STATISTICAL ANALYSIS PLAN

Protocol No.:	ACP-103-064
Protocol Title:	A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of Pimavanserin as Adjunctive Treatment for the Negative Symptoms of Schizophrenia (ADVANCE-2)
Drug:	Pimavanserin
Sponsor:	Acadia Pharmaceuticals Inc.
Version No. and Date	Version 4.0, January 11, 2024

SIGNATURE/APPROVAL PAGE

AUTHOR

See appended electronic signature page

PPD
Director, Biostatistics
Acadia Pharmaceuticals Inc.

Date

APPROVERS

See appended electronic signature page

PPD MD
PPD Clinical Development
Acadia Pharmaceuticals Inc.

Date

See appended electronic signature page

PPD Ph.D.
Executive Director, Biostatistics
Acadia Pharmaceuticals Inc.

Date

Signature Page for ACP-103-064 Statistical Analysis Plan v4

Approve	PPD s 12-Jan-2024 22:19:24 GMT+0000

Approve	PPD 12-Jan-2024 22:26:46 GMT+0000

Approve	PPD ics 12-Jan-2024 22:37:53 GMT+0000

Signature Page for ACP-103-064 Statistical Analysis Plan v4

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AE	adverse event
AIMS	Abnormal Involuntary Movement Scale
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
BACS	Brief Assessment of Cognition in Schizophrenia
BARS	Barnes Akathisia Rating Scale
BLQ	below the limit of quantification
BMI	body mass index
CDSS	Calgary Depression Scale for Schizophrenia
CGI-SCH-I	Clinical Global Impression – Improvement scale
CGI-SCH-S	Clinical Global Impression – Severity scale
CMH	Cochran-Mantel-Haenszel
COVID-19	coronavirus disease 2019
C-SSRS	Columbia-Suicide Severity Rating Scale
DSMB	Data and Safety Monitoring Board
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ECG	Electrocardiogram
eCRF	electronic case report form
EOT	End-of-Treatment
ET	Early Termination
FAS	Full Analysis Set
FDA	Food and Drug Administration
GSD	Guidance for Site Documentation and Data Management Querying of Data Impacted by COVID-19
ICE	Intercurrent event
LOE	lack of efficacy
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MMRM	mixed model for repeated measures
MNAR	missing not at random
Msec	Milliseconds
NSA-16	Negative Symptom Assessment-16
OC	observed cases
PANSS	Positive and Negative Syndrome Scale

PCI	potentially clinically important
PD	Pharmacodynamic
PK	Pharmacokinetic
PMM	Pattern Mixture Model
PSP	Personal and Social Performance scale
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
SAS	Simpson-Angus Extrapyramidal Side Effects Scale
SAS®	Statistical Analysis System
SD	standard deviation
SE	standard error
SOC	system organ class
TEAE	treatment-emergent adverse event
WoRQ	Work Readiness Questionnaire

1. INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the statistical analyses of efficacy and safety data as described in the final study protocol Amendment 3 dated 05OCT2022.

For Argentina, a country-specific protocol amendment (Amendment 2-AR finalized 19 March 2021) specifies additional hepatitis C virus (HCV) and human immunodeficiency virus (HIV) testing at Screening and additional pregnancy tests at Week 12, 18 and 24 for all female subjects of childbearing potential to be conducted only for subjects enrolled in Argentina.

Specifications for tables, figures, and listings are contained in a separate document. Statistical analyses for population pharmacokinetic (PK) and PK/pharmacodynamics (PD) modeling will be presented in a separate report and therefore will not be included in this SAP.

This plan should be read in conjunction with the study protocol and the electronic case report forms (eCRFs).

2. OBJECTIVES

2.1. Primary Objective

The primary objective of the study is to evaluate the efficacy of pimavanserin compared with placebo in the adjunctive treatment of the negative symptoms of schizophrenia.

2.2. Secondary Objective

The secondary objectives of the study are to evaluate the effect of adjunctive pimavanserin compared with adjunctive placebo on

- Global impression of severity of negative symptoms of schizophrenia
- Global improvement of symptoms of negative symptoms of schizophrenia
- Personal and social performance
- Response to treatment in adults experiencing negative symptoms of schizophrenia

3. STUDY DESIGN

3.1. General Study Design

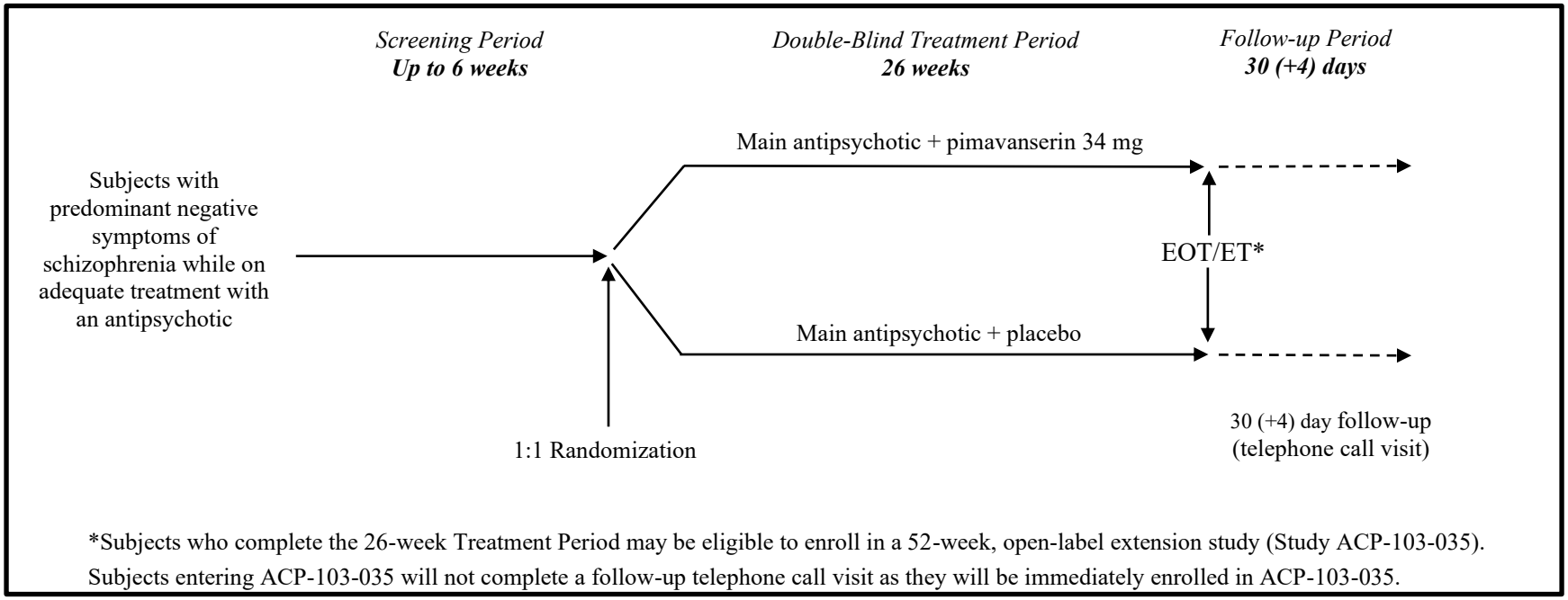
This study will be conducted as a Phase 3, 26-week, randomized, double-blind, placebo-controlled, multicenter, multinational outpatient study in subjects with schizophrenia who have predominant negative symptoms while on adequate treatment with an antipsychotic. Schizophrenia is defined by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), confirmed by a customized module of the Structured Clinical Interview for DSM-5, Clinical Trials Version (SCID-5-CT). Negative symptoms are considered predominant when other symptoms of schizophrenia, particularly positive symptoms such as delusions and hallucinations, are relatively mild and well controlled.

This study will screen approximately 692 subjects and randomize approximately 426 subjects (213 subjects per treatment group) with predominant negative symptoms of schizophrenia across approximately 88 study sites in 12 countries. The daily dose of the main antipsychotic must remain stable and may not be changed from Screening through the duration of the study. Subjects will participate in the study for up to 36 weeks, including a Screening Period of up to 6 weeks, a 26-week Treatment Period, and a 30 (+4) day safety follow-up (telephone visit) for those subjects who discontinue prematurely from the study or who do not enroll in the 52-week, open-label extension study (Study ACP-103-035).

Study drug will be administered under double-blind conditions throughout the Treatment Period. Clinic visits occurring after Baseline will be conducted at Weeks 2, 4, 8, 14, 20, and 26 (End-of-Treatment [EOT]/Early Termination [ET] visit).

Figure 1 illustrates the study design.

Figure 1 Schematic of Study Design



*Subjects who complete the 26-week Treatment Period may be eligible to enroll in a 52-week, open-label extension study (Study ACP-103-035). Subjects entering ACP-103-035 will not complete a follow-up telephone call visit as they will be immediately enrolled in ACP-103-035.

3.2. Schedule of Assessments

The schedule of events and assessments for the study is presented in Appendix A.

3.3. Randomization

Eligible subjects will be randomized into 1 of 2 treatment groups (pimavanserin or placebo) in a 1:1 ratio using an interactive response technology (IRT) system. The randomization will be stratified by geographic region (North America, Europe, or rest of world). The assignments will be based on a pre-generated permuted-block randomization schedule.

3.4. Blinding

This is a double-blind study. Blinding will be assured by restricting access of Investigators and Sponsor personnel and/or designee to the treatment codes, and providing identical tablets and packaging for the pimavanserin and placebo treatments. Neither the subjects nor the study personnel at the clinical sites will know which treatment is administered to each subject.

3.5. Determination of Sample Size

The primary efficacy endpoint is the change from Baseline to Week 26 in the Negative Symptom Assessment-16 (NSA-16) total score.

Let Δ be the difference in the mean change from Baseline to Week 26 in the NSA-16 total score between the pimavanserin and the placebo groups:

The null hypothesis for the primary efficacy endpoint is: $\Delta = 0$

The alternative hypothesis for the primary efficacy endpoint is: $\Delta \neq 0$

Assuming the true difference in the mean change in the NSA-16 total score from Baseline to Week 26 is 3.0 points between the pimavanserin group and the placebo group, and the common standard deviation is 9 points, 191 subjects who complete the study per treatment group will provide at least 90% power to detect a difference between the pimavanserin group and the placebo group at a significance level of 0.05, using a 2-sided t-test.

Acadia used a dropout rate of 10% to calculate the sample size for this study. This dropout rate was informed based on the whole Full Analysis Set (FAS) dropout rate of 13.5% and the dropout rate of 8.8% based on the last dose of 34 mg in FAS observed in the recently completed study ACP-103-038 (ADVANCE). Adjusting for a potential dropout rate of up to 10%, approximately 426 subjects (213 subjects per treatment group) will be randomized.

3.6. Coronavirus Disease 2019

In the context of the global COVID-19 pandemic or other natural disaster, the mechanisms to record information on the potential impact of the COVID-19 pandemic on data itself, as well as data collection and integrity, will be implemented (as detailed in the "Guidance for Site Documentation and Data Management Querying of Data Impacted by COVID-19" [GSD] of the Data Management Plan). The impact of COVID-19 on the statistical analysis is discussed in each of the relevant sections of this SAP.

4. ANALYSIS SETS

Randomized Analysis Set

The Randomized Analysis Set will consist of all unique subjects who were randomized.

Note that if there are duplicate subjects (see Section 5.13 for details), only the identity under the first enrollment will be included in the Randomized Analysis Set.

Subjects will be classified according to the randomized treatment assignment.

Safety Analysis Set

The Safety Analysis Set will consist of all subjects who were randomized and received at least one dose of study drug.

