Official Title: A 52-Week, Open-Label, Extension Study of Pimavanserin for the Adjunctive Treatment of Schizophrenia

NCT Number: NCT03121586

Document Date: 26 January 2024

STATISTICAL ANALYSIS PLAN

Protocol No.:	ACP-103-035
Protocol Title:	A 52-Week, Open-Label, Extension Study of Pimavanserin for the Adjunctive Treatment of Schizophrenia
Drug:	Pimavanserin
Sponsor:	Acadia Pharmaceuticals Inc.
Version No. and Date	Version 1.0, January 26, 2024

SIGNATURE/APPROVAL PAGE

AUTHOR

See appended electronic signature page

PPD
Director, Biostatistics, Biometrics
Acadia Pharmaceuticals Inc.

Date

APPROVERS

See appended electronic signature page

PPD
PPD M.D.
Clinical Development
Acadia Pharmaceuticals Inc.

Date

See appended electronic signature page

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

Date

Signature Page for ACP-103-035 Statistical Analysis Plan V1.0

Approve	PPD Biostatistics 27-Jan-2024 00:00:29 GMT+0000

Approve	PPD Clinical 27-Jan-2024 07:52:46 GMT+0000

Approve	PPD Biostatistics 27-Jan-2024 19:20:59 GMT+0000

Signature Page for ACP-103-035 Statistical Analysis Plan V1.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AE	adverse event
AIMS	Abnormal Involuntary Movement Scale
ATC	Anatomical Therapeutic Chemical
BARS	Barnes Akathisia Rating Scale
BMI	body mass index
CGI-S	Clinical Global Impression scale – Severity
CGI-SCH-S	Clinical Global Impression of Schizophrenia scale – Severity
COVID-19	coronavirus disease 2019
C-SSRS	Columbia-Suicide Severity Rating Scale
DAI-10	10-item Drug Attitude Inventory
DB	Double-blind
DSMB	Data and Safety Monitoring Board
ECG	electrocardiogram
eCRF	electronic case report form
EOS	End-of-Study
ET	Early Termination
GSD	Guidance for Site Documentation and Data Management Querying of Data Impacted by COVID-19
MedDRA	Medical Dictionary for Regulatory Activities
msec	milliseconds
NSA-16	Negative Symptom Assessment-16
OL	Open-label
PANSS	Positive and Negative Syndrome Scale
PCI	potentially clinically important
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
SF-36	36-item Short Form Health Survey
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 safety and efficacy data as described in the study protocol Amendment 3 dated 11 August 2020. Specifications for tables, figures, and listings are contained in a separate document.

For Argentina, a country-specific protocol amendment (Amendment 3-AR finalized 27 April 2021) specifies additional pregnancy tests at Week 10, 16, 24, 32, 40 and 48 for all female subjects of childbearing potential to be conducted only for subjects enrolled in Argentina.

2 OBJECTIVES

2.1 Primary Objective

- To evaluate the long-term safety and tolerability of pimavanserin after 52 weeks of adjunctive treatment in subjects with schizophrenia

2.2 Secondary Objective

- Clinical global assessment overall severity of symptoms

2.3 Exploratory Objectives

To evaluate the continued efficacy of pimavanserin treatment with respect to:

- Overall symptoms of schizophrenia
- Negative symptoms of schizophrenia (for subjects from Study ACP-103-038 and -064 only)
- Personal and social performance

3 STUDY DESIGN

3.1 General Study Design

ACP-103-035 is an open-label extension study to determine the long-term safety and tolerability of pimavanserin for the adjunctive treatment of schizophrenia. This study will be conducted as a 52-week, open-label, flexible-dose extension of Studies ACP-103-034, -038, and -064.

Subjects who have completed Studies ACP-103-034, -038, or -064, and who may continue to benefit from adjunctive pimavanserin treatment based on the Investigator’s judgment will be included in this long-term extension study.

Study ACP-103-035 subjects **must be** consented prior to the final procedures being performed at Week 6 for Studies ACP-103-034 or at Week 26 for Study ACP-103-038 or -064. Procedures performed at the End-of-Study (EOS) visits of these three double blind studies (ACP-103-034, -038, or -064) will be carried over to the ACP-103-035 study to be included as baseline information, and this visit will be considered the Baseline visit (Visit 1) of the ACP-103-035 study.

