, 20, 28, 36, 44, and 52/ET visits.

The CMAI is a 29-item scale designed to systematically assess agitation, rated on a 7-point (1-7) scale of frequency (1 = Never; 2 = Less than once a week; 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). Subjects are rated by their primary caregiver regarding the frequency with which they manifest physically aggressive, physically non-aggressive, and verbally agitated behaviors. The CMAI in this study is to be completed by interview of the caregiver. Ratings are inclusive of the 2 or 4 weeks prior to the administration of the scale.

In addition to the 7-point frequency scale, there are two other options for rating the behavior: 1) “8 - would occur if not prevented” (e.g., a person is physically restrained so he/she cannot pace), and 2) “9 - not applicable” (e.g., a non-verbal resident not being able to repeat sentences or questions, or a person who cannot walk or move a wheelchair not being able to pace, an amputated person not being able to kick). These two ratings are used only if the

behavior really has never occurred in the past two weeks. If it has occurred, then the 1-7 point frequency scale should be used. These ratings of 8 and 9 are not included in the CMAI total score calculation. The CMAI has a range of 29-203 points, with higher scores indicating more severe agitation symptoms. If there are less than 6 missing items then the total score will be imputed by replacing each missing item with the mean (rounded to the nearest integer) of the non-missing values for that subject and timepoint. If there are more than 5 missing items then the total score will be missing.

In addition to the total score, behavioral subscales will be defined as follows (Rabinowitz et al, 2005):

- Aggressive Behavior: 12 Items (minimum = 12, maximum = 84)
- Sum of scores from items 3, 4, 7, 8, 9, 10, 11, 13, 14, 15, 21 and 25
- Physically Non-aggressive Behavior: 6 Items (minimum = 6, maximum = 42)
- Sum of scores from items 1, 2, 16, 22, 26, and 29
- Verbally Agitated Behavior: 4 Items (minimum = 4, maximum = 28)
  Sum of scores from items 5, 6, 18, and 19
- Hiding and Hoarding : 2 Items (minimum = 2, maximum =14)
  Sum of score from items 23 and 24

For the aggressive behavior subscale, if less than 3 items are missing then the aggressive behavior subscale score will be imputed by replacing the missing item with the mean of the non-missing values (rounded to the nearest integer) for that subject and timepoint within aggressive behavior subscale. If more than 2 items are missing then the aggressive behavior subscale score will be missing.

For the physically non-aggressive subscale, if only 1 item is missing then the physically non-aggressive subscale score will be imputed by replacing the missing item with the mean of the non-missing values (rounded to the nearest integer) for that subject and timepoint within physically non-aggressive subscale. If more than 1 item is missing then the physically non-aggressive subscale score will be missing.

For the verbally agitated behavior and hiding and hoarding subscales, if any of the items are missing then the subscale score will be missing.

5.2.2 Zarit Burden Interview (ZBI)

The ZBI is assessed at Baseline, Weeks 2, 4, 8, 12, 20, 28, 36, 44, and 52/ET visits.

The ZBI was designed to assess the stresses experienced by caregivers of patients with dementia. Caregivers are asked to respond to a series of 22 questions about the impact of the patient's disabilities on their life. For each item, caregivers are to indicate how often they felt that way (never, rarely, sometimes, quite frequently, or nearly always) (0-4). The total score of ZBI will range from 0 to 88, with higher scores denoting more stresses experienced by caregivers. If there are less than 5 missing items then the total score will be imputed by replacing each missing item with the mean (rounded to the nearest integer) of the non-missing values for that subject and timepoint. If there are more than 4 missing items then the total score will be missing.

If the informant of ZBI is a professional caregiver, their ZBI total scores will not be included in the summaries.

5.2.3 Alzheimer's Disease Cooperative Study – Activities of Daily Living Inventory (ADCS-ADL)

The ADCS-ADL is assessed at Baseline and Weeks 12, 28 and 52/ET visits.

The ADCS-ADL is an inventory to assess activities of daily living (ADL) in subjects with AD. This is a caregiver-rated 23-item questionnaire, with sub-questions and sub-items for some items, including both basic (e.g. personal hygiene, eating, bathing) and instrumental (e.g. preparing a meal, using a telephone) ADLs. The total score can range from 0 to 78, with lower scores indicating worse functioning.

A response of "DON'T KNOW" is not considered a missing value, rather it will be scored similarly to a "NO" response (i.e., that item score will be set to 0). The following table includes the item score range and the scoring rule. These rules will only apply in the analysis dataset; the raw data values will remain unchanged.

Table 2 ADCS-ADL Item Score Range

Item Number	Score Range	Note*
1-5, 8-12, 14, 17-19, 21, 22	0-3	For item 8, the score is sum of 8A, 8B, and 8C. The sub-question will be equal to 0 if the main or sub-question checks NO or DON'T KNOW. For item 18, the score is sum of 18A, 18B, and 18C. The sub-question will be equal to 0 if 1) the main or sub-question checks NO or DON'T KNOW; or 2) "patient is institutionalized" box is checked. For item 22, if NO or DON'T KNOW is checked then assign item score = 0. If YES is checked and one or more sub-item is

Item Number	Score Range	Note*
		checked, but the sub-question is not answered, then assign item score = 1. Otherwise, item score is equal to the non-missing sub-question score.
6	0-7	Sum of 6A (0-3; 0 if the main question checks NO or DON'T KNOW) and 6B (0-4).
7	0-5	
13, 15, 16, 23	0-4	For item 16, the score is sum of 16A (0-3; 0 if the main question checks NO or DON'T KNOW) and 16B (0-1; 0 if the main or sub-question checks NO or DON'T KNOW). For item 23, if NO or DON'T KNOW is checked then assign item score = 0. If YES is checked and one or more sub-item is checked, but the sub-question is not answered, then assign item score = 1. Otherwise, item score is equal to the non-missing sub-question score.
20	0-2	

* Item 7-21. If NO or DON'T KNOW is checked, then assign item score = 0. If YES is checked, the score of the item is sum of non-missing score(s) of the sub-question(s).