Subjects will be classified according to the actual treatment received.

Full Analysis Set

The Full Analysis Set will consist of a subset of subjects in the Randomized Analysis Set who receive at least one dose of study drug, and have both a Baseline value and at least one post-Baseline value for the NSA-16 total score.

Subjects will be classified according to the randomized treatment assignment.

Per-protocol Analysis Set

The Per-protocol Analysis Set will consist of a subset of subjects in the Full Analysis Set without any protocol deviation that could potentially have a substantial impact on the primary efficacy outcome. The precise reasons for excluding subjects from the Per-protocol Analysis Set will be fully defined and documented prior to the clinical database lock.

Subjects will be classified according to the randomized treatment assignment.

Pharmacokinetics Analysis Set

For pimavanserin and AC-279 plasma concentration summaries, the Pharmacokinetics Analysis Set will consist of subjects with at least one measurable pimavanserin plasma concentration.

Subjects will be classified according to the actual treatment received.

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 standard deviations and standard errors will be presented to 2 more decimal places than the raw data. Unless otherwise specified, the maximum number of decimal places is 4 and values will be truncated to 4 decimal places in situations where there are more than 4 decimal places. Wherever possible data will be decimal aligned.

For categorical variables, summaries will include the number and percentage of subjects in each category. For demographic and baseline characteristics, the number of subjects and the percentage of subjects with missing data will be summarized (if applicable), and the denominator for percentages will be the total number of subjects in the given treatment group. Categories with zero counts will not have zero percentages displayed. Percentages will be presented to 1 decimal place.

Unless specified otherwise, all statistical tests will be 2-sided hypothesis tests performed at the significance level of 5% for main effects and all confidence intervals (CIs) will be 2-sided 95% CIs. P-values will generally be presented to 4 decimal places; values less than 0.0001 will be presented as <0.0001.

Values that are collected with “<” or “>” signs will generally be analysed as the numerical value without the sign in tables and figures. In listings, these data will be reported as collected with the sign.

5.2. Derived Variables

In general, the assessment scale total scores and subscores will be derived within the analysis datasets. In the event that total scores and/or subscores are also collected on the electronic case

report form (eCRF), the derived values will be used for all analyses. Both the raw and derived scores will be presented in listings.

5.2.1. Negative Symptom Assessment-16 (NSA-16)

The NSA-16 is assessed at Baseline, Weeks 2, 4, 8, 14, 20, and Week 26/ET visits.

The NSA-16 is a semi-structured interview and a validated scale containing 16 items for evaluating negative symptoms of schizophrenia. Negative symptoms represent the reduction or absence of emotional expression and volitional behaviors normally present in a healthy person. Items are scored based on behaviors during the interview (items 1 to 4, 6, 7, 9, 11, 15, 16) or previous 7 days (items 5, 8, 10, 12 to 14) on a 6-point scale from 1 to 6 (a score of 9 indicates the item is not ratable and therefore is equivalent to missing). The NSA-16 total score can range from a minimum of 16 to a maximum of 96, with higher scores denoting more severe negative symptoms in schizophrenia. As per the NSA-16 Manual Version 3.0, the "normal" (score = 1) reference is based on the comparison to a young person in their twenties without schizophrenia. For most items, this modification may not be necessary as there are rules for rating NSA-16 items which would supersede the age reference rule. In general, it is recommended that this condition not be changed in the limited instances where it is applicable as it is difficult if not impossible to know what is normal for someone in their 30s, 40s, 50s, 60s, 70s, 80s, etc. There are only two instances where this condition for rating is mentioned. The description for item 8 (reduced social drive) states, "the reference range is a normal 20 year old". Similarly, the description for item 14 (reduced daily activity) states, "if the subject is hospitalized rate his/her daily activity as you would for a young person who is not hospitalized and without regard for the limitations that the hospital routine may place on him/her." Raters are typically trained to focus on the scoring anchors, most of which do not require a reference population. It is not: (1) the same person at another point in time; (2) a healthy person of similar age, living under similar circumstances; or (3) another hospitalized person. The addition of the wording "without schizophrenia" was an update to Version 3.0 of the instruction manual.

The NSA-16 includes 5 domains:

- Communication: 4 Items (minimum = 4, maximum = 24)
  - Item 1 Prolonged time to respond
  - Item 2 Restricted speech quantity
  - Item 3 Impoverished speech content
  - Item 4 Inarticulate speech
- Emotion/affect: 3 Items (minimum = 3, maximum = 18)
  - Item 5 Emotion: Reduced range
  - Item 6 Affect: Reduced modulation of intensity
  - Item 7 Affect: Reduced display on demand

- Social involvement: 3 Items (minimum = 3, maximum = 18)
  - Item 8 Reduced social drive
  - Item 9 Poor rapport with interviewer
  - Item 10 Interest in emotional and physical intimacy
- Motivation: 4 Items (minimum = 4, maximum = 24)
  - Item 11 Poor grooming and hygiene
  - Item 12 Reduced sense of purpose
  - Item 13 Reduced hobbies and interest
  - Item 14 Reduced daily activity
- Retardation: 2 Items (minimum = 2, maximum = 12)
  - Item 15 Reduced expressive gestures
  - Item 16 Slowed movements

When calculating the domain score, if any item score is missing in a domain, then that particular domain score will be missing. When calculating the NSA-16 total score, if there are no more than 3 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 4 or more missing items then the total score will be missing.

In addition to the 5 domains above, there is a global negative symptoms rating which assesses the overall severity on a 7-point scale from 1 to 7, with higher scores denoting more severe negative symptoms in schizophrenia. Global negative symptoms rating score should not depend on any specific item from the NSA-16 or any other similar instrument. Instead, it represents the overall assessment of the frequency, severity and functional impact of the negative symptoms during the interview and rating period. Missing global negative symptoms rating scores will not be imputed.

5.2.2. Clinical Global Impression of Schizophrenia Scale – Severity (CGI-SCH-S)

The CGI-SCH-S is assessed at Screening, Baseline, Weeks 2, 4, 8, 14, 20, and Week 26/ET visits.

The CGI-SCH-S is a clinician-rated, 7-point scale that is designed to evaluate positive, negative, depressive, cognitive symptoms and overall severity in schizophrenia. For purpose of this study, only the negative symptoms are evaluated. The 7-point scores are: 1 = normal, not ill; 2 = minimally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill; 7 = among the most severely ill. Higher scores denote more severe negative symptoms in schizophrenia.

Missing CGI-SCH-S scores will not be imputed.

5.2.3. Clinical Global Impression of Schizophrenia Scale – Improvement (CGI-SCH-I)

The CGI-SCH-I is assessed at Weeks 2, 4, 8, 14, 20, and Week 26/ET visits.

The CGI-SCH-I is a clinician-rated, 7-point scale that is designed to evaluate change from Baseline in positive, negative, depressive, cognitive symptoms and overall severity in schizophrenia. For purpose of this study, only the changes in negative symptoms from Baseline are evaluated. The 7-point scores are: 1 = very much improved; 2 = much improved; 3 = minimally improved; 4 = no change; 5 = minimally worse; 6 = much worse; 7 = very much worse. Higher scores denote worse or less improvement of negative symptoms in schizophrenia.

Missing CGI-SCH-I scores will not be imputed.

5.2.4. Positive and Negative Syndrome Scale (PANSS)

The PANSS is assessed at Screening, Baseline, Week 14 and Week 26/ET visits.

The PANSS is a 30-item scale used to evaluate the presence, absence, and severity of schizophrenia symptoms. Items are scored over the past week (7 days) on the following 7-point scale: 1 = absent; 2 = minimal; 3 = mild; 4 = moderate; 5 = moderate-severe; 6 = severe; 7 = extreme. The PANSS total score can range from a minimum of 30 to a maximum of 210, with higher scores denoting more severe symptoms.

The following are the 3 PANSS subscales and the PANSS items that define each subscale:

- Positive Scale: 7 Items (minimum = 7, maximum = 49)
  - P1 Delusions
  - P2 Conceptual disorganization
  - P3 Hallucinatory behavior
  - P4 Excitement
  - P5 Grandiosity
  - P6 Suspiciousness/persecution
  - P7 Hostility
- Negative Scale: 7 Items (minimum = 7, maximum = 49)
  - N1 Blunted affect
  - N2 Emotional withdrawal
  - N3 Poor rapport
  - N4 Passive/apathetic social withdrawal
  - N5 Difficulty in abstract thinking
  - N6 Lack of spontaneity and flow of conversation
  - N7 Stereotyped thinking
- General Psychopathology Scale: 16 Items (minimum = 16, maximum = 112)
  - G1 Somatic concern
  - G2 Anxiety
  - G3 Guilt feelings
  - G4 Tension
  - G5 Mannerisms and posturing
  - G6 Depression

- G7 Motor retardation
- G8 Uncooperativeness
- G9 Unusual thought content
- G10 Disorientation
- G11 Poor attention
- G12 Lack of judgment and insight
- G13 Disturbance of volition
- G14 Poor impulse control
- G15 Preoccupation
- G16 Active social avoidance

The following are the 5 Marder factors and the PANSS items that define each factor:

- Negative Symptoms: 7 Items (minimum = 7, maximum = 49)
  - N1 Blunted affect
  - N2 Emotional withdrawal
  - N3 Poor rapport
  - N4 Passive/apathetic social withdrawal
  - N6 Lack of spontaneity and flow of conversation
  - G7 Motor retardation
  - G16 Active social avoidance
- Positive Symptoms: 8 Items (minimum = 8, maximum = 56)
  - P1 Delusions
  - P3 Hallucinatory behavior
  - P5 Grandiosity
  - P6 Suspiciousness/persecution
  - N7 Stereotyped thinking
  - G1 Somatic concern
  - G9 Unusual thought content
  - G12 Lack of judgment and insight
- Disorganized Thought: 7 Items (minimum = 7, maximum = 49)
  - P2 Conceptual disorganization
  - N5 Difficulty in abstract thinking
  - G5 Mannerisms and posturing
  - G10 Disorientation
  - G11 Poor attention
  - G13 Disturbance of volition
  - G15 Preoccupation
- Uncontrolled Hostility/Excitement: 4 Items (minimum = 4, maximum = 28)
  - P4 Excitement
  - P7 Hostility
  - G8 Uncooperativeness
  - G14 Poor impulse control

- Anxiety/Depression: 4 Items (minimum = 4, maximum = 28)
  - G2 Anxiety
  - G3 Guilt feelings
  - G4 Tension
  - G6 Depression

For the PANSS subscales, if more than 1 item score is missing within the subscale, then that particular subscale score will be missing. When there is only 1 item missing for a subscale, then the subscale score will be calculated by the arithmetic mean of non-missing items for that given subscale, subject and timepoint multiplied by the number of items within the given subscale, rounded to the nearest integer. After the missing item(s) is set as the arithmetic mean within the subscale(s), then the PANSS total score and Marder factor scores will be computed without further imputation. The PANSS total score is calculated as the sum of 3 subscale scores. If any of 3 subscale scores is missing, the total score will be missing. A Marder factor score will be missing if any item score is missing for that factor, after the imputations within the subscales were applied.

5.2.5. Personal and Social Performance Scale (PSP)

The PSP is assessed at Baseline, Weeks 2, 8, 14 and Week 26/ET visits.

The PSP is a validated 100-point (1 to100) single-item rating scale to assess the psychosocial functioning of subjects with schizophrenia. Ratings are based on four main areas: (a) socially useful activities, including work and study; (b) personal and social relationships, (c) self-care; and (d) disturbing and aggressive behaviors. The time period assessed is “past month” and does not take into account suicidal risk and behavior. Higher scores denote better psychosocial functioning: scores of 71-100 reflect only mild difficulties; 31-70 reflect manifest disabilities of various degrees; 1-30 reflect functioning so poor that intensive support or supervision is needed.