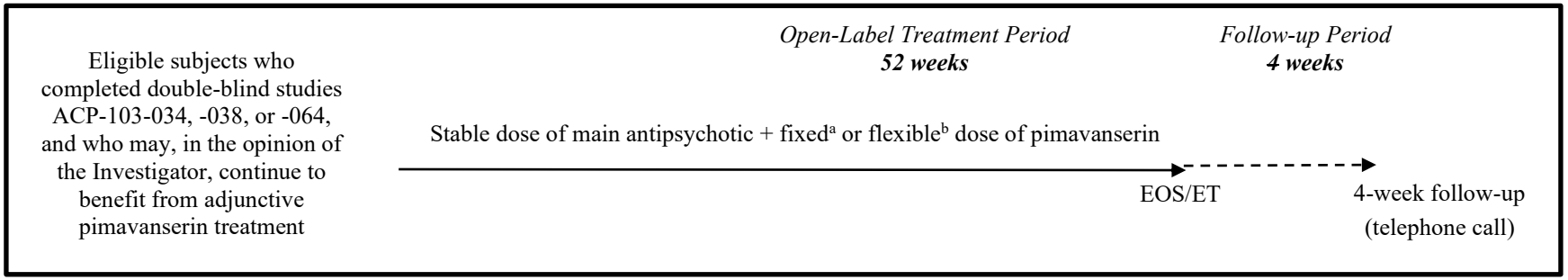
All subjects will receive once daily (QD) doses of pimavanserin over 52 weeks of treatment. Subjects transitioning from Studies ACP-103-034 and 038 will start at a dose of 20 mg pimavanserin for the first 2 weeks. After the Week 2 visit, the daily dose can be adjusted to 34 mg, 20 mg, or 10 mg pimavanserin throughout the study, based on the Investigator's assessment of clinical response. Dose adjustments may be made at scheduled or unscheduled visits (which may occur prior to the Week 2 visit). Subjects transitioning from the ACP-103-064 study will start at a dose of 34 mg and remain at 34 mg for the duration of the Treatment Period. During the Treatment Period, clinic visits will be conducted at Baseline and Weeks 2, 6, 12, 20, 28, 36, 44, and 52, or upon early termination (ET) from the study.

Study drug will be dispensed to the subject to take home at the Baseline visit and at each subsequent visit. The subject and their study partner/caregiver will be provided instructions for subject's first dose of study drug on the day after the Baseline visit. It is recommended that the subject take the study drug at approximately the same time each day as a single, oral dose. The main antipsychotic, all concomitant antidepressants, anxiolytics, and other permitted medications should remain at a stable dose throughout the study, if possible. Adjustments in the dose of the main antipsychotic after Baseline are discouraged to minimize confounding interpretation of the pimavanserin dose changes and to better understand the reason for treatment discontinuation.

In addition to the EOT or ET visit performed at time of study completion or premature discontinuation, a follow-up safety assessment will be conducted by telephone call at approximately 4 weeks after the last dose of study drug.

Figure 1 illustrates the study design.

Figure 1 Schematic of Study Design



[a] For subjects from Study ACP-103-064 only
[b] For subjects from Studies ACP-103-034 and 038 only

3.2 Schedule of Assessments

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

Table 1 Schedule of Events and Assessments

Period	Treatment Period									Follow-Up[a]
Visit[b]	1	2	3	4	5	6	7	8	9 (EOT/ET)	Telephone call
Day or Week	Baseline[c]	Week 2	Week 6	Week 12	Week 20	Week 28	Week 36	Week 44	Week 52	Week 56
Allowable visit window (# days)	0	+3	+3	+3	+7	+7	+7	+7	+7	+7
Informed consent[d]	X									
Inclusion/exclusion criteria	X									
Physical examination[e]	X								X	
Vital signs	X	X	X	X	X	X	X	X	X	
Weight	X	X	X	X	X	X	X	X	X	
12-lead ECG[f]	X	X				X			X	
Verbal confirmation of continued antipsychotic use from double-blind studies ACP-103-034, 038, or 064	X	X	X	X	X	X	X	X	X	
Clinical laboratory tests[g]	X					X			X	
Urine pregnancy test (female subjects of childbearing potential only)[h]	X	X	X	X	X	X	X	X	X	
AIMS, BARS, and SAS	X	X		X		X		X	X	
C-SSRS (Since Last Visit version)	X	X	X	X	X	X	X	X	X	
CGI-S and/or CGI-SCH-S[i]	X	X	X	X	X	X	X	X	X	
PANSS and IQ-PANSS	X			X		X			X	
NSA-16 (for ACP-103-038 and 064 subjects only)	X					X			X	
PSP	X			X		X			X	
SF-36 and DAI-10 (for ACP-103-034 and 038 subjects only)	X			X		X			X	
WoRQ (for ACP-103-064 subjects only)	X			X		X			X	
Assessment of concomitant medications	X	X	X	X	X	X	X	X	X	X
Assessment of adverse events[j]	X	X	X	X	X	X	X	X	X	X
Dispense study drug	X	X[k]	X[k]	X[k]	X[k]	X[k]	X[k]	X[k]		
Study drug accountability		X	X	X	X	X	X	X	X	