In addition to the total score, the following subscale scores will be defined:

- Basic subscale score: 6 Items (minimum = 0, maximum = 22)

  Sum of scores from items 1, 2, 3, 4, 5, 6

- Instrumental subscale score: 17 Items (minimum = 0, maximum = 56)
  Sum of scores from items 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23

After applying the preceding data handling rules, if missing values remain then the 23-item total, basic, and instrumental scores will be derived as follows:

For the 23-item total score, if the sum of the maximum possible scores among all of the missing items is less than or equal to 16 then the total score will be derived as 78 multiplied by (sum of observed item scores divided by the sum of maximum possible scores among observed items), rounded to the nearest integer. Otherwise, the total score will be missing.

For the basic subscale score, if the sum of the maximum possible scores among the missing items is less than or equal to 4 then the basic subscale score will be derived as 22 multiplied by (sum of observed basic item scores divided by the sum of the maximum possible scores among observed basic item scores), rounded to the nearest integer. If the sum of the maximum possible scores among the missing items is greater than 4 then the basic subscale score will be missing.

For the instrumental subscale score, if the sum of the maximum possible scores among the missing items is less than or equal to 11 then the instrumental subscale score will be derived as 56 multiplied by (sum of observed instrumental item scores divided by sum of the maximum possible scores among observed instrumental item scores), rounded to the nearest integer. If the sum of the maximum possible scores among the missing items is greater than 11 then the instrumental subscale score will be missing.

5.2.4 Neuropsychiatric Inventory – Clinician Rating Scale (NPI-C)

The complete NPI-C, which includes agitation and aggression domains, will be administered at Baseline and Weeks 12, 28 and 52/ET visits while only the NPI-C agitation and aggression domains will be administered at Weeks 2, 4, and 8 visits.

The NPI-C is a scale to assess the 14 neuropsychiatric symptoms (domains) in patients with dementia. For each domain, there are three NPI-C sections: caregiver interview, patient interview and clinician rating. The caregiver interview includes the frequency (0-4), severity (0-3) and distress (0-5) scores. The subject interview consists of the subject-reported frequency score (0-4). The clinician rating is a clinical impression severity score (0-3) based on all available interview information. For the caregiver interview, if a response of 0 is selected for the frequency score, then the severity and distress scores will be set to 0. For each domain, the domain score is the sum of the clinical impression severity scores within each domain, with higher scores denoting more serious symptoms. Before deriving any combined domain scores, if there are any missing items, the individual domain score will be imputed as follows:

For the delusions, hallucinations, aggression, elation/euphoria, aberrant motor disturbance, sleep disorders, appetite and eating disorders, and aberrant vocalizations domains, if only 1 item is missing then the respective domain score will be imputed by replacing the missing item with the mean of the non-missing items (rounded to the nearest integer) for that subject and timepoint within the domain. If more than 1 item is missing then the respective domain score will be missing.

For the agitation, dysphoria, anxiety, apathy/indifference, and irritability/lability domains, if up to 2 items are missing then the respective domain score will be imputed by replacing each missing item with the mean of the non-missing items (rounded to the nearest integer) for that subject and timepoint within the domain. If more than 2 items are missing then the respective domain score will be missing.

For the disinhibition domain, if up to 3 items are missing then the respective domain score will be imputed by replacing each missing item with the mean of the non-missing values

(rounded to the nearest integer) for that subject and timepoint within the domain. If more than 3 items are missing then the respective domain score will be missing.

Table 3 NPI-C Domain Score

NPI-C Domain	# items	Domain Score Range	Maximum Missing Items Allowed for Deriving the Domain Score*
Delusions	8	0 – 24	1
Hallucinations	7	0 – 21	1
Agitation	13	0 – 39	2
Aggression	8	0 – 24	1
Dysphoria	13	0 – 39	2
Anxiety	14	0 – 42	2
Elation/Euphoria	6	0 – 18	1
Apathy/Indifference	11	0 – 33	2
Disinhibition	16	0 – 48	3
Irritability/Lability	12	0 – 36	2
Aberrant Motor Disturbance	9	0 – 27	1
Sleep Disorders	8	0 – 24	1
Appetite and Eating Disorders	9	0 – 27	1
Aberrant Vocalizations	8	0 – 24	1

* For each domain, the maximum number of missing items allowed for deriving the domain score is the nearest integer to but ≤ 20% of the total items in that domain. If there are any missing items, the individual domain score will be imputed first before calculating the combined domain scores.

The NPI-C combined agitation and aggression domain score is calculated as the sum of the agitation domain score and the aggression domain score. If either the agitation or aggression domain score is missing, the NPI-C combined agitation and aggression domain score will be missing.

The NPI-C combined delusions and hallucinations domain score is calculated as the sum of the delusions domain score and the hallucinations domain score. If either the delusions or hallucinations domain score is missing, the NPI-C combined delusions and hallucinations domain score will be missing.

The NPI-C combined dysphoria and apathy/indifference domain score is calculated as the sum of the dysphoria domain score and the apathy/indifference domain score. If either the

dysphoria or apathy/indifference domain score is missing, the NPI-C combined dysphoria and apathy/indifference domain score will be missing.

The NPI-C total score is calculated as the sum of all 14 individual domain scores. The NPI-C total score ranges from 0 to 426, with higher scores denoting more serious symptoms. After applying the preceding data handling rules, if any of the domain scores remain missing, the total score will be missing.