Missing PSP scores will not be imputed.

5.2.6. Brief Assessment of Cognition in Schizophrenia (BACS)

The BACS is assessed at Baseline and Week 26/ET visits.

The BACS is a performance-based assessment that measure treatment-related changes in cognition and assesses 6 domains, including:

- Verbal memory and learning (verbal memory task): a subject is given 5 attempts to remember 15 words and recall as many words as possible; raw scores can range from 0 to 75.

- Working memory (digit sequencing task): a subject is presented with clusters of numbers of increasing length and then asked to repeat in order from the lowest to highest length; raw scores can range from 0 to 28.
- Motor function (token motor task): a subject is given 100 plastic tokens and asked to place 2 tokens at a time within a container as quickly as possible within 60 seconds; raw scores can range from 0 to 100.
- Verbal fluency (semantic and letter fluency): a subject is asked to name as many words as possible within a specific category (e.g., animal names), and to name words that begin with a specific letter (e.g., F and S) within 60 seconds, respectively; raw scores can range from 0 to 225.
- Attention and speed of processing (symbol coding task): a subject is asked to write matching numbers from 1 to 9 to symbols within 90 seconds; raw scores can range from 0 to 110.
- Executive function (Tower of London): a subject is shown two pictures of three balls of different colors arranged on three different pegs, whereby the balls were arranged differently on each picture and the subjects were asked to give the total number of times the balls in one picture needed to be moved in order to end with the arrangement in the other picture; raw scores can range from 0 to 22.

For each domain, higher scores reflect better cognition and raw scores will be converted to age and gender corrected normalized scores. The BACS composite score (measure of overall cognitive functioning) will be calculated as the mean of the normalized scores from the 6 domains. The composite score for subjects with missing domain are imputed using the average normalized score of the non-missing domains. For ≤ 2 missing BACS domains, the average of the non-missing domain normalized scores are imputed for the composite score. If there are more than 2 missing BACS domains, the composite score should be set to missing. If a domain score is missing, the above imputation is only used to impute the composite score, but the domain score should be set to missing. See Appendix B for details.

5.2.7. Work Readiness Questionnaire (WoRQ)

The WoRQ is assessed at Baseline, Weeks 8, 14 and Week 26/ET visits.

The WoRQ consists of 7 statements that the Investigator rates on a 4-point scale (“strongly agree” [1 point] to “strongly disagree” [4 points], with “strongly agree” being the most indicative of work readiness). Using the ratings of these 7 statements as an aid, the Investigator provides a global yes/no judgment about the subject’s readiness to work. The WoRQ total score is calculated by adding scores for the 7 statements. Lower scores denote better work readiness. If

more than 1 score is missing, then the total score will be missing. If there is only 1 missing item score, the total score will be calculated by the arithmetic mean of non-missing items for that subject and timepoint multiplied by 7, rounded to the nearest integer.

5.2.8. Calgary Depression Scale for Schizophrenia (CDSS)

The CDSS is assessed at Screening, Baseline, Week 14 and Week 26/ET visits.

The CDSS is a 9-item scale that was developed specifically to assess the level of depression in schizophrenia. It was originally developed to differentiate depressive symptoms from negative symptoms. Items are scored over the past 2 weeks (14 days) on the following 4-point scale: 0 = absent; 1 = mild; 2 = moderate; 3 = severe. The total score is the sum of the 9 item scores which can range from 0 to 27, with higher scores denoting more severe depression.

Missing CDSS item scores will not be imputed. The total score will be missing if any item score is missing.

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

The C-SSRS is assessed at Screening, Baseline, Weeks 2, 4, 8, 14, 20, and Week 26/ET visits.

The C-SSRS monitors changes in suicidal thinking and behavior over time, in order to determine risk. Four constructs are measured: the severity of ideation, the intensity of ideation, behavior, and lethality.

The C-SSRS Baseline/Screening version will be completed at the Screening visit and the version assessing information since the last visit will be completed at all following visits (including the Baseline visit). The C-SSRS results for each subject should be reviewed by the Investigator at each visit. If at any time the C-SSRS results for a given subject reveal potential suicidality, then the Investigator should assess the clinical significance of such results. If a clinically significant risk of suicidality is identified for a subject, then the Investigator should discontinue the subject and implement appropriate treatment.

There are 5 questions about suicidal ideation, representing 5 types of suicidal ideation: wish to be dead; non-specific active suicidal thoughts; active suicidal ideation with any methods (not plan) without intent to act; active suicidal ideation with some intent to act, without specific plan; active suicidal ideation with specific plan and intent. If a subject answers “yes” to any of these 5 questions, this subject will be counted as having suicidal ideation.

There are 5 questions about suicidal behavior, representing 5 types of suicidal behavior: preparatory acts or behavior; aborted attempt; interrupted attempt; actual attempt; suicide. If a subject answers “yes” to any of these 5 questions, this subject will be counted as having suicidal behavior.

Missing C-SSRS item scores will not be imputed.

5.2.10. Abnormal Involuntary Movement Scale (AIMS)

The AIMS is assessed at Screening, Baseline, Weeks 2, 14 and Week 26/ET visits.

The AIMS is a rating scale that was designed to measure involuntary movements known as tardive dyskinesia. The AIMS has a total of 12 items rating involuntary movements of various areas of the subject's body.

- Items 1 to 7 assess the severity of dyskinesia (orofacial, extremity and truncal movements) are rated on a 5-point scale of severity: 0 (none), 1 (minimal; may be extreme normal), 2 (mild), 3 (moderate), and 4 (severe).
- Items 8 and 9 assess the overall severity and incapacitation, and are also rated on a 5-point scale of severity: 0 (none, normal), 1 (minimal), 2 (mild), 3 (moderate), and 4 (severe).
- Item 10 assesses the subject’s level of awareness of the movements with associated distress and is rated on a 5-point scale: 0 (no awareness), 1 (aware, no distress), 2 (aware, mild distress), 3 (aware, moderate distress), and 4 (aware, severe distress).
- Items 11 and 12 refer to dental status and the responses are yes (scored as 1) or no (scored as 0).

The AIMS total score is the sum of the 12 item scores which can range from 0 to 42, with higher scores denoting more severe dyskinesia symptoms.

Missing AIMS item scores will not be imputed. The total score will be missing if any item score is missing.

5.2.11. Barnes Akathisia Scale (BARS)

The BARS is assessed at Screening, Baseline, Weeks 2, 14 and Week 26/ET visits.

The BARS is a 4-item, physician-administered scale that assesses the severity of drug-induced akathisia. Items 1 to 3 assess the objective presence and frequency of akathisia, the subjective awareness of restlessness, and the subjective distress related to restlessness. These 3 items are rated on a 4-point scale from 0 to 3 and the total score is the sum of these 3 item scores, which can range from 0 to 9. Additionally, there is a global clinical assessment of akathisia which is rated on a 6-point scale from 0 to 5. For total or global clinical assessment scores, higher scores denote more severe akathisia symptoms.

Missing BARS item scores will not be imputed. The total score will be missing if any non-global item score is missing.

5.2.12. Simpson Angus Extrapyramidal Side Effect Scale (SAS)

The SAS is assessed at Screening, Baseline, Weeks 2, 14 and Week 26/ET visits.

The SAS is a 10-item physician-administered scale commonly used for the assessment of parkinsonian movement disorder related to psychiatric drug treatment. One item on the SAS measures gait/hypokinesia; 6 items measure rigidity (arm-dropping, shoulder shaking, elbow rigidity, wrist rigidity or fixation of position, head rotation, and akathisia); and 3 items measure glabella tap, tremor, and salivation, respectively. The grade of severity of each item is rated using a 5-point scale from 0 to 4. The SAS total score is the sum of the 10 items, which can range from 0 to 40, with higher scores denoting more severe parkinsonian symptoms.

Missing SAS item scores will not be imputed. The total score will be missing if any item score is missing.

5.3. Data Presentation Conventions

- 1 year = 365.25 days. Year is calculated as (days/365.25) and will be rounded to 1 digit after the decimal point (tenths) for presentation purposes.
- 1 month = 30.4375 days. Month is calculated as (days/30.4375) and will be rounded to 1 digit after the decimal point (tenths) for presentation purposes.
- 1 pound = 0.454 kg and 1 kg = 2.2 pounds.
- 1 inch = 2.54 cm and 1 cm = 0.3937 inches.
- Body mass index (BMI) calculated as [weight (kg)/height $(m)^2$]

5.4. Study Day

If the date of assessment occurs on or after the first dose date, then study day will be calculated as (date of assessment – date of first dose) + 1. If the date of assessment occurs prior to the first dose date, then study day will be calculated as (date of assessment – date of first dose). There is no study day 0.

5.5. Baseline Definition

Baseline data are defined as data collected which are prior to the administration of the first dose. If there is more than one value on or prior to Study Day 1, the value closest to and prior to the receipt of the first dose, whether scheduled or unscheduled, will be used as the Baseline value.

5.6. Analysis Visit Windows

Efficacy, safety and PK assessments will be summarized by analysis visit as presented in Table 1 below.

Table 1 Analysis Visit Windows

Endpoint Type	Analysis Visit Name	Target Study Day[1]	Study Day Interval
Efficacy[2]	Baseline	1	≤1
	Week 2	15	7 to 21
	Week 4	29	22 to 36
	Week 8	57	50 to 64
	Week 14	99	92 to 106
	Week 20	141	134 to 148
	Week 26	183	176 to 190
Safety and plasma concentration data	Baseline	1	≤1
	Week 2	15	2 to 21
	Week 4	29	22 to 42
	Week 8	57	43 to 77
	Week 14	99	78 to 119
	Week 20	141	120 to 161
	Week 26	183	≥162

1 Study day = assessment date - first dose date + 1 if the assessment date ≥ first dose date, otherwise study day = assessment date – first dose date. Study day 1 is the day of first administration of study drug (adjunctive pimavanserin or adjunctive placebo).

2 Efficacy assessment collected more than seven days after the last double-blind dosing date will be excluded from the analyses.

5.6.1. Unscheduled Assessments

Both scheduled and unscheduled assessments, including early termination visits, will be considered for planned timepoint analyses based on the above analysis visit windowing rules. All assessments will be presented in data listings.

5.6.2. 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 analyses. In these analyses, if two assessments are equidistant from the target day then the chronologically last assessment will be used.

For safety analysis where the most extreme values should be selected (e.g. overall post-baseline minimum, overall post-baseline maximum, and potentially clinically important values), 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.

When ECG is collected in triplicate, the average of the triplicate will be considered as one assessment for the analyses.

5.7. 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 missing last dose date of study drug will be imputed using the last expected dosing date which is defined as the minimum of the non-missing drug return date of the last dispensed drug kits, the last drug dispensed date plus the number of days that the dosing would continue per protocol or the end-of-treatment/early termination visit (EOT/ET) date, whichever occurs earlier. For the incomplete last dose date of the study drug, the imputation will be compared against the last expected dosing date defined as above. Detailed algorithms will be documented in a separate specifications document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

For 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.

5.8. 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 medication use 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 a separate specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

5.9. 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 a separate specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

5.10. Missing Severity Assessment for Adverse Events

If the severity is missing for an AE starting on or after the date of the first dose of study drug, then a severity of "Severe" will be assigned. The imputed values for severity assessment will be used for incidence summaries, and the actual values will be presented in data listings.