Footnotes for Table 1 on next page.

Abbreviations: AIMS=Abnormal Involuntary Movement Scale; BARS=Barnes Akathisia Rating Scale; CGI-S=Clinical Global Impression – Severity; CGI-SCH-S=Clinical Global Impression of Schizophrenia – Severity; C-SSRS=Columbia-Suicide Severity Rating Scale; DAI-10=10-item Drug Attitude Inventory; ECG=electrocardiogram; EOS=End of Study; EOT=End of Treatment; ET=Early Termination; IQ-PANSS=Informant Questionnaire for the Positive and Negative Syndrome Scale; NSA-16=Negative Symptom Assessment-16 scale; PANSS=Positive and Negative Syndrome Scale; PSP=Personal and Social Performance Scale; SAS=Simpson-Angus Extrapyramidal Side Effects Scale; SF-36=36-item Short Form Health Survey; WoRQ=Work Readiness Questionnaire

[a] For subjects who complete the study or who discontinue prematurely from the study, a safety follow-up telephone call will occur approximately 4 weeks 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 2, 3, and 4; ±7 days for Visits 5, 6, 7, 8, and 9; and +7 for a follow-up telephone call approximately 4 weeks after last dose of study drug. Circumstances may arise (e.g., pandemic, natural disaster, or political upheaval) when on-site assessments of efficacy and 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] Procedures performed at the EOS visits for Studies ACP-103-034, 038, or EOT visit for 064, will be carried over as baseline information, if applicable.

[d] Study ACP-103-035 subjects **must be** consented prior to final procedures being performed at the Week 6 visit for Studies ACP-103-034 or at the Week 26 visit for Studies ACP-103-038 and 064. The subject's caregiver must provide written agreement indicating their agreement to participate in the study in the caregiver role.

[e] A complete physical examination should be performed at Baseline and Week 52. Pelvic and/or urogenital examination may be deferred, unless the Investigator deems this to be clinically indicated.

[f] A single 12-lead ECG can be performed any time before blood sampling or at least 30 minutes after blood sampling during clinic visits. The ECG should not be recorded from the same arm as the blood draw if taken after blood draw. The subject must rest in a supine position for 5 minutes before the ECG is obtained. ECG tracings (paper or electronic) will be reviewed and interpreted by a qualified clinician. ECG tracings and results (ventricular rate, PR, QRS, QT, QTcF and QTcB intervals) will be included in the subject's study records. 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.

[g] To include hematology, serum chemistry, prolactin levels, and urinalysis (note: UDS and 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 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.

[h] Urine pregnancy tests will be completed at all visits for all female subjects of childbearing potential. Additionally, home urine pregnancy tests should be performed at Weeks 10, 16, 24, 32, 40, and 48 for all female subjects of childbearing potential; sites should ascertain the results of these home tests by telephone. Any positive urine pregnancy test at home or the site must be followed by a confirmatory serum pregnancy test.

[i] The CGI-S will be administered to subjects from all three antecedent studies. The CGI-SCH-S will be administered only to subjects who completed Studies ACP-103-038 and 064; for the purposes of this study, only the negative symptoms will be evaluated.

[j] Adverse events will be recorded in study ACP-103-035 from the first dose of open-label study drug until 30 days after the last dose of study drug.

[k] Subjects are to return unused study drug and study drug bottles at each subsequent visit; new study drug bottle(s) 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.

3.3 Randomization

Not applicable.

3.4 Blinding

Not applicable.

3.5 Determination of Sample Size

The planned sample size for this study is not based on statistical power but will depend on the number of subjects who complete double-blind Studies ACP-103-034, -038, and -064 and who transition into this open-label extension study.