5.2.5 Modified Alzheimer's Disease Cooperative Study – Clinical Global Impression-Change (mADCS-CGIC)

The mADCS-CGIC is assessed at Baseline, Weeks 2, 4, 8, 12, 20, 28, 36, 44, and 52/ET visits.

The mADCS-CGIC scale will be used to allow the Investigator to determine the subject's overall clinical condition as it relates to their symptoms of agitation and aggression, and to address the clinical significance of changes from Baseline in other psychometric measures. The mADCS-CGIC interview will be performed by the Investigator or a medically qualified rater. After completion of the interview, the rater will be asked to rate the subject's symptoms of agitation and aggression relative to the Baseline interview from the parent study (ACP-103-032), using a standardized 7-point scale (1 = marked improvement to 7 = marked worsening). Higher scores denote worsening or less improvement in agitation and aggression symptoms.

Missing mADCS-CGIC scores will not be imputed.

5.2.6 Karolinska Sleepiness Scale (KSS)

KSS is assessed at Baseline, Weeks 12, 28 and 52/ET visits.

The KSS is a self-reported subjective measure of a subject's level of drowsiness. With the modified version, respondents must choose which of nine statements (1 = extremely alert to 9 = very sleepy, great effort to keep awake, fighting sleep) most accurately describes their level of sleepiness over a period of time, which for this study will be "on average over the previous week". Higher scores denote more drowsiness.

Missing KSS scores will not be imputed.

5.2.7 Mini-Mental State Examination (MMSE)

MMSE is assessed at Baseline, Weeks 4, 8, 12, 20, 28, 36, 44, and 52/ET visits.

The MMSE is an 11-area, 30 items questionnaire that is used to measure cognitive impairment, with lower scores indicating more severe cognitive impairment.

The total score (0-30) is calculated as the sum of the 30 item scores. If there are less than or equal to 6 missing items then the total score will be imputed as the mean of the non-missing values multiplied by 30 and rounded to the nearest integer for that subject and timepoint. If there are more than 6 missing items then the total score will be missing.

5.2.8 Global Clinician Assessment of Suicidality (GCAS)

Global Clinician Assessment of Suicidality (GCAS) is assessed at Baseline, Weeks 2, 4, 8, 12, 20, 28, 36, 44, 52/ET and Follow-up (if applicable) visits.

The GCAS will be used to assess the occurrence of treatment-emergent suicidal ideation and behavior. The GCAS is a clinician-rated, 5-point scale (0–4) that is designed to rate the subject's suicidality based on the report of the subject, the report of the study partner/caregiver, and the clinician's global assessment. Ratings can be 0 ("Absent"), 1 ("Feels life is not worth living"), 2 ("Wishes he/she were dead or any thoughts of possible death to self"), 3 ("Suicidal ideas or gesture"), or 4 ("Attempt at suicide"). The Investigator will record a subject rating, a partner/caregiver rating, and a clinician rating. For a rating of 3 or 4 based on the clinician's assessment, the date of event will be recorded. At each visit, suicidality since the previous visit will be assessed.

Missing GCAS scores will not be imputed.

5.3 Analysis Visit Windows

In general, the open-label Baseline assessment will be defined as the last non-missing assessment, including those from repeated and unscheduled measurements, before or on the first open-label dose date. Exceptions to this definition will be handled on a case by case basis.

Efficacy and safety assessments will be summarized by analysis visit as presented in Table 4 below.

Table 4 Analysis Visit Windows

Open-label (OL) Analysis Visit Name	Target Study Day[1]	Study Day Interval
OL Baseline	1	≤1(based on the first open-label dose date)
OL Week 2	15	2 to 21
OL Week 4	29	22 to 42
OL Week 8	57	43 to 70
OL Week 12	85	71 to 112

OL Week 20	141	113 to 168
OL Week 28	197	169 to 224
OL Week 36	253	225 to 280
OL Week 44	309	281 to 336
OL Week 52	365	337 to 379
OL Follow-up	395	380 – maximum

[1] If the assessment date ≥ first open-label dose date, study day = assessment date – first open-label dose date + 1; otherwise study day = assessment date – first open-label dose date. Study day 1 is the first open-label dose date.

5.3.1 Unscheduled Assessments

Both scheduled and unscheduled assessments, including early termination visits, will be considered for planned timepoint summaries. All assessments will be presented in data listings.

5.3.2 Multiple Measurements within Visit Windows

If more than one assessment falls within a given window then the assessment closest to the target study day will be selected for the by-visit summaries. If two assessments are equidistant from the target day then the chronologically last assessment will be used for summary. Exceptions may be made for incomplete assessments, in which case, more complete assessments may be given priority. Details are provided in a separate programming conventions document.

For safety summaries where the most extreme values should be selected, e.g. overall post-Baseline minimum, overall post-Baseline maximum and potentially clinically important values for overall post-Baseline summaries, all non-missing post-Baseline values should be considered, regardless of whether the value is selected for the by-visit summaries. All results will be presented in data listings.

5.4 Handling of Missing Date

5.4.1 Missing or Incomplete Date for Last Dose of Study Drug

In the Safety Analysis Set, if the last dose date of study drug is missing for a subject who completed or early terminated from the study, then the date of the end-of-study/early termination visit will be used in the calculation of treatment duration. For the incomplete last dose date of the study drug, the imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

For the data summarization before final database lock, if a subject is still ongoing, then this subject's last dose date will be imputed using the database extract date.