5.11. Missing Relationship to Study Drug for Adverse Events

If the relationship to study drug is missing for an AE starting on or after the date of the first dose of study drug, a causality of "Related" will be assigned. The imputed values for relationship to study drug will be used for incidence summaries, and the actual values will be presented in data listings.

5.12. Character Values of Clinical Laboratory Variables

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

5.13. Duplicate Subjects

Duplicate-subject check will be performed throughout study to identify whether there are individuals who are randomized more than once into ACP-103-064 study. Confirmed duplicates will only be included for efficacy analyses under the subject number to which they were first randomized in ACP-103-064. In the event of the confirmed duplicate subjects, additional details for handling the duplicates in the efficacy and safety analyses will be provided in a separate document prior to the final database lock. Data collected under other subject number(s) will be listed but will not be analyzed or summarized. Case narratives will be provided for duplicate subjects.

6. SUBJECT DISPOSITION

For subjects who participate in the screening phase but are not randomized (screen failures), their demographics information (including age, sex, and primary race), screen failure reasons (the specific inclusion/exclusion criterion (or criteria) not met or other reason including the reason due to coronavirus disease 2019 [COVID-19]) and protocol version will be listed. If a subject is re-screened, then the re-screening subject ID and the final enrollment status (whether eventually enrolled) will also be displayed in this listing. In addition, the frequency that the screen failure

reasons are reported will also be summarized. Note that one subject may be deemed ineligible for multiple inclusion/exclusion criteria. If the screen failure was due to a temporary condition that subsequently resolved, it may be allowed to rescreen with the permission of the Medical Monitor.

The number of sites that screened at least 1 subject, number of sites that randomized at least 1 subject, number of subjects screened, and number of unique subjects screened will be summarized by region and overall. In addition, the number of subjects enrolled at each site will also be tabulated by Analysis Set and by treatment group and overall.

For randomized subjects, number and percentage of subjects in Safety Analysis Set, Full Analysis Set, and Per-protocol Analysis Set will be summarized by treatment group and overall. A listing will be provided displaying all subjects excluded from the Safety, Full or Per-protocol Analysis Sets, and will include reason(s) for exclusion. The number and percentage of subjects who are excluded from the Per-protocol Analysis Set will be presented in a summary table by reason, and by treatment group and overall.

For Randomized Analysis Set, Safety Analysis Set, Full Analysis Set, and Per-protocol Analysis Set, the number and percentage of subjects who completed the study or discontinued (all discontinued and by discontinuation reason including the reason due to COVID-19) will also be summarized by treatment group and overall. The number and percentage of subjects completed by visit will be summarized by visit and treatment group.

7. PROTOCOL DEVIATIONS

Protocol deviations will be reviewed regularly over the course of the study. The review process, definition of the deviation categories, and the classification of a deviation as major or minor are detailed in the Protocol Deviation Management Plan. Protocol deviations will also be assessed with respect to relationship to COVID-19.

A summary of the number and percentage of subjects with major protocol deviations for each deviation category will be presented by treatment group and overall in three ways: all protocol deviations, COVID-19 related protocol deviations, and non COVID-19 related protocol deviations.

Three listings of protocol deviations will be provided: all deviations, COVID-19 related protocol deviations and non COVID-19 related protocol deviations.

8. DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Demographics and baseline characteristics will be summarized by treatment group and overall for Randomized Analysis Set, Safety Analysis Set, Full Analysis Set and Per-protocol Analysis Set using descriptive statistics. For the Full Analysis Set, summaries by region will also be presented. Variables include age, sex, primary race (subjects of multi-racial background can only identify/select one primary race on eCRF, or choose "other" and specify), ethnicity, height, weight, body mass index (BMI), region, current smoking status, highest education level, marital status, employment status, Baseline NSA-16 total score, Baseline NSA-16 domain scores, Baseline PANSS total score, Baseline PANSS subscore, Baseline PANSS Marder factor score, Baseline CGI-SCH-S score, Baseline PSP score, Baseline BACS score, Baseline WoRQ total score and Baseline WoRQ readiness to work status.

The reported age reflects a subject's age at the informed consent date. Age, Baseline NSA-16 total score, Baseline PSP score, and Baseline CGI-SCH-S score will be presented as both continuous and categorical variables. Age categories will be presented as 18 to 35 and >35 years old. Baseline NSA-16 total score will be presented as $\leq$55 (lower symptom severity) or >55 (higher symptom severity). Baseline PSP scores will be presented by deciles (e.g. scores of 31 to 40, 41 to 50, etc.). Baseline CGI-SCH-S categories will be presented as scores 1 to 7 as well as 4 (lower disease severity) or $\geq$5 (higher disease severity).

Schizophrenia disease history will be summarized by treatment group for Randomized Analysis Set, Safety Analysis Set, Full Analysis Set and Per-protocol Analysis Set using descriptive statistics. Variables include:

- Age (years) at diagnosis of schizophrenia disease
- Age (years) when received first antipsychotic medication for schizophrenia
- Duration (years) of schizophrenia disease
- Duration of negative symptoms (<1, 1 to 5, or >5 years)
- Time (years) since first antipsychotic treatment
- Number of hospitalizations for treatment of schizophrenia (0, 1 to 5, 6 to 10, or >10)
- Time (years) since last hospitalization
- Current main background antipsychotic medication
- Duration (months) of current main background antipsychotic medication
- Ever had suicidal ideation or behavior (yes or no)

- Had suicidal ideation or behavior within past 6 months (yes or no)

Informed consent date will be used as the reference date for calculating the durations listed above.

Additional information will be listed, including the name of first antipsychotic medication and the dose and frequency of the current main background antipsychotic medication. For subjects who had been hospitalized before, date of last hospitalization and treatment received in the hospital will also be listed.

9. MEDICAL HISTORY

Medical and surgical history data will be coded using Medical Dictionary for Regulatory activities (MedDRA) version 24.0 or newer. The subject incidence will be summarized for each system organ class (SOC) and preferred term by treatment group and overall for Safety Analysis Set. A subject will be counted only once per SOC or per 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

For each subject, the duration of exposure to study drug (last dose date – first dose date + 1), cumulative dose (first multiply the number of doses taken for each kit utilized, then sum the results from all kits), and average daily dose (cumulative dose in mg divided by duration of exposure in days) will be calculated and summarized by treatment group. Duration of 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: <2 weeks (1 to 13 days), 2 to <4 weeks (14 to 27 days), 4 to <8 weeks (28 to 55 days), 8 to <14 weeks (56 to 97 days), 14 to <20 weeks (98 to 139 days), 20 to <26 weeks (140 to 181 days), and ≥26 weeks (182 days or longer). Kaplan-Meier curves of duration on study drug will also be presented for each treatment group.

The pimavanserin dose levels are expressed as free base.

10.2. Measurement of Treatment Compliance

Study drug dosing compliance (in percentage) for a subject is defined as the total number of tablets actually taken divided by the number of tablets expected to be taken and then multiplied by 100. The total number of tablets actually taken is calculated by the total number of tablets dispensed minus the total number of tablets returned. The number of tablets expected to be taken is calculated as the duration of exposure multiplied by 2 (the planned number of tablets taken per day). Additional details for handling missing number of tablets returned will be provided in a separate specification document.

Compliance will be summarized as both continuous and categorical variables by treatment groups. 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. PRIOR, CONCOMITANT AND POST-TREATMENT MEDICATION

For a subject, prior medication is defined as any medication with stop date prior to the date of the first dose of study drug. Concomitant medication is defined as any medication that is ongoing at the first dose of study drug or with a start date between the dates of the first and last doses of study rug, inclusive. Any medication with a start date after the date of the last dose of study drug will be considered as post-treatment medication. Prior, concomitant, or post-treatment medications will be summarized separately. Medications will be coded using WHO Drug Dictionary 2021 March or newer version. The number and percentage of subjects taking each drug class (ATC Level 3) and medication preferred term will be tabulated by treatment group and overall for Safety Analysis Set. Multiple medication usage by a subject in the same category will be counted only once.

Relationship to COVID-19 will be assessed for selected medications as detailed in the GSD. Concomitant and post-treatment medication analyses described above will also be summarized and listed by relationship to COVID-19 (Not related to COVID-19 vs. Related to COVID-19).

12. EFFICACY ANALYSES CONSIDERATIONS

All efficacy analyses will be performed using the planned treatment assignments based on the randomization schedule for the Full Analysis Set

12.1. Efficacy Variables

Primary Efficacy Endpoint

The primary efficacy endpoint is the change from Baseline to Week 26 in the NSA-16 total score.

Key Secondary Efficacy Endpoint

The key secondary efficacy endpoint is the change from Baseline to Week 26 in the CGI-SCH-S of negative symptoms score.

Other Secondary Efficacy Endpoints

Other secondary efficacy endpoints are the following:

- CGI-SCH-I of negative symptoms score at Week 26
- Proportion of CGI-SCH-I of negative symptoms responders (CGI-SCH-I of negative symptoms score of 1 or 2) at Week 26
- Change from Baseline to Week 26 in PSP scale score
- Proportion of NSA-16 responders (≥20% and ≥30% reduction in NSA-16 total score) at Week 26
- Change from Baseline to Week 26 in the PANSS total score
- Change from Baseline to Week 26 in PANSS negative subscores
- Change from Baseline to Week 26 in PANSS Marder factor (negative symptoms) score

Exploratory endpoints are the following:

- Change from Baseline to Week 26 in PANSS subscores (positive scale and general psychopathology scale)
- Change from Baseline to Week 26 in PANSS Marder factor scores (positive symptoms, disorganized thought, uncontrolled hostility/excitement, and anxiety/depression factors)
- Change from Baseline to Week 26 in Work Readiness Questionnaire (WoRQ) total score
- Change from Baseline to Week 26 in Work Readiness Questionnaire (WoRQ) readiness to work question (item 8)
- Change from Baseline to Week 26 in Brief Assessment of Cognition in Schizophrenia (BACS) score
- Change from Baseline to Week 26 in CDSS score

12.2. Adjustment for Covariates

For continuous variables (except CGI-SCH-I) analyzed using the mixed model for repeated measures (MMRM) or analysis of covariance (ANCOVA), the Baseline value of the endpoint being analyzed and geographic region (North America, Europe, or rest of world) will be included as covariates as described in Section 13. For CGI-SCH-I, the Baseline CGI-SCH-S score and geographic region will be included as covariates in the MMRM analysis.

12.3. Handling of Missing Data

The primary analysis of the primary efficacy variable will be performed assuming missing at random (MAR) using an MMRM method. Total scores that are missing, after any imputation of individual missing items as described in Section 5.2.1, will not be imputed. The MMRM method is unbiased in the estimation of treatment effect under the MAR assumption and can be thought of as aiming to estimate the treatment effect that could have been observed if all subjects had continued on treatment for the full study duration (EMA, 2009). Sensitivity analyses of the primary efficacy variable will be performed as described in Section 13.

12.4. Multiple Comparisons / Multiplicity

A hierarchical testing procedure will be used to control the type 1 error rate across the primary and key secondary endpoint. The key secondary endpoint will be tested at the 2-sided significance level of 0.05, if and only if the pimavanserin treatment is superior to the placebo with respect to the primary endpoint at the 2-sided significance level of 0.05. This hierarchical testing procedure of the primary and key secondary endpoints will provide a strong control of the family-wise error rate at the nominal α=5%.