4 ANALYSIS SETS

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

5 DATA HANDLING CONVENTIONS

All data collected in the study will be listed.

5.1 General Data Reporting Conventions

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

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.

No hypothesis testing is planned. Descriptive summaries of all safety and efficacy endpoints will be provided.

All safety endpoints will be summarized for the overall Safety Analysis Set by treatment groups includes PIM 34 mg and PIM Flexible dose. Treatment group of "PIM 34 mg" will include subject rolled over from Study 064 and subjects rolled over from Studies 034 or 038 who changed from the initial dose of 20 to 34 mg and stayed on 34 mg through the end of study (only one dose change). Treatment group of "PIM Flexible dose" will include subject rolled over from Studies 034 or 038. In addition, an "All PIM" group will include all subjects who received pimavanserin at any dose level in Study 035.

In addition, safety summaries by cohorts based on the antecedent study (ACP-103-034 or ACP-103-038/-064) and treatment group according to the original double–blinded treatment in the antecedent study (placebo or pimavanserin) will also be provided.

For efficacy endpoints, summaries by cohorts based on the antecedent study (ACP-103-034, -038, and -064) and treatment group according to the original double–blinded treatment in the antecedent study (placebo or pimavanserin) will be provided.
For each continuous measure in safety and efficacy analyses, change from Baseline results will be presented in two ways:

1. Main analysis: using the Baseline from Study ACP-103-035 and reporting the changes across Study ACP-103-035 timepoints
2. Exploratory analysis: using the Baseline from the antecedent study (ACP-103-034, -038, and -064), and reporting the changes across the timepoints of both the double-blind antecedent study and the open-label Study ACP-103-035.

The analysis using the Baseline from the first dose of pimavanserin in either the double-blind or the open-label study is included in the two methods outlined above: for the cohort of subjects who received placebo in the antecedent double-blind study, method 1 results represent the changes from the first dose of pimavanserin; for the cohort of subjects who received pimavanserin in the antecedent double-blind study, method 2 results represent the changes from the first dose of pimavanserin.

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 Clinical Global Impressions Scale – Severity (CGI-S)

The CGI-S is assessed at Baseline, Weeks 2, 6, 12, 20, 28, 36, 44 and Week 52/ET visits. Subjects rolled over from all antecedent studies will be administered the CGI-S scale.

The CGI-S is a clinician-rated, 7-point scale that is designed to rate the severity of the subject's illness, in this case schizophrenia, at the time of assessment, making use of the clinician's judgment and past experience with subjects who have the same disorder. 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 schizophrenia. The licensed version used in this study is a blend of Guy's (Guy W; Early Clinical Drug Evaluation Unit Assessment Manual for Psychopharmacology; US Department of Health, Education, and Welfare; National Institute of Mental Health; 1976) and Haro's (Haro JM; The Clinical Global Impression-schizophrenia scale: a simple instrument to measure the diversity of symptoms present in schizophrenia; 2003) versions.

Missing CGI-S 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 Baseline, Weeks 2, 6, 12, 20, 28, 36, 44 and Week 52/ET visits. Only subjects rolled over from antecedent Study ACP-103-038 and ACP-103-064 will be administered the CGI-SCH-S scale and only the negative symptoms will be assessed for this scale.

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 Positive and Negative Syndrome Scale (PANSS)

The PANSS is assessed at Baseline, Weeks 12, 28 and Week 52/ET visits. Subjects rolled over from all antecedent studies will be administered the PANSS scale.

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 each of the subscales (positive, negative, or general), if more than 1 item score is missing, then that particular subscale score and the PANSS total score will be missing. When there is only 1 item missing for a subscale, then the missing single item will be imputed using the average of the non-missing item scores for that subscale, subject and timepoint, rounded to the nearest integer. After the missing item(s) is imputed within the subscale(s), then the PANSS total score and Marder factor scores will be computed without further imputation. A Marder factor score

will be missing if any item score is missing for that factor, after the imputations within the subscales were applied.

5.2.4 Negative Symptom Assessment-16 (NSA-16)

The NSA-16 is assessed at Baseline, Week 28 and Week 52/ET visits. Only subjects rolled over from antecedent Study ACP-103-038 and ACP-103-064 will be administered the NSA-16 scale.