5.4.2 Missing or Incomplete Dates for Prior or Concomitant Medications

Missing or incomplete medication start or end dates will be imputed for the purpose of determining whether the medications are taken concomitantly (see Section 11 for definition). When the chronological order of a medication's utilization relative to the study drug treatment period is unclear due to missing or incomplete date(s), the medication will be considered as concomitant. The imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

5.4.3 Missing or Incomplete Dates for Adverse Events

Missing or incomplete adverse event (AE) start dates will be imputed for the purpose of determining whether the AEs are treatment emergent (see Section 14.1 for definition). When the chronological order of an AE onset relative to the study drug treatment period is unclear due to missing or incomplete date(s), the AE will be considered as treatment-emergent. The imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

5.4.4 Missing Severity Assessment for Adverse Events

If the severity is missing for a TEAE, then a severity of "Severe" will be assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be presented in data listings.

5.4.5 Missing Relationship to Study Drug for Adverse Events

If the relationship to study drug is missing for a TEAE, then a causality of "Related" will be assigned. The imputed values for relationship to study drug will be used for incidence summaries, while the actual values will be presented in data listings.

5.4.6 Character Values of Clinical Laboratory Variables

If the reported value of a clinical laboratory variable cannot be used in a summary due to, for example, a character string reported for a numerical variable, an appropriately determined coded value will be used in the summary. The coding algorithms will be detailed in the analysis dataset specification document. The actual values as reported in the database will be presented in data listings.

6 SUBJECT DISPOSITION

The number of sites that enrolled at least 1 subject and number of subjects enrolled will be summarized by region and overall. Enrolled subjects are subjects who signed informed consent for Study ACP-103-033 and whose status in the Interactive Web Randomization System (IWRS) is “Entered Extension”, “Treatment Completed EXT” or “Discontinued Treatment EXT”. Subjects who signed informed consent for Study ACP-103-033 but failed to roll over (recorded as “Rollover Failed” in IWRS) will not be counted as enrolled subjects.

For enrolled subjects, number and percentage of subjects in the Safety Analysis Set will be summarized. A listing will be provided displaying all subjects excluded from the Safety Analysis Set (if any), and will include reason(s) for exclusion.

The number and percentage of subjects in the Safety Analysis Set who completed the study or discontinued (all discontinued and by discontinuation reason) will also be summarized. Summaries by region for the Safety Analysis Set will also be presented.

7 PROTOCOL DEVIATIONS

Protocol deviations will be reviewed periodically over the course of the study. The review process, definition of the deviation categories, and the classification of a deviation as major (significant) or minor (not significant) are detailed in the Study Deviation Rules Document.

For enrolled subjects, a summary of the number and percentage of subjects with major protocol deviations for each deviation category will be presented. A listing of protocol deviations by site and subject will be provided.

8 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Demographics and open-label Baseline characteristics will be summarized for Safety Analysis Set using descriptive statistics. Summaries by region will also be presented. Variables include age, sex, race, ethnicity, height, weight, body mass index (BMI), region, Baseline living situation category, Baseline caregiver category, Baseline NPI-C combined agitation and aggression domain score, Baseline MMSE total score, Baseline CMAI total score, Baseline ZBI total score, and concomitant selective serotonin reuptake inhibitor (SSRI) usage (yes/no).

Age (the reported age reflects a subject’s age at the open-label Baseline visit date), Baseline CMAI total score, Baseline MMSE total score, and Baseline NPI-C combined agitation and aggression domain score will be presented as both continuous and categorical variables. Age will be presented as ≤85 and >85 years old. Category of Baseline CMAI will be presented as <65 and ≥65. Category of Baseline NPI-C combined agitation and aggression will be

presented as <22 and ≥22. Category of Baseline MMSE will be presented as ≤10, 11-20 and >20.

Alzheimer's disease history will be summarized for Safety Analysis Set using descriptive statistics. Variables include:

- Time (year) since diagnosis of probable Alzheimer's disease
- Duration (year) of symptoms of Alzheimer's disease
- Duration (year) of symptoms of agitation/aggression

Baseline visit date will be used as the reference date for calculating the durations listed above.

A listing of living situation and caregiver information by subject and visit will also be provided.

9 MEDICAL HISTORY

Medical and surgical history data will be coded using Medical Dictionary for Regulatory activities (MedDRA) version 19.0 or newer. The subject incidence of relevant medical and surgical history data will be summarized by system organ class (SOC) and preferred term for the Safety Analysis Set. A subject will be counted only once per SOC or preferred term for the summary.

A listing of the SOC, preferred term, body system, verbatim term for the medical history condition/event, start and stop dates (when available), and an indicator for whether or not the condition is ongoing will be provided.

10 EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE

Summaries of exposure and compliance to study drug will be provided for the Safety Analysis Set.

10.1 Exposure to Study Drug

Pimavanserin dose levels are expressed as free base.

10.1.1 Exposure to Open-label Pimavanserin

For each subject, the duration of exposure (in days) to open-label pimavanserin will be calculated as:

(date of last open-label dose) – (date of first open-label dose) + 1

The cumulative dose of pimavanserin will be calculated by first multiplying the number of doses taken by the dose level (in mg) for each kit returned to the site during the open-label study period and then summing the results for all kits. The average daily dose of pimavanserin (in mg) will be calculated as the cumulative dose (in mg) divided by duration of open-label exposure (in days). Duration of exposure, cumulative dose, and average daily dose for pimavanserin during open-label study period will be summarized descriptively.

Duration of open-label exposure will be summarized as both continuous and categorical variables. For categorical presentation, the number and percentage of subjects in each of the following categories will be presented: <4 weeks (1 to 27 days), 4 to <8 weeks (28 to 55 days), 8 to <12 weeks (56 to 83 days), 12 to <20 weeks (84 to 139 days), 20 to <28 weeks (140 to 195 days), 28 to <36 weeks (196 to 251 days), 36 to <44 weeks (252 to 307 days), 44 to <52 weeks (308 to 363 days) and ≥52 weeks (364 days or longer). Kaplan-Meier curves of duration on pimavanserin will also be presented.