12.5. Examination of Subgroups

Treatment comparisons will be made with respect to the primary and key secondary efficacy variables using the MMRM analysis described in Section 13.1.1 separately for each subgroup by:

- region (North America, Europe, or rest of world)
- age group (18 to 35 or >35 years old)
- sex (male or female)
- primary race (white, non-white)
- main background antipsychotic medication (aripiprazole, asenapine, brexpiprazole, cariprazine, lurasidone, olanzapine, risperidone, or paliperidone)
- duration of schizophrenia (≤5 or >5 years)

- duration of negative symptoms (<1, 1 to 5, or >5 years)
- Baseline smoking status (smoker or non-smoker)
- Baseline BMI (<25, 25 to 30, or >30)
- Baseline NSA-16 total score (≤55 or >55)
- Baseline disease severity measured by CGI-SCH-S score (4 or ≥5)

The LS mean differences with corresponding 95% CIs from the subgroups will also be graphically presented in forest plots for the primary efficacy variable and for the key secondary efficacy variable.

13. METHODS OF EFFICACY ANALYSES

13.1. Primary Efficacy Endpoints

13.1.1. Primary Analysis

Estimand

The primary clinical question of interest for the primary objective is: what is the difference in the mean changes from Baseline to Week 26 in the NSA-16 total score comparing pimavanserin 34 mg vs. placebo as adjunctive treatment, in subjects with negative symptoms of schizophrenia, if all subjects complete 26 weeks of treatment with a stable dose of main background antipsychotic, without extended use of rescue medications or use of prohibited medications?

The primary estimand defining the treatment effect of interest in this study uses the hypothetical strategy specified in the ICH E9 (R1) Addendum. The estimand, or target of estimation, following the hypothetical strategy is the pharmacological effect seen, had no study treatment discontinuation occurred. This hypothetical estimand is justifiable in this case, since the focus is on the pharmacological effect of the drug additional to non-specific effects. Subjects who discontinue from a study treatment either could have lost their treatment effect or could have had their treatment effect confounded by other medication taken after the discontinuation of study treatment. This means that any observations taken after subjects stop study drug will most likely not contribute relevant information about the pharmacological effect of the study drug.

By this strategy, the last collected efficacy assessment after premature treatment discontinuation will be done only at the early termination (ET) visit. Every effort will be made to complete the ET evaluations prior to administering any additional medications for the treatment of negative symptoms in schizophrenia (NSS) or other prohibited medications. In the case of lost to follow-

up events or death, no ET evaluations are expected, and only scheduled assessments performed before such an event are expected.

In this hypothetical strategy, the event of premature discontinuation of study medication is considered missing at random (MAR), and the primary endpoint of the study could be considered as a combination of the responses of on-treatment completers at Week 26 and the imputation of the endpoint to Week 26 following the trend in each treatment group using the mixed-effect model repeated measures (MMRM) method for subjects who discontinuation study drug during the study. All data collected during the study treatment period will be used for statistical analysis. Under the MAR assumption, MMRM provides an unbiased estimate of treatment effect for the treatment period.

The estimand is described by the following attributes:

Target Population: subjects with negative symptoms of schizophrenia (NSS) with a stable dose of main background antipsychotic as defined by the inclusion/exclusion criteria of the study.

Variable (primary endpoint): change from Baseline to Week 26 in the NSA-16 total score

Treatment condition: pimavanserin 34 mg and placebo daily for 26 weeks, with a stable dose of main background antipsychotic, and without extended use of rescue medications or use of or prohibited medications.

Intercurrent events (ICE) and strategies:

- The two intercurrent events “premature treatment discontinuation due to COVID-19”, “premature treatment discontinuation not due to COVID-19” (i.e. early dropout) prior to Week 26 will be addressed by the hypothetical strategy, i.e., assuming that subjects with these ICEs evolve in the same way after ICE occurrence as the subjects in the same treatment group who complete the treatment.
- The three intercurrent events “use of rescue medications beyond maximum allowed period”, “use of prohibited medications” (as listed in the Protocol Appendices A and B) and “change in the main background antipsychotic” resulted in treatment discontinuation (as major protocol violation defined by the protocol), then the event will be considered as an ICE of treatment discontinuation and will be handled in the analyses according to the strategy for treatment discontinuation ICEs.
- The intercurrent event “remote assessments” will be addressed by the treatment policy strategy, i.e., utilizing measurements of the primary efficacy endpoint regardless of the occurrence of this ICE.
- Alternative approaches to handling intercurrent events will be addressed in the sensitivity analyses. Details on the sensitivity analyses will be provided in Section 13.1.2.

Population-level summary: difference (pimavanserin 34 mg vs. placebo) in the mean changes from Baseline to Week 26 in the NSA-16 total score.

Hypotheses

Let Δ be the difference in the mean change from Baseline to Week 26 in the NSA-16 total score between the pimavanserin and placebo groups:

The null hypothesis for the primary efficacy endpoint is: $\Delta = 0$

The alternative hypothesis for the primary efficacy endpoint is: $\Delta \neq 0$

Primary Estimator

The hypothesis testing will be performed for Full Analysis Set using the direct likelihood MMRM method assuming missing data are missing at random (MAR). The dependent variable will be the change from Baseline in the NSA-16 total score. The independent variables in the model will include the following: treatment group (pimavanserin or placebo), visit (Week 2, Week 4, Week 8, Week 14, Week 20, or Week 26), treatment-by-visit interaction, Baseline-by-visit interaction, geographic region (North America, Europe, or rest of world), and the Baseline NSA-16 total score. An unstructured covariance matrix will be used and the Kenward-Roger approximation will be used to adjust the denominator degrees of freedom. Observed cases will be analyzed in the model and the least squares (LS) means will be estimated using observed margins. Sample SAS® code for MMRM analysis is provided in Appendix C.

In the event that the model fails to converge using the unstructured covariance matrix, the following alternative covariance structures will be modeled in the order given: heterogeneous Toeplitz, heterogeneous compound symmetry, heterogeneous autoregressive(1), Toeplitz, compound symmetry, autoregressive(1), variance components. The first alternative covariance structure that allows for convergence will be selected for the final model and the sandwich estimator will be used to estimate the standard errors of the fixed-effect parameters.

At each visit, the effect size (Cohen's d) for the change from Baseline between the treatment groups will be calculated using the following formula:

$$Effect\ size = \frac{LS\ mean\ difference}{\sqrt{variance}}$$

The variance at a given visit will be obtained from the covariance matrix estimated for the MMRM model. The sign (+ or -) of the effect size will be chosen so that a positive value favors pimavanserin.

Summary statistics for the NSA-16 total score (observed and change from baseline) will be presented for all visits from Baseline through Week 26. For change from Baseline values at each

post-Baseline visit, LS means and standard errors (SE), the between-group difference in LS means with the corresponding 95% confidence interval, p-value and effect size will also be presented. In addition, LS mean ± SE over time for the change from Baseline values by treatment group will also be presented in line plots.

The hypothesis testing will be performed based on the difference in LS means at Week 26 and will be tested at an alpha level of 0.05 (2-sided). The treatment-group comparisons at each of the other timepoints (Weeks 2, 4, 8, 14, or 20) using the same MMRM model will be considered exploratory.

The LS mean differences with corresponding 95% CIs from the primary analysis will also be graphically presented in a forest plot.

13.1.2. Sensitivity Analysis with MNAR Using Multiple Imputation (MI)

MMRM assumes data are MAR, which is a reasonable assumption in longitudinal clinical trials. However, the possibility of missing not at random (MNAR) data can never be ruled out. In order to further evaluate robustness of the primary results to deviations from MAR assumptions, additional sensitivity analyses may be conducted. Sensitivity analyses based on pattern-mixture model (PMM) will be performed in order to explore data missing mechanisms of MNAR and investigate the response profile of dropout reasons.

PMM based on MI with mixed missing data mechanisms will be used to investigate the response profile of dropout subjects by dropout reason under MNAR mechanism for the following scenarios:

1) Dropout reason due to Adverse Event (AE) as MNAR
2) Dropout reasons due to AE, lack of efficacy (LOE), subject withdraw consent, COVID-19, use of prohibited medication or change background therapy as MNAR
3) All dropouts not due to COVID-19 as MNAR
4) All dropouts as MNAR

Delta Adjustment Imputation Methods (Tipping Point Analyses)

This sensitivity analysis will be implemented for Full Analysis Set using delta adjustment imputation. This MNAR sensitivity analysis is to departure from MAR assumption by progressively increasing the delta until conclusion from the primary analysis is overturned. The delta is 0%, 5%, 10%, 15%, .., 100% of the observed treatment difference between pimavanserin and placebo from the primary analysis of MMRM model until conclusion of the primary analysis

is overturned. When delta = 0 the missing data are assumed to be MAR. When delta > 0, the missing data are assumed to be MNAR.

1) Using Monte Carlo Markov Chain (MCMC) methodology from PROC MI to impute the intermittent missing data to a monotone missing pattern
2) Using a standard MAR-based multiple imputation approach from PROC MI to impute the monotone missingness data
3) For subjects in the pimavanserin group with a dropout reason due to AE, LOE, subject withdraw consent, COVID-19, use of prohibited medication or change background therapy, a delta will be added for all the values after the dropout time
4) Using MMRM model in the primary analysis to analyze the completed data using PROC MIXED on the multiple imputed data
5) Obtaining the overall results using PROC MIANALYZE.

Placebo Based Imputation Methods (Treatment Policy Strategy)

Similar to "Standard" multiple imputations, except parameters for imputation model obtained from only the placebo (control) group. Missing data for both placebo and pimavanserin group are imputed based on the imputation model derived from placebo data within each stage. If pimavanserin improved outcomes prior to dropout, this benefit is carried into subsequent imputed values, but will diminish over time in accordance with the correlation structure.

13.1.2.1. Remote Assessment

To assess the impact of the intercurrent events of remote assessments on the primary analysis, an indicator variable will be assigned to flag subjects that had assessments done remotely for the primary efficacy endpoint at any visit including baseline (Yes/No). The remote visit flag indicator variable and the interaction between treatment group and remote visit flag indicator variable will be included as factors in the MMRM model.

13.1.2.2. Additional Summaries Evaluating the Impact of COVID-19 on the Primary Endpoint

The following descriptive summaries will be presented to assess the major potential impact of COVID-19 on NSA-16 total score:

- Number and percentage of subjects having each of the following COVID-19-related ICEs and overall
    - Treatment discontinuation due to COVID-19

- Intermediate missing data due to COVID-19
- Remote assessments

- Number and percentage of subjects with missed assessments due to COVID-19 for the analysis of the primary endpoint by visit
- Number and percentage of subjects with remote assessments for the analysis of the primary endpoint by visit

13.1.3. Supportive Analysis

Supportive analysis of the primary endpoint will be analyzed, similarly to the primary analysis as described in Section 13.1.1, using the Per-Protocol Analysis set.

13.2. Key Secondary Efficacy Analysis

The key secondary endpoint, change from baseline to Week 26 in the CGI-SCH-S of negative symptoms score, will be analyzed in the same manner as the primary efficacy endpoint as described in Section 13.1.1, and the same subgroup and supportive analyses will be performed.

13.3. Other Secondary Efficacy Analyses

13.3.1. Responder Analysis

Multiple responder criteria will be used to define a NSA-16 responder (Schooler et al., 2015):

- ≥20%, ≥30%, ≥40%, ≥50%, ≥75% and 100% reduction in total score from Baseline

Responder will also be evaluated based on CGI-SCH-I: score of 1 (very much improved) or 2 (much improved)

Note that since NSA-16 is an interval scale that lacks a natural zero point, the percent change in NSA-16 total score will be calculated based on corrected scores after subtracting 16 points from the raw scores (Leucht et al., 2007; Leucht et al., 2009). For example, if a subject's Baseline NSA-16 total score is 60 and Week 26 NSA-16 total score is 16 (absent of all symptoms), the percent change from Baseline to Week 26 in NSA-16 total score will be calculated as $[(16 - 16) - (60 - 16)] \div (60 - 16) \times 100\% = -100\%$.