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. The “normal” (score = 1) reference is based on the comparison to a young person in their twenties without schizophrenia. 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 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.

Additionally, 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. Missing global negative symptoms rating scores will not be imputed.

5.2.5 Personal and Social Performance Scale (PSP)

The PSP is assessed at Baseline, Weeks 12, 28 and Week 52/ET visits. Subjects rolled over from all antecedent studies will be administered the PSP scale.

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 Work Readiness Questionnaire (WoRQ)

The WoRQ is assessed at Baseline, Weeks 12, 28 and Week 52/ET visits. Only subjects rolled over from antecedent Study ACP-103-064 will be administered the WoRQ.

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.

5.2.7 36-Item Short Form Health Survey (SF-36)

The SF-36 is assessed at Baseline, Weeks 12, 28 and Week 52/ET visits. Subjects rolled over from all antecedent studies will be administered the SF-36 scale.

The SF-36 is a 36-item survey that measures the overall health status of a subject. The SF-36 assesses eight health concepts during the past 4 weeks: physical functioning, bodily pain, role limitations due to physical health problems, role limitations due to personal or emotional problems, emotional well-being, social functioning, energy/fatigue, and general health perceptions. It also includes a single item that provides an indication of perceived change in health compared to one year ago. In Study ACP-103-035, the RAND 36-Item Health Survey (Version 1.0) is used.

Scoring the SF-36 is a 2-step process. First, pre-coded numeric values are re-coded per the scoring key given in Table 2. Note that all items are scored so that a high score defines a more favorable health state and less disability. In addition, each item is scored on a 0 to 100 range so that the lowest and highest possible scores are 0 and 100, respectively. Scores represent the percentage of total possible score achieved, with higher scores denoting higher level of functioning. In step 2, items in the same scale are averaged together to create the 8 scale scores. Table 3 lists the items averaged together to create each scale. Items that are left blank (missing data) are not taken into account when calculating the scale scores. Hence, scale scores represent the average for all items in the scale that the respondent answered.

Table 2 SF-36 Re-Coding Rules for Individual Items

Item Number	Original Response Category[1]	Re-Coded Value
2, 20, 34, 36	1	100
	2	75
	3	50
	4	25
	5	0
3, 4, 5, 6, 7, 8, 9, 10, 11, 12	1	0
	2	50
	3	100
13, 14, 15, 16, 17, 18, 19	1	0
	2	100
23, 26, 27, 30	1	100
	2	80
	3	60
	4	40
	5	20
	6	0
24, 25, 28, 29, 31	1	0
	2	20
	3	40
	4	60
	5	80
	6	100
32, 33, 35	1	0
	2	25
	3	50
	4	75
	5	100
1	1	100
	2	85
	3	60
	4	25
	5	0
21	1	100
	2	88
	3	64
	4	42
	5	24
	6	0

Table 2 SF-36 Re-Coding Rules for Individual Items (Continued)

Item Number	Original Response Category[1]	Re-Coded Value
22	1 (Item 21 = 1 or .)	100
	1 (Item 21 = 2 to 6)	80
	2 (Item 21 = .)	75
	2 (Item 21 ^= .)	60
	3 (Item 21 = .)	50
	3 (Item 21 ^= .)	40
	4 (Item 21 = .)	25
	4 (Item 21 ^= .)	20
	5	0

[1] Pre-coded response choices as printed in the questionnaire.

Table 3 SF-36 Scale Components

Scale	Items
PHYFUN10: Physical functioning	3, 4, 5, 6, 7, 8, 9, 10, 11, 12
ROLEP4: Role limitations due to physical health	13, 14, 15, 16
ROLEE3: Role limitations due to emotional problems	17, 18, 19
ENFAT4: Energy/fatigue	23, 27, 29, 31
EMOT5: Emotional well-being	24, 25, 26, 28, 30
SOCFUN2: Social functioning	20, 32
SFPAIN2: Pain	21, 22
SFGENH5: General health	1, 33, 34, 35, 36

To enable meaningful comparisons across scales, the 8 scale scores are converted to norm-based z-scores based on 1998 General Population statistics:

PF_Z = (PHYFUN10 - 82.96845) / 23.83795 ;
RP_Z = (ROLEP4 - 77.93107) / 35.34865 ;
RE_Z = (ROLEE3 - 83.10276) / 31.64149 ;
EN_Z = (ENFAT4 - 56.99917) / 21.12677 ;
EM_Z = (EMOT5 - 75.21913) / 17.60698 ;
SF_Z = (SOCFUN2 - 83.56494) / 23.02758 ;
BP_Z = (SFPAIN2 - 70.22865) / 23.35310 ;
GH_Z = (SFGENH5 - 70.10060) / 21.35900 ;