In addition, summaries of whether subjects had any dose change (yes vs. no), their highest dose level (20 mg or 34 mg) and their last dose level (20 mg or 34 mg) will also be provided.

10.1.2 Total Exposure to Pimavanserin across Double-blind and Open-label Periods

For subjects who received pimavanserin in double-blind Study ACP-103-032, the total duration of exposure (in days) to pimavanserin will be calculated as the sum of the exposure durations in the double-blind and open-label periods. The cumulative dose of pimavanserin will be calculated by first multiplying the number of doses taken by the dose level (in mg) for each kit returned to the site during both the double-blind and the open-label study periods and then summing the results for all kits. The average daily dose of pimavanserin (in mg) will be calculated as the cumulative dose (in mg) divided by total duration of exposure (in days).

For subjects who received placebo in double-blind Study ACP-103-032, the total duration of exposure (in days) to pimavanserin, the cumulative dose of pimavanserin, and the average daily dose of pimavanserin across the double-blind and open-label periods will be the same as the values calculated for the open-label study period alone. Duration of exposure, cumulative dose, and average daily dose for pimavanserin across the double-blind and open-label periods will be summarized descriptively.

Duration of total exposure will be summarized as both continuous and categorical variables. For categorical presentation, the number and percentage of subjects in each of the following categories will be presented: <4 weeks (1 to 27 days), 4 to <8 weeks (28 days to 55 days), 8 to <12 weeks (56 days to 83 days), 12 to <16 weeks (84 days to 111 days), 16 to <28 weeks

(112 days to 195 days), 28 to <40 weeks (196 days to 279 days), 40 to <52 weeks (280 days to 363 days) , 52 to <64 weeks (364 days to 447 days), and ≥64 weeks (448 days or longer).

In addition, summaries of whether subjects had any dose change (yes vs. no), their highest dose level (20 mg or 34 mg) and their last dose level (20 mg or 34 mg) will also be provided.

10.2 Measurement of Treatment Compliance

Study drug dosing compliance (in percentage) during the open-label study period for a subject is defined as:

$$\frac{(total\ number\ of\ tablets\ actually\ taken)}{(total\ number\ of\ tablets\ expected\ to\ be\ taken)} \times 100$$

With the total number of tablets actually taken calculated as:

(the total number of tablets dispensed) – (the total number of tablets returned)

The number of tablets expected to be taken is calculated as:

(the duration of exposure) x 2.

The duration is multiplied by 2 because the subject is instructed to take 2 tablets per day.

Compliance will be summarized as both continuous and categorical variables. For categorical presentation, the number and percentage of subjects in each of the following categories will be presented: <80%, 80 to 120%, and >120%.

11 CONCOMITANT AND POST-TREATMENT MEDICATION

Concomitant medication is defined as any medication with a start date prior to the first dose of the open-label study drug and continuing after the first dose of the open-label study drug or with a start date between the first dose of the open-label study drug and last dose of open-label study drug, inclusive. Any medication with a start date after the date of the last dose of open-label study drug will be considered as post-treatment medication. Concomitant and post-treatment medications will be summarized separately.

Medications will be coded using WHO Drug Dictionary (WHODD) 2016 March or newer version. The number and percentage of subjects taking each drug class (ATC Level 3) and medication preferred term will be tabulated for the Safety Analysis Set. Multiple medication usage by a subject in the same ATC category will be counted only once.

12 EFFICACY ANALYSES

All efficacy summaries will be performed for the Safety Analysis Set. No hypothesis testing is planned. Descriptive summaries of all efficacy endpoints will be provided.

12.1 Exploratory Efficacy Variables

The exploratory efficacy endpoints include the following:

- Change from Baseline in Cohen-Mansfield Agitation Inventory (CMAI) total score
- Change from Baseline in Zarit Burden Interview (ZBI) total score
- Modified Alzheimer's Disease Cooperative Study - Clinical Global Impression of Change (mADCS-CGIC) agitation score
- Change from Baseline in Neuropsychiatric Inventory-Clinician Rating Scale (NPI-C) combined agitation and aggression domain scores
- Change from Baseline in NPI-C total score
- Change from Baseline in NPI-C individual domain scores
- Change from Baseline in Mini-Mental State Examination (MMSE) score
- Change from Baseline in Karolinska Sleepiness Scale (KSS) score
- Change from Baseline in CMAI subscale scores
- Change from Baseline in Alzheimer's Disease Cooperative Study - Activities of Daily Living Inventory (ADCS-ADL) score

12.2 Adjustment for Covariates

Not applicable.

12.3 Handling of Missing Data

Any derived scores (i.e. total, domain or subscale scores) that are missing after any imputation of individual missing items as described in Section 5.2 will not be imputed.

12.4 Multiple Comparisons / Multiplicity

No hypothesis testing is planned.

12.5 Examination of Subgroups

Subgroup analyses by geographic region (North America, Europe, or Rest of World) will be performed for the change from Baseline in the CMAI total score and the ZBI total score.

13 METHODS OF EFFICACY ANALYSES

13.1 Exploratory Efficacy Analyses

Descriptive statistics for all exploratory efficacy endpoints listed in Section 12.1 will be tabulated by double-blind treatment group at each scheduled timepoint. The summaries of the change from Baseline results will be presented in two ways as specified in Section 5.1, by double-blind treatment group at each scheduled timepoint.

13.1.1 Responder Analyses

In the responder analyses, since CMAI is an interval scale that lacks a natural zero point (1 = Never), the percent change in CMAI total score will be calculated based on corrected scores after subtracting 29 points from the raw scores. For example, if a subject's Baseline CMAI total score is 79 and Week 8 CMAI total score is 29 (absent of all agitation behaviors), the percent change from Baseline to Week 8 in CMAI total score will be calculated as [(29–29) – (79 – 29)] ÷ (79– 29) × 100% = -100%.