For each responder analysis at each timepoint, the proportion of responders will be summarized by treatment group. The treatment groups will be compared using a Cochran-Mantel-Haenszel (CMH) test stratified by geographic region (North America, Europe, or rest of world). The adjusted difference in percent responders between the treatment groups (pimavanserin group

minus placebo group) using the weighting scheme of CMH and Newcombe's 95% CI will also be presented.

At any given visit, the subjects with missing values will be imputed as non-responders for that visit. Using this imputation method, the proportion of NSA-16 responders at Week 26 by treatment group will also be presented in a bar chart.

In addition, an observed-case responder analysis (subjects with missing values at a given visit are excluded from the analysis for that visit) will also be performed.

For change in NSA-16 total score from Baseline, the cumulative distribution function at Week 26 for the observed cases will be plotted by treatment group.

The proportion of subjects who were worsened or no change, based on CGI-SCH-I evaluation, will be summarized by treatment group for each visit. The treatment groups will be compared using a similar CMH test as described above for the responder analyses. Results will be presented using 2 different missing-handling methods: (1) subjects with missing CGI-SCH-I at a given visit are considered as worsened or no change for that visit, and (2) observed-case analysis (subjects with missing values at a given visit are excluded from the analysis for that visit).

13.3.2. Other Secondary Efficacy Analysis for Continuous Variables

The change from Baseline to each post-Baseline timepoint in the NSA-16 global negative symptoms rating, NSA-16 domain scores (communication, emotion/affect, social involvement, motivation, and retardation), PSP score, PANSS total score, PANSS subscores (negative scale), and PANSS Marder factor (negative symptoms) score will be analyzed using an MMRM model. The model will include effects for treatment group, visit, the treatment-by-visit interaction, geographic region, the Baseline value of the endpoint being analyzed, and the Baseline-by-visit interaction.

The CGI-SCH-I of negative symptoms score at each post-Baseline timepoint will be analyzed using an MMRM model. The model will include effects for treatment group, visit, the treatment-by-visit interaction, geographic region, the Baseline CGI-SCH-S of negative symptoms score, and the Baseline-by-visit interaction.

To assess whether there is a linear correlation between the change in NSA-16 total score from Baseline and CGI-SCH-I rating, the change in NSA-16 total score from Baseline will also be summarized by CGI-SCH-I rating and treatment group for all post-baseline visits. The mean $\pm$ SE change in NSA-16 total score from Baseline by CGI-SCH-I rating will also be plotted for Week 26 observed cases.

13.4. Exploratory Efficacy Analyses

The change from Baseline to each post-Baseline timepoint in the PANSS subscores (positive scale and general psychopathology scale), PANSS Marder factor (positive symptoms, disorganized thought, uncontrolled hostility/excitement, and anxiety/depression factors) score, WoRQ total score, and CDSS score will be analyzed using an MMRM model. The model will include effects for treatment group, visit, the treatment-by-visit interaction, geographic region, the Baseline value of the endpoint being analyzed, and the Baseline-by-visit interaction.

The proportion of subjects who are rated as ready for work in the WoRQ readiness to work question (item 8) and the shift from the Baseline status of the readiness to work will be summarized at each timepoint. In addition, the odds of being ready for work at Week 26 will be estimated using logistic regression adjusted for Baseline status.

The change from Baseline to Week 26 in the normalized BACS composite and domain scores will be analyzed using an analysis of covariance (ANCOVA) model with effects for treatment group (pimavanserin or placebo), geographic region, and the Baseline BACS score.

At Week 26, the effect size (Cohen's d) for the change from Baseline between the treatment groups will be calculated using the following formula:

$$Effect\ size = \frac{LS\ mean\ difference}{\sqrt{MSE}}$$

Where MSE is the mean squared error from the ANCOVA model. The sign (+ or -) of the effect size will be chosen so that a positive value favors pimavanserin.

14. SAFETY ANALYSES

The safety analysis will be performed based on the Safety Analysis Set using actual treatment. Safety variables include adverse events (AEs), clinical laboratory variables, vital signs, body weight, BMI, physical examinations, electrocardiogram (ECG), C-SSRS, AIMS, BARS and SAS variables.

14.1. Adverse Events

All Adverse events (AEs) will be coded using MedDRA version 24.0 or newer.

An AE (classified by preferred term) will be considered a treatment-emergent AE (TEAE) if it started after first study dose administration and no later than last study dose date + 30. AEs reported on Day 1 based on Baseline (pre-dose) findings (e.g. clinically significantly abnormal

vital signs, laboratory test results, or electrocardiogram parameters) will not be considered as TEAEs.

The event counts and the number and percentage of subjects reporting TEAEs in each treatment group 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, the subject will be counted only once for that preferred term using the most severe and more related occurrence for the summarization by severity and by relationship to study drug, respectively. The display in these tables will be sorted alphabetically by SOC and then by descending subject frequency for the preferred terms 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 within each treatment group. This table will be sorted by descending subject frequency.

The relationship of selected AEs to COVID-19 will be assessed as detailed in the GSD. TEAEs tabulated by SOC and preferred term will be presented with and without COVID-19 related TEAEs.

The incidence of most frequently reported (preferred terms reported by ≥ 5% of subjects in either treatment group) TEAEs, treatment emergent SAEs, and TEAEs leading to discontinuation of study drug, TEAEs related to study drug will be summarized by SOC, preferred term, and treatment group. The tables will be sorted alphabetically by SOC and then by descending frequency within each SOC. In addition, the incidence of fatal treatment-emergent AEs (i.e., events that cause death) will be summarized separately by SOC, preferred term and treatment group. These summary tables except for the most frequently reported TEAEs table will also be presented with and without COVID-19 related events.

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 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 treatment-emergent SAEs, related TEAEs, TEAEs leading to discontinuation, fatal TEAEs (if any), and all COVID-19 related events.

14.2. Clinical Laboratory Variables

Clinical laboratory tests are performed at Screening, Baseline and Week 26/ET visits. Due to various circumstances (e.g., pandemic, natural disaster, or political upheaval), it is possible that

some test results may be collected from a local laboratory. Local laboratory results and the associated normal ranges will be converted to SI units; the local laboratory results, in SI units, will then be normalized to central lab ranges to be included in summary data analysis together with the central laboratory results. The normalization will be performed using the following scale transformation equation:

$$s = L_s + (x - L_x)\frac{U_s - L_s}{U_x - L_x}$$

where s is the normalized individual laboratory value to be used for summary; x is the original value from the local lab; L_x and U_x are the lower and upper limits from the local lab; L_s and U_s are the lower and upper limits from the central lab.

For labs with only a single upper (or lower) limit, the following scale transformation equation will be used:

$$s = x\frac{U_s}{U_x}$$

where s is the normalized individual laboratory value to be used for summary; x is the original value from the local lab; U_s is the upper (or lower) limit from the central lab; U_x is the upper (or lower) limit from the local laboratory. Local laboratory results and normalized results will be included in data listings. Only central lab and normalized local lab results will be used for summary of change from baseline, shift, and potentially clinically important (PCI) analyses.

- Clinical chemistry serum tests
    - Sodium (Na), potassium (K), chloride (Cl), phosphorus (P), calcium (Ca), magnesium (Mg)(at Screening only), carbon dioxide (CO_2), blood urea nitrogen (BUN), creatinine (CR), uric acid
    - Alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), total bilirubin (TBIL), lactate dehydrogenase (LDH)
    - Vitamin B12 (at Screening only),
    - HbA_{1c} (at Screening only),
    - Glucose
    - Albumin (ALB), total protein
    - Prolactin
    - Thyroid stimulating hormone (TSH) and free T4
    - Creatine kinase (CK)/creatine phosphokinase (CPK)
    - Lipid panel

        - Total cholesterol, HDL-cholesterol, triglycerides, LDL-cholesterol, cholesterol/HDL ratio, non-HDL cholesterol
- Pregnancy test
    - A serum pregnancy test should only be performed at Visit 1 (Screening) for women of childbearing potential
    - A urine pregnancy test should be performed at all designated visits after Visit 1 (Screening) for women of child-bearing potential
        - If urine cannot be obtained in women of childbearing potential, a serum pregnancy test should be done in its place
- Hematology tests
    - Complete blood count (CBC) including:
        - White blood cell (WBC) count
        - Complete differential (relative and absolute)
        - Hematocrit (Hct), hemoglobin, red blood cells (RBC), platelets
        - Reticulocyte
- Urinalysis
    - Blood, leukocyte esterase (plus reflexed microscopic analysis if test is positive), protein, glucose, ketones, specific gravity, pH
- Urine toxicity (drug) screen
    - The urine toxicity (drug) screen (UDS) will include testing for the following substances: tetrahydrocannabinol (THC), benzodiazepines, barbiturates, cocaine, amphetamine, methamphetamine, ecstasy, opiates, methadone, oxycodone, buprenorphine, and phencyclidine.

All laboratory test results 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 for hematology, chemistry and urinalysis will be summarized by treatment group using descriptive statistics at Baseline and Week 26 visits. The change from Baseline values will also be summarized by treatment group at the Week 26 visit. 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 (blood, protein, glucose, and ketones), the number and percentage of subjects will be tabulated by category at Baseline and Week 26, and the denominator is the number of subjects with non-missing values for the given parameter, visit and treatment group.

Laboratory values (except the ones that were only assessed at screening and urine drug screen)

will also be summarized in shift tables by treatment group, to determine the number and percentage of subjects with values classified as below (low), within (normal), and above (high) normal ranges at Week 26, overall post-Baseline minimum and overall post-Baseline maximum, relative to the same classification at the Baseline visit. 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 treatment group. For the 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 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 laboratory values (PCI) at Week 26 and overall post-baseline will be summarized by treatment group for selected parameters. PCI criteria are listed in Table 2 and Table 3. For the overall post-Baseline summaries of PCI values, 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 value for the given parameter, visit and treatment group, and the denominator is the number of subjects with non-missing values for the given parameter, visit and 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 treatment group, and the denominator is the number of subjects with at least 1 post-Baseline laboratory value for the given parameter and 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. Subjects with any PCI values will be presented in an additional listing.

Table 2 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 3 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 and the urine drug screen results will be presented in a listing.

14.3. Vital Signs

Vital signs are assessed at Screening, Baseline, Weeks 2, 4, 8, 14, 20, and Week 26/ET visits.

Vital signs including weight, height (only at screening), and the derived BMI will be summarized by treatment group using descriptive statistics at Baseline and all post-Baseline visits. The change from Baseline values will also be summarized by treatment group at the 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 4. The number and percentage of subjects with post-Baseline vital signs that are PCI will be summarized by treatment group at each 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 treatment group, and the denominator is the number of subjects with non-missing values for the given parameter, visit and 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 the given parameter and treatment group, and the denominator is the number of subjects with at least 1 post-Baseline vital sign for the given parameter and treatment group. A listing of all subjects with any PCI values will be provided.

Table 4 Criteria for Potentially Clinically Important Vital Signs

Vital Sign Parameter	Unit	Criteria[a]		
		Observed Value	And/Or	Change Relative to Baseline
Systolic blood pressure (supine or sitting)	mmHg	≥180	And	Increase of ≥20
		≤90	And	Decrease of ≥20
Diastolic blood pressure (supine or sitting)	mmHg	≥105	And	Increase of ≥15
		≤50	And	Decrease of ≥15
Pulse (supine or sitting)	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)

Electrocardiogram is performed at Screening, Baseline, Week 14 and Week 26/ET visits.