Multiple types of responders will be summarized at each scheduled post-Baseline visit:

- ≥20%, ≥30%, ≥50%, ≥75% and 100% reduction in the CMAI total score from Baseline
- ≥20%, ≥30%, ≥50%, ≥75% and 100% reduction in NPI-C combined agitation and aggression domain scores from Baseline
- ≥20%, ≥30%, ≥50%, ≥75% and 100% reduction in the individual agitation and aggression NPI-C domain scores from Baseline
- mADCS-CGIC score of 1 (marked improvement) or 2 (moderate improvement)

Other responder criteria based on point changes may be included based on examination of cumulative distribution functions and anchoring to the ADCS-CGIC.

For each of these responder analyses, the proportion of responders will be summarized by double-blind treatment group at each timepoint using observed cases and also with missing values imputed as non-response.

14 SAFETY ANALYSES

The safety summaries will be presented based on the Safety Analysis Set. Safety variables include AEs, clinical laboratory variables, vital signs, physical examinations, ECG and GCAS variables. Safety variables will be summarized by double-blind treatment group and overall using descriptive statistics. For each continuous measure in clinical laboratory variables, vital signs, and electrocardiogram, change from Baseline results will be presented in two ways as specified in Section 5.1.

14.1 Adverse Events

All AEs will be coded using the MedDRA Version 19.0 or newer.

An AE (classified by preferred term) will be considered a TEAE if it started on and after the first open-label study dose date and no later than the last open-label study dose date + 30 days. If any AE started before the completion of procedures at the Week 12 visit in Study ACP-103-032, it will be considered as TEAE in Study ACP-103-032 but not in Study ACP-103-033.

The event counts and the number and percentage of subjects reporting TEAEs will be tabulated by SOC and preferred term; and, by SOC, preferred term, and maximum severity. If more than 1 AE occurs with the same preferred term for the same subject, then the subject will be counted only once for that preferred term using the most severe occurrence for the summarization by severity. In addition, the event counts and the number and percentage of subjects with TEAEs classified by the Investigators as related to the study drug, with most frequently reported TEAEs (preferred terms reported by ≥5% of subjects in the Safety Analysis Set, with treatment-emergent serious AEs (TESAEs), with fatal AEs (i.e. events that cause death), and with TEAEs leading to discontinuation of study drug will be summarized by SOC and preferred term. These tables will be sorted alphabetically by SOC and then by descending subject frequency for the preferred terms in the Safety Analysis Set within each SOC.

The event counts and the number and percentage of subjects with any TEAEs will also be tabulated by preferred term without SOCs. This table will be sorted by descending subject frequency in the Safety Analysis Set.

An AE listing by subject will display all events, including those which are not treatment-emergent, and will include the verbatim term in addition to the MedDRA SOC and preferred term. This listing will also include all relevant eCRF data associated with the event: e.g. date of onset, date resolved, date of open-label first dose, date of open-label last dose, severity, frequency, outcome, relationship to study drug, action taken with study drug, and required therapy. When a date is presented, the study day associated with the date will also be displayed. Separate listings will be presented for subjects with SAEs, subjects with AEs leading to discontinuation and subjects who died (if any). In these listings, an indicator for treatment-emergent events and the dose at onset will also be included.

14.2 Clinical Laboratory Variables

Clinical laboratory tests are performed at Baseline, Weeks 4, 12, 28 and 52/ET visits.

- Chemistry serum tests include the following

- Sodium (Na), potassium (K), chloride (Cl), phosphorus (P), calcium (Ca), carbon dioxide (CO_2), blood urea nitrogen (BUN), creatinine (CR), uric acid
- Alanine aminotransferase (ALT/SGPT), aspartate aminotransferase (AST/SGOT), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), total bilirubin (TBIL), lactate dehydrogenase (LDH)
- HbA1c, glucose
- Albumin (ALB), total protein
- Prolactin
- Creatine kinase (CK)/creatine phosphokinase (CPK)
- Lipid panel
    - Total cholesterol, HDL-cholesterol, triglycerides, LDL-cholesterol, cholesterol/HDL ratio, Non-HDL cholesterol
- Serum pregnancy test for women of childbearing potential

- Hematology tests include the following:
  - Complete blood count (CBC) including:
    - White blood cell (WBC) count
    - Complete differential (relative and absolute)
    - Hematocrit (Hct), hemoglobin, red blood cells (RBC), platelets
    - Reticulocytes

- Urinalysis tests include the following:
  - Blood, RBCs, WBCs, protein, glucose, ketones, specific gravity, pH
  - Urine pregnancy test for women of childbearing potential.

All laboratory test results are from a central laboratory and will be listed. The listings will include date and study day of collection. All units will be displayed in Système International [SI] units. Out of range values will be flagged in the data listings (e.g., 'L' or 'H').

Clinical laboratory values reported as continuous values for hematology, chemistry and urinalysis will be summarized using descriptive statistics at Baseline and each scheduled post-Baseline visits. The change from Baseline values will also be summarized at each scheduled post-Baseline visits. The overall minimum and maximum post-Baseline observed

and change from Baseline values will also be summarized. For hemoglobin, hematocrit and uric acid, the above summaries will be presented for each gender as well as for both genders combined. For urinalysis with categorical results, the number and percentage of subjects will be tabulated by category at Baseline and each scheduled post-Baseline visits. For the categorical urinalysis by-visit summary, the denominator is the total number of subjects with non-missing values for the given parameter, visit and double-blind treatment group.

The laboratory values will also be summarized in shift tables to determine the number and percentage of subjects with values classified as below (low), within (normal) and above (high) normal ranges at each scheduled post-Baseline visits relative to the same classification at the Baseline visit. The shifts from Baseline to overall post-Baseline minimum and overall post-Baseline maximum will also be presented. For the by-visit shift summary, the denominator is the number of subjects with non-missing values at Baseline and the given visit for the given parameter and double-blind treatment group. For the summaries of shift to the overall post-Baseline minimum or maximum, the denominator is the number of subjects with non-missing Baseline and at least 1 post-Baseline value for the given parameter and double-blind treatment group. For hemoglobin, hematocrit and uric acid, the shift summaries will be presented for each gender as well as for both genders combined.

Number and percentage of subjects with potentially clinically important (PCI) laboratory values at each scheduled post-Baseline visits and overall post-Baseline will be summarized by Baseline status (all and within normal range) for selected parameters. PCI criteria are listed in Table 5 and Table 6. For the overall post-Baseline summaries, all post-Baseline values will be considered, including unscheduled and out of window values. Subjects with multiple PCI values for a given parameter will be counted only once for that parameter. For the by-visit summary, the numerator for the percentage is the number of subjects with a post-Baseline PCI laboratory value for the given parameter, visit and double-blind treatment group, and the denominator is the number of subjects with non-missing values for the given parameter, visit and double-blind treatment group. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI laboratory value for the given parameter and double-blind treatment group and the denominator is the number of subjects with at least 1 post-Baseline laboratory value for the given parameter and double-blind treatment group. For hemoglobin, hematocrit and uric acid, the count and percentage of subjects with PCI values will be presented for each gender as well as for both genders combined. A listing of all PCI values in study ACP-103-033 will be provided. This listing will include all observations from study ACP-103-033 for those subjects and parameters for which at least 1 PCI value (including Baseline) was observed.

Table 5 Criteria for Potentially Clinically Important Laboratory Values – Hematology and Chemistry

Analyte	Conventional Unit	Low PCI Criteria	High PCI Criteria	SI Unit	Low PCI Criteria	High PCI Criteria
Hematology (whole blood)						
Hemoglobin (male)	g/dL	<11	>18	g/L	<110	>180
Hemoglobin (female)	g/dL	<10	>17	g/L	<100	>170
Hematocrit (male)	%	<30	>55	L/L	<0.3	>0.55
Hematocrit (female)	%	<30	>50	L/L	<0.3	>0.5
Leukocyte (White Blood Cell Count)	x 10^3/uL	≤2.8	≥15	x 10^9/L	≤2.8	≥15
Neutrophils	x 10^3/uL	≤1.5	No upper limit	x 10^9/L	≤1.5	No upper limit
Platelet Count	x 10^3/uL	≤75	≥700	10^9/L	≤75	≥700
Chemistry (serum or plasma)						
ALT (SGPT)	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
AST (SGOT)	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
Total Bilirubin	mg/dL	No lower limit	≥1.5 ULN	umol/L	No lower limit	≥1.5 ULN
BUN	mg/dL	No lower limit	≥30.0	mmol/L	No lower limit	≥10.71
Creatine Kinase (CK)	U/L	No lower limit	≥3 ULN	U/L	No lower limit	≥3 ULN
Sodium	mEq/L	≤125	≥155	mmol/L	≤125	≥155
Potassium	mEq/L	≤3.0	≥5.5	mmol/L	≤3.0	≥5.5
Calcium, total	mg/dL	<8.0	>11.0	mmol/L	<2.0	>2.75
Lactate Dehydrogenase (LDH)	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
Alkaline Phosphatase	U/L	No lower limit	≥3 X ULN	U/L	No lower limit	≥3 X ULN
Uric acid (male)	mg/dL	No lower limit	≥10.5	umol/L	No lower limit	≥624.75
Uric acid (female)	mg/dL	No lower limit	≥8.5	umol/L	No lower limit	≥505.75
Albumin	g/dL	≤2.6	≥6.0	g/L	≤26	≥60
Total Protein	g/dL	≤5.0	≥10.0	g/L	≤50	≥100
Chloride	mEq/L	≤85	≥120	mmol/L	≤85	≥120
Glucose (random)	mg/dL	≤45.1	≥200.0	mmol/L	≤2.48	≥11
Serum Creatinine	mg/dL	Not Applicable	>1.5 ULN	umol/L	Not Applicable	>1.5 ULN
Triglycerides	mg/dL	Not Applicable	>300	mmol/L	Not Applicable	>3.39
Gamma-Glutamyl Transferase (GGT)	U/L	Not Applicable	≥3 ULN	U/L	Not Applicable	≥3 ULN

Table 6 Criteria for Potentially Clinically Important Laboratory Values – Urinalysis

Urinalysis (qualitative dipstick)	Low PCI Criteria	High PCI Criteria
Blood (Occult Blood)	Not Applicable	≥+2
Protein	Not Applicable	≥+2
Glucose	Not Applicable	≥+2

The pregnancy results (positive or negative) for female subjects will be presented in a listing.

14.3 Vital Signs

Vital signs are assessed at Baseline, Weeks 2, 4, 8, 12, 20, 28, 36, 44, and 52/ET visits.

Vital signs including weight, height (measured at ACP-103-032 Screening visit), and the derived BMI will be summarized using descriptive statistics at Baseline and all scheduled post-Baseline visits. The change from Baseline values will also be summarized at the scheduled post-Baseline visits.