All tracings will be evaluated by a central reading laboratory. At the Baseline visit, the machine-read results will also be recorded. All data, including the machine-read Baseline results, will be listed. 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 by treatment group. 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 by treatment group at each visit and for the 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

For cardiologist's interpretations, the number and percentage of subjects with ECG results that are determined as normal or abnormal will be summarized at scheduled visits. The overall post-baseline worst interpretation will also be summarized (i.e. if a subject has one or more post-baseline ECG results that is/are considered as abnormal, this subject will be counted in the abnormal category). Cardiologist's interpretations will also be summarized in a shift table to determine the number and percentage of subjects with ECG results classified as normal or abnormal at scheduled post-Baseline visits relative to the same classification at the Baseline visit. The shifts from Baseline to overall post-Baseline worst interpretation will also be presented. For the by-visit shift summary, the denominator is the number of subjects with non-missing cardiologist's interpretation at Baseline and the given visit for the given treatment group. For the summaries of shift from Baseline to the overall post-Baseline worst interpretation, the denominator is the number of subjects with non-missing Baseline and at least 1 post-Baseline cardiologist's interpretation for the given treatment group.

Electrocardiogram variable values will be considered PCI if they meet or exceed the upper limit values listed in Table 5. The number and percentage of subjects with post-baseline PCI values will be summarized by treatment group at each post-Baseline visit and for overall post-Baseline. 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 treatment group, and the denominator is the number of subjects with non-missing values for the given parameter, visit and 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 treatment group, and the denominator is the number of subjects with at least 1 post-Baseline ECG value for the given parameter and treatment group. A listing of all subjects with any PCI values will be provided.

Table 5 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 Screening, Baseline, Weeks 2, 14 and Week 26/ET visits.

Physical examination results (normal, abnormal, and not done) at Baseline and all post-Baseline visits will be summarized in a frequency table by treatment group, body system and visit.

A listing of physical examination data will be listed.

14.6. Other Safety Variables

14.6.1. Suicidality

The C-SSRS is assessed at Screening, Baseline, Weeks 2, 4, 8, 14, 20, and Week 26/ET visits.

The event counts and the number and percentage of subjects reporting any post-Baseline suicidal ideation (wish to be dead; non-specific active suicidal thoughts; active suicidal ideation with any methods (not plan) without intent to act; active suicidal ideation with some intent to act, without specific plan; active suicidal ideation with specific plan and intent), suicidal behavior (preparatory acts or behavior; aborted attempt; interrupted attempt; actual attempt; suicide), or suicidality (any suicidal ideation or behavior) will be tabulated for each treatment group.

14.6.2. Extrapyramidal Symptom Measures

14.6.2.1. Abnormal Involuntary Movement Scale (AIMS)

The AIMS is assessed at Screening, Baseline, Weeks 2, 14, and Week 26/ET visits.

The AIMS total score will be summarized by treatment group using descriptive statistics at Baseline, Weeks 2, 14 and 26 visits. The change from Baseline scores will also be summarized by treatment group at Weeks 2, 14 and 26 visits.

In addition, the number and percentage of subjects with dyskinesia will be summarized by treatment group at each visit and for overall post-Baseline. Dyskinesia is defined as having a score of 3 or more on any of the first 7 AIMS items or a score of 2 or more on any two of the first 7 AIMS items. If there are multiple assessments performed within the same visit window, all assessments regardless of whether the value is selected for the by-visit summaries will be evaluated for dyskinesia using the above criteria. The tabulations will be presented by subjects' status at Baseline:

- Subjects without Dyskinesia at Baseline.
- All subjects who have any non-missing AIMS assessment regardless the status of Dyskinesia at Baseline

The individual item scores will be listed but not summarized.

14.6.2.2. Barnes Akathisia Scale (BARS)

The BARS is assessed at Screening, Baseline, Weeks 2, 14, and Week 26/ET visits.

The BARS total score and the Global Clinical Assessment of Akathisia score will be summarized by treatment group using descriptive statistics at Baseline, Weeks 2, 14 and 26 visits. The change from Baseline scores will also be summarized by treatment group at Weeks 2, 14 and 26 visits.

In addition, the number and percentage of subjects with akathisia will be summarized by treatment group at each visit and for overall post-Baseline. Akathisia is defined as having a Global Clinical Assessment of Akathisia score ≥2. If there are multiple assessments performed within the same visit window, all assessments regardless of whether the value is selected for the by-visit summaries will be evaluated for Akathisia using the above criteria. The tabulations will be presented by subjects' status at Baseline:

- Subjects without Akathisia at Baseline.
- All subjects who have any non-missing BARS assessment regardless the status of Akathisia at Baseline

The individual item scores will be listed but not summarized.

14.6.2.3. Simpson Angus Scale (SAS)

The SAS is assessed at Screening, Baseline, Weeks 2, 14, and Week 26/ET visits.

The SAS total score will be summarized by treatment group using descriptive statistics at Baseline, Weeks 2, 14 and 26 visits. The change from Baseline scores will also be summarized by treatment group at Weeks 2, 14 and 26 visits.

The number and percentage of subjects with Parkinsonism will be summarized by treatment group at each visit and for overall post-Baseline. Parkinsonism is defined as having a SAS total score >3. If there are multiple assessments performed within the same visit window, all assessments regardless of whether the value is selected for the by-visit summaries will be evaluated for Parkinsonism using the above criteria. The tabulations will be presented by subjects' status at Baseline:

- Subjects without Parkinsonism at Baseline.
- All subjects who have any non-missing SAS assessment regardless the status of Parkinsonism at Baseline

for subjects who have at least 1 SAS assessment as well as for a subset of these subjects who do not have Parkinsonism at Baseline.

The individual item scores will be listed but not summarized.

15. CLINICAL PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

PK samples for pimavanserin, its metabolite (AC-279), and the main background antipsychotic medications are collected pre-dose at Baseline, Weeks 2, 8, 14, and Week 26/ET visits.

For pimavanserin-treated subjects, plasma concentration for pimavanserin and AC-279 will be listed. For pimavanserin-treated subjects in the Pharmacokinetics Analysis Set, plasma concentration for pimavanserin and AC-279 will be summarized using descriptive statistics by visit. Concentrations that are below the limit of quantification (BLQ) will be displayed as "BLQ" in the data listings and imputed as 0 for computing summary statistics.

Plasma concentration data for the background main antipsychotics will be listed.

If data allow, population PK and PK/PD analyses will be performed to further characterize the PK profile and exposure response relationship of pimavanserin and its metabolite using measures of safety, and efficacy parameters. The results of population PK and PK/PD modeling will be presented in a separate report. Pimavanserin and AC-279 plasma concentration data will remain blinded to the Investigators and the Sponsor until the unblinding of the clinical database at the end of the study.

16. UNBLINDED INTERIM ANALYSIS

No interim analysis for efficacy is planned for this study.

17. DATA MONITORING/REVIEW COMMITTEE

An independent Data and Safety Monitoring Board (DSMB) will review interim safety data including data on AEs and SAEs. The DSMB will be independent of the Sponsor and the Investigators and will be empowered to recommend stopping the study due to safety concerns. The roles and responsibilities of DSMB members and planned frequency of meetings are detailed in the DSMB Charter.

An independent statistician (and/or programmer) not affiliated with the Sponsor will produce unblinded statistical outputs and provide these outputs to DSMB members using a secure method. The Sponsor and the Investigators will remain blinded until the official unblinding of the database at the end of the study. The outputs presented to DSMB members will include but are not limited to summaries of enrollment and disposition, demographics and baseline characteristics, medical and schizophrenia histories, concomitant medications, study drug exposure, all treatment-emergent adverse events (including deaths, SAE and AEs leading to discontinuation), vital signs, laboratory test results, and ECG parameters. Subject profiles,

boxplots of Baseline and most extreme post-Baseline values for clinical laboratory data, and listings of AEs and potentially clinically important laboratory and QTcF results will also be provided to DSMB members for their review.

18. COMPUTER METHODS

Statistical analyses will be performed using Version 9.4 (or newer) of SAS® on a suitably qualified and validated environment.

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

19. CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

No changes are made to the analyses specified in the protocol.

20. REFERENCES

Leucht S, et al. Defining 'response' in antipsychotic drug trials: recommendations for the use of scale-derived cutoffs. *Neuropsychopharmacology.* 2007; 32: 1903-1910.

Leucht S, et al. Definitions of response and remission in schizophrenia: recommendations for their use and their presentation. *Acta Psychiatr Scand.* 2009; 119 (Suppl. 438): 7-14.

Schooler N, et al. Defining therapeutic benefit for people with schizophrenia: focus on negative symptoms. *Schizophr Res.* 2015; 162(1-3): 169-174.

EMA (2009). *Guideline on Missing Data in Confirmatory Clinical Trials, European Medicines Agency*, London, UK.

FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic | Guidance for Industry, Investigators, and Institutional Review Boards (March 2020).

21. APPENDICES

21.1. Appendix A Schedule of Events and Assessments

Period	Screening Period	Treatment Period							Follow-Up[a]
Visit[b]	1	2 (Baseline)	3	4	5	6	7	8 (EOT/ET)	Safety Follow-up
Day or Week	(Up to 42 days)	Day 1/Week 0	Week 2	Week 4	Week 8	Week 14	Week 20	Week 26	Week 30
Allowable visit window (# days)			±3	±3	±3	±7	±7	±7	+4
Informed consent[c]	X								
Inclusion/exclusion criteria	X	X							
Demography	X								
Medical and psychiatric history[d]	X								
Schizophrenia disease history	X								
SCID-5-CT customized module	X								
Interview by an independent clinician[e]	X								
Physical examination[f]	X	X	X			X		X	
Vital signs	X	X	X	X	X	X	X	X	
Height and weight[g]	X	X						X	
12-lead ECG[h]	X	X				X		X	
Clinical laboratory tests[i]	X	X						X	
Confirmation of main antipsychotic[j]	X								
Pregnancy test[k]	X	X	X	X	X	X	X	X	
Urine toxicity (drug) screen	X	X						X	
PK blood draws[l]		X	X		X	X		X	
NSA-16		X	X	X	X	X	X	X	
PANSS and IQ-PANSS	X	X				X		X	
CGI-SCH-S of negative symptoms	X	X	X	X	X	X	X	X	
CGI-SCH-I of negative symptoms			X	X	X	X	X	X	
PSP		X	X		X	X		X	
BACS		X						X	
CDSS	X	X				X		X	
WoRQ		X			X	X		X	
C-SSRS[m]	X	X	X	X	X	X	X	X	
AIMS, BARS, and SAS	X	X	X			X		X	
Assessment of prior and concomitant medications[n]	X	X	X	X	X	X	X	X	X
Assessment of adverse events	X	X	X	X	X	X	X	X	X
Dispense study drug		X	X[o]	X[o]	X[o]	X[o]	X[o]		
Study drug accountability			X	X	X	X	X	X	

Abbreviations: AIMS=Abnormal Involuntary Movement Scale; BACS=Brief Assessment of Cognition for Schizophrenia; BARS=Barnes Akathisia Rating Scale; CDSS=Calgary Depression Scale for Schizophrenia; CGI-SCH-I=Clinical Global Impression of Schizophrenia–Improvement; CGI-SCH-S=Clinical Global Impression of Schizophrenia–Severity; C-SSRS=Columbia–Suicide Severity Rating Scale; ECG=electrocardiogram; EOT=end of treatment; ET=early termination; HCV=hepatitis C virus; HIV=human immunodeficiency virus; IQ-PANSS=Informant Questionnaire for the Positive and Negative Syndrome Scale; NSA-16=Negative Symptom Assessment–16 scale; PANSS=Positive and Negative Syndrome Scale; PK=pharmacokinetic(s); PSP=Personal and Social Performance; SAS=Simpson-Angus Extrapyramidal Side Effects Scale; SCID-5-CT=Structured Clinical Interview for DSM-5, Clinical Trials Version; WoRQ=Work Readiness Questionnaire

[a] For subjects who discontinue prematurely from the study or who do not enroll in the 52-week, open-label extension study (ACP-103-035), a safety follow-up telephone call visit will occur 30 (+4) days after the last dose of study drug.