Vital sign values will be considered PCI if they meet both the observed value criteria and the change from Baseline criteria listed in Table 7. The number and percentage of subjects with post-Baseline values that are PCI will be summarized at each scheduled post-Baseline visit and for overall post-Baseline. For the overall post-Baseline summaries, all post-Baseline values will be considered, including unscheduled and out of window values. Subjects with multiple PCI values for a given parameter will be counted only once for that parameter. For the by-visit summary, the numerator for the percentage is the number of subjects with a post-Baseline PCI vital sign for the given parameter, visit and double-blind treatment group and the denominator is the number of subjects with non-missing values for the given parameter, visit and double-blind treatment group. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI vital sign for a given parameter and double-blind treatment group and the denominator is the number of subjects at least 1 post-Baseline vital sign for the given parameter and double-blind treatment group. A listing of all PCI values in study ACP-103-033 will be provided. This listing will include all observations from study ACP-103-033 for those subjects and parameters for which at least 1 PCI value (including Baseline) was observed.

Table 7 Criteria for Potentially Clinically Important Vital Signs

Vital Sign Parameter	Unit	Criteria[a]		
		Observed Value	And/Or	Change Relative to Baseline
Systolic blood pressure	mmHg	≥180	And	Increase of ≥20
		≤90	And	Decrease of ≥20
Diastolic blood pressure	mmHg	≥105	And	Increase of ≥15
		≤50	And	Decrease of ≥15
Pulse	bpm	≥120	And	Increase of ≥15
		≤50	And	Decrease of ≥15
Weight	kg	Not Applicable		Increase of ≥7%
				Decrease of ≥7%

[a] A post-Baseline value is considered as a PCI value if it meets both criteria for observed value and change from Baseline.

14.4 Electrocardiogram (ECG)

Electrocardiograms are performed at Baseline, Weeks 4, 12, 28 and 52/ET visits. All tracings will be evaluated by a central reading laboratory. ECG data summaries will be performed using the centrally evaluated data.

Observed values of ECG variables (heart rate, PR interval, RR interval, QRS interval, QT interval, and QTc intervals) and the changes from Baseline at each assessment timepoint will be summarized at Baseline and each scheduled post-Baseline visit.

QTc intervals include QTcB (Bazett's formula) and QTcF (Fridericia's formula). QTcF will also be categorized into the following categories (msec) and the number and percentage of subjects in each category will be summarized at each scheduled visit and overall post-Baseline maximum:

- Observed: ≤450, 451 to 480, 481 to 500, and >500
- Change from Baseline: ≤10, 11 to 30, 31 to 60, and >60

Electrocardiogram values will be considered PCI if they meet the criteria listed in Table 8. The number and percentage of subjects with post-Baseline PCI values will be summarized at each scheduled post-Baseline visit and for overall post-Baseline. For the overall post-Baseline summaries, all post-Baseline values will be considered, including unscheduled and out of window values. Subjects with multiple PCI values for a given parameter will be counted only once for that parameter. For the by-visit summary, the numerator for the percentage is the number of subjects with a post-Baseline PCI ECG for the given parameter, visit and double-blind treatment group and the denominator is the number of subjects with non-missing values for the given parameter, visit and double-blind treatment group. For the

overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI ECG for the given parameter and double-blind treatment group and the denominator is the number of subjects with at least 1 post-Baseline ECG value for the given parameter and double-blind treatment group. A listing of all PCI values in study ACP-103-033 will be provided. This listing will include all observations from study ACP-103-033 for those subjects and parameters for which at least 1 PCI value (including Baseline) was observed

Table 8 Criteria for Potentially Clinically Important ECG Values

ECG Parameter	Unit	High PCI Criteria
QRS Interval	msec	≥120
PR Interval	msec	≥220
QTcB or QTcF	msec	>500
QTcB or QTcF: change from Baseline	>60 msec	

14.5 Physical Examination

Physical examination is performed at Baseline, Weeks 12 and 52/ET visits.

Physical examination results (normal, abnormal, and not done) at Baseline, Weeks 12 and 52 will be summarized in a frequency table by body system and visit.

14.6 Other Safety Variables

14.6.1 Suicidal Ideation and Behavior

GCAS is assessed at Baseline, Weeks 2, 4, 8, 12, 20, 28, 36, 44, 52/ET and Follow-up (if applicable) visits.

The number and percentage of subjects for each GCAS rating (0-4) based on clinician's assessment will be tabulated by visit, using the number of subjects with non-missing GCAS score for the given visit and double-blind treatment group as the denominator for the percentages. The number and percentage of subjects reporting any post-Baseline GCAS score of 3 or 4 based on clinician's assessment will also be tabulated, using the number of subjects with at least 1 post-Baseline GCAS score for the given double-blind treatment group as the denominator for the percentages.

15 CLINICAL PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

Not applicable.

16 INTERIM ANALYSIS

No interim analysis is planned in this study.

17 DATA MONITORING/REVIEW COMMITTEE

There is no Safety Monitoring Committee (SMC) for this study. Safety data are monitored throughout the study and aggregate safety reports are produced and reviewed approximately quarterly.

18 COMPUTER METHODS

All data summaries will be performed using Version 9.4 (or newer) of SAS® (SAS® Institute, Inc., Cary, North Carolina) on a suitably qualified and validated environment.

Validation and quality control of the tables, listings and figures containing the results of the data summaries will follow appropriate standard operating procedures.

19 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

None.

20 REFERENCES

Rabinowitz J, Davidson M., De Deyn PP., Katz I., Brodaty H., Cohen-Mansfield J. (2005). "Factor Analysis of the Cohen-Mansfield Agitation Inventory in Three Large Samples of Nursing Home Patients With Dementia and Behavioral Disturbance" The American Journal of Geriatric Psychiatry; 13(11): 991–998.

21 APPENDICES

21.1 Summary of Version Changes

Version No:	Document History Description of Update	Author	Version Date
1.0	Original version	[redacted]	06 March 2019