[b] Study visits are designated by weeks and have a window, calculated from the Baseline visit, of ±3 days for Visits 3, 4, and 5 and of ±7 days for Visits 6, 7, and 8. The Screening Period is up to 42 days long (Day -42 to Day -1) and all Screening procedures should be completed as early in the Screening Period as possible. The window for the 30-day follow-up telephone call visit is +4 days. Circumstances may arise (e.g., pandemic, natural disaster, or political upheaval) when on-site assessments of efficacy and/or safety are not possible. In those cases, assessments may be performed at the subject's place of residence by raters either in person, or via video technology or telephone where possible. The Investigator **must** contact the Medical Monitor for approval with the plan. Sites must keep a log to identify details of all visits that are administered remotely. For some remote efficacy assessments (i.e., NSA-16, PANSS, CGI), the vendor will provide additional training to ensure calibration to reduce discrepancy between on-site and remote assessments.

[c] The subject's caregiver must provide written agreement prior to any Screening procedures being performed indicating their agreement to participate in the study in the caregiver role.

[d] Medical history is to include a history of tobacco and nicotine use. A review of any history of HIV, hepatitis B, or HCV will also be performed.

[e] A structured telemedicine interview of the subject by an independent clinician will be performed during the Screening Period. The interview will be conducted by video and will not be recorded.

[f] A complete physical examination should be performed at Screening, Baseline, and Week 26; symptom-directed examinations will be performed at Weeks 2 and 14. Pelvic and/or urogenital examination may be deferred, unless the Investigator deems this to be clinically indicated.

[g] Height will only be measured at the Screening visit.

[h] A 12-lead ECG will be performed in triplicate at Screening. Single 12-lead ECG recordings will be performed at Baseline, Week 14, and at the Week 26 (EOT/ET) visit. A single 12-lead ECG can be performed any time before blood sampling or at least 30 minutes after blood sampling during clinic visits. Circumstances may arise (e.g., pandemic, natural disaster, or political upheaval) when on-site ECG assessment is not possible. In those cases, ECG assessments may be performed at the subject's place of residence by study staff. The Investigator **must** contact the Medical Monitor for approval with the plan. Sites must keep a log to identify details of all visits that are administered remotely.

[i] To include hematology, serum chemistry, prolactin levels, and urinalysis (note: additional laboratory studies [in addition to scheduled timepoints shown in the table] for a given subject may be repeated at any time throughout the Treatment Period, at the discretion of the Investigator). It is preferable but not required that subjects be in a fasted state (e.g., fasting for approximately 10 hours) before the blood sample for clinical chemistry is obtained. Circumstances may arise (e.g., pandemic, natural disaster, or political upheaval) when on-site clinical laboratory tests are not possible. In those cases, clinical laboratory tests may be performed at the subject's place of residence by study staff or at a local laboratory. The Investigator **must** contact the Medical Monitor for approval with the plan. Sites must keep a log to identify details of all visits that are administered remotely.

[j] Blood samples for measurement of the following will be obtained at Screening only: glycosylated hemoglobin (HbA_{1c}), thyroid stimulating hormone (TSH), and main antipsychotic detection. Measurement of a full thyroid panel will be conducted only if the TSH value is outside of the laboratory reference range.

[k] A serum pregnancy test will be completed at the Screening visit for all female subjects of childbearing potential; urine pregnancy tests will be completed at all other scheduled time-points for all female subjects of childbearing potential.

[l] At the Screening visit, a PK sample will be collected for the presence or absence of the subject's main antipsychotic. At each subsequent timepoint, a PK sample will be collected for pimavanserin, the metabolite AC-279, and the main antipsychotic. The Baseline PK sample should be collected pre-dose. When possible, an additional PK sample will be collected from subjects who experience an SAE or an AE leading to discontinuation, as soon as possible after the occurrence of that event. For all PK samples (scheduled and unscheduled), the dates and times of administration of the last 3 doses of both study drug and the main antipsychotic should be recorded. For samples collected from subjects who experience an SAE or an AE leading to discontinuation, the date and time of the last dose prior to the SAE or AE should also be recorded. Circumstances may arise (e.g., pandemic, natural disaster, or political upheaval) when on-site assessment of PK is not possible. In those cases, PK assessments may be performed at the subject's place of residence by study staff. The Investigator **must** contact the Medical Monitor for approval with the plan. Sites must keep a log to identify details of all visits that are administered remotely.

[m] The Baseline/Screening version of the C-SSRS will be administered at Screening, and the Since Last Visit version of the C-SSRS will be administered at all subsequent visits.

[n] Prior medication history will only be collected at the Screening visit.

[o] Subjects are to return unused study drug and all kit materials at each subsequent visit; a new kit will be dispensed at each identified visit. In addition to the study drug dispensed at the site, investigational product may be delivered directly to the subject's place of residence; related procedures will be described in the study-specific pharmacy manual.

21.2. Appendix B Composite z score and T-score Transformation for Brief Assessment of Cognition in Schizophrenia (BACS)

PSYCHOMETRIC AND ANALYSIS METHODOLOGY

ACP-103-064

Brief Assessment of Cognition in Schizophrenia (BACS)

Project Lead:

PPD

Table of Contents

Description of the Brief Assessment of Cognition in Schizophrenia (BACS)

The Brief Assessment of Cognition in Schizophrenia (BACS) was used to assess cognition. The validity and reliability properties of the BACS have been established in patients with schizophrenia and healthy controls, and the BACS composite score has proven high test-retest reliability, increasing the likelihood of detecting a treatment-related effect (Keefe et al., 2004, 2006, 2008).

BACS Component Test Scores

The BACS consists of 6 subscales:

The Brief Assessment of Cognition in Schizophrenia (BACS) battery was administered to subjects. The BACS includes the following tests with the primary domain(s) of cognition they measure in parentheses:

1. Verbal memory task (verbal memory) – a subject was given 5 attempts to remember 15 words and recall as many words as possible;

2. Digit sequencing task (working memory) – a subject was presented with clusters of numbers of increasing length and then asked to repeat in order from the lowest to highest length;

3. Token motor task (motor speed) – a subject was given 100 plastic tokens and asked to place 2 tokens at a time within a container as quickly as possible within 60 seconds;

4. Verbal fluency (language, speed of processing) – divided into semantic and letter fluency, whereby a subject was asked to name as many words as possible within a specific category (e.g., supermarket items), and to name words that begin with a specific letter (e.g., F and S) within 60 seconds, respectively;

5. Symbol coding task (attention, speed of processing) – a subject was asked to write matching numbers from 1 to 9 to symbols within 90 seconds;

6. Tower of London (executive functions) –a subject was shown two pictures of three balls of different colors arranged on three different pegs, whereby the balls were arranged differently on each picture and the subjects were asked to give the total number of times the balls in one picture needed to be moved in order to end with the arrangement in the other picture.

Scaled Test Component Scores and the BACS Composite

For each subscale, higher scores reflect better cognition. For each subscale, a Standard Deviation Score was calculated based on normative data (Keefe et al. 2008). The BACS composite score is calculated as the average Standard Deviation Score of the 6 subscale scores, standardized so that the mean of the composite score in the healthy normative sample is 50 and the standard deviation is 10. The change in

BACS composite score was calculated as the BACS composite score at Week 26 minus the BACS composite score at baseline. BACS assessments were completed in the following languages: Bulgarian, Croatian, Czech, Hungarian, Italian, Lithuanian, Polish, Russian, Serbian, Spanish, Ukrainian and Romanian. Individual test scores were converted into standardized (T) scores and composite scores that were corrected for age and gender (Keefe et al., 2008).

STATISTICAL ANALYSES

BACS z score and T Score computation

i= each of the 6 subscales of the BACS

z{i}=(raw score for subscale{i}-mean score from the norms{i})/standard deviation of the norms {i}
z{i}=round(z{i},.01)
t{i}=round((z{i} x 10)+50)

Composite z score = (Verbal Memory z + Digit Sequencing z + Token Motor z + Verbal Fluency z + Symbol Coding z + Tower of London z)/denominator

Composite T Score = round ((composite z x 10)+50)

BACS Composite Score

The BACS composite score was calculated by averaging all z scores from the BACS subscales, including Verbal Memory, Digit Sequencing, Token Motor, Symbol Coding, Verbal Fluency (Semantic Fluency and Letter Fluency) and Tower of London. The z scores were produced by available age and gender corrected norms. The composite score is a measure of overall cognitive functioning. Raw score means (Standard Deviation) and z scores (corrected for gender and age) across visits were presented. The BACS composite scores are calculated by averaging all z-scores of the six subtests from the BACS.

Data Ranges for each BACS Subtest

Subtest Ranges	Minimum Value	Maximum Value
Verbal Memory	0	75
Digit Sequencing	0	28
Token Motor	0	100
Verbal Fluency	0	300
Symbol Coding	0	110
Tower of London	0	22

Handling Missing Scores on the BACS Composite

The Composite score for subjects with missing subtests are imputed using the average z or T score of the remaining subtests. For ≤ 2 missing BACS subtests, the average of the other non-missing subtests z-scores are imputed for the composite score. If there are more than 2 missing BACS subtests, the

composite score should be set to missing. If a subtest score is missing, the above is only used to impute the composite score, but the subtest score should be set to missing.

References

Keefe RS, Goldberg TE, Harvey PD, Gold JM, Poe MP, Coughenour L. The brief assessment of cognition in schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery.Schizophr Res. 2004;68:283–97.

Keefe RS, Harvey PD, Goldberg TE, Gold JM, Walker TM, Kennel C, et al. Norms and standardization of the Brief Assessment of Cognition in Schizophrenia (BACS) Schizophr Res. 2008;102:108–15.

Keefe RS, Poe M, Walker TM, Kang JW, Harvey PD. The schizophrenia cognition rating scale: an interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity.Am J Psychiatry. 2006;163:426–32. doi: 10.1176/appi.ajp.163.3.426.

Keefe, R.S.E., Poe, M., Walker, T.M., Harvey, P.D., The relationship of the Brief Assessment of Cognition in Schizophrenia (BACS) to functional capacity and real-world functional outcome. J. Clin. Exp. Neuropsychol. 2006b. 28, 260-269.

21.3. Appendix C Sample SAS Codes

CCI

[redacted]


CCI

21.4. Summary of Version Changes

Version No:	Document History Description of Update	Author(s)	Version Date
1.0	Original version	PPD	03 DEC 2020
2.0	Updated to include analyses related to COVID-19	PPD	12 APR 2021
3.0	Updated to include the additional language for estimand and use the sandwich estimator for the alternative covariance structures	PPD	24 AUG 2021
4.0	Updated for changes in protocol amendment 3; Revised NSA-16 imputation rules to the ACP-103-038 method; Corrected the WoRQ score direction; Removed the upper limit of Week 26 analysis visit window for safety and PK endpoints; Redefined analysis visit windows for efficacy endpoints and to exclude data collected more than 7 days after last dose date; Clarified the languages about estimand; Added additional sensitivity analyses; Added 40% responder cut-off for NSA responder analyses.	PPD	11 JAN 2024