In addition, physical and mental health composite scores will also be computed using the individual-scale z-scores:

Physical health composite scores AGPHYS_Z
= (PF_Z * 0.42402) + (RP_Z * 0.35119) + (BP_Z * 0.31754) +
(GH_Z * 0.24954) + (EM_Z * -.22069) + (RE_Z * -.19206) +
(SF_Z * -.00753) + (EN_Z * 0.02877);

Mental health composite scores AGMENT_Z
= (PF_Z * -.22999) + (RP_Z * -.12329) + (BP_Z * -.09731) +
(GH_Z * -.01571) + (EM_Z * 0.48581) + (RE_Z * 0.43407) +
(SF_Z * 0.26876) + (EN_Z * 0.23534);

The 2 composite scores will not be calculated if any individual scale score is missing.

Finally, the z-scores will be transformed to t-scores for data summaries_Exploratory_Efficacy_Analyses.

```
AGPHYS_T = 50 + (AGPHYS_Z * 10);
AGMENT_T = 50 + (AGMENT_Z * 10);
PF_T = 50 + (PF_Z * 10) ;
RP_T = 50 + (RP_Z * 10) ;
RE_T = 50 + (RE_Z * 10) ;
EN_T = 50 + (EN_Z * 10) ;
EM_T = 50 + (EM_Z * 10) ;
SF_T = 50 + (SF_Z * 10) ;
BP_T = 50 + (BP_Z * 10) ;
GH_T = 50 + (GH_Z * 10) ;
```

5.2.8 Drug Attitude Inventory (DAI-10)

The DAI-10 is assessed at Baseline, Weeks 12, 28 and Week 52/ET visits. Subjects rolled over from all antecedent studies will be administered the DAI-10 scale.

The DAI-10 contains 6 items (1, 3, 4, 7, 9, and 10) that a subject who is fully adherent to the prescribed medication would answer as "True" and 4 items (2, 5, 6, and 8) that a subject who is fully adherent to the prescribed medication would answer as "False." A correct answer is scored +1 and an incorrect answer is scored -1. The total score is the sum of pluses and minuses, which can range from -10 to 10 in increments of 2. A positive total score indicates a positive subjective response (adherent) and a negative total score indicates a negative subjective response (non-adherent). Higher scores denote better adherence.

Missing DAI-10 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 Baseline, Weeks 2, 6, 12, 20, 28, 36, 44 and Week 52/ET visits. C-SSRS is a safety assessment and subjects rolled over from all antecedent studies will be assessed.

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 version assessing information since the last visit will be completed at all 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 Baseline, Weeks 2, 12, 28, 44 and Week 52/ET visits. AIMS is a safety assessment and subjects rolled over from all antecedent studies will be assessed.

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 Baseline, Weeks 2, 12, 28, 44 and Week 52/ET visits. BARS is a safety assessment and subjects rolled over from all antecedent studies will be assessed.

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 Baseline, Weeks 2, 12, 28, 44 and Week 52/ET visits. SAS is a safety assessment and subjects rolled over from all antecedent studies will be assessed.

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

5.5 Baseline Definition

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

5.6 Analysis Visit Windows

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

Table 4 Analysis Visit Windows

Open-label (OL) Analysis Visit Name	Target Study Day[1]	Study Day Interval
OL Baseline	1	≤1 (based on the first OL dose date)
OL Week 2	15	2 to 28
OL Week 6	43	29 to 63
OL Week 12	85	64 to 112
OL Week 20	141	113 to 168
OL Week 28	197	169 to 224
OL Week 36	253	225 to 280
OL Week 44	309	281 to 336
OL Week 52	365	≥ 337[2]

1 If the assessment date ≥ first OL dose date, study day = assessment date - first OL dose date + 1, otherwise study day = assessment date – first OL dose. Study day 1 is the first OL dose date.

2 Efficacy assessment collected more than seven days after the last OL dosing date will be excluded from the efficacy analyses.

5.6.1 Unscheduled Assessments

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

5.6.2 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 summaries. If two assessments are equidistant from the target day then the chronologically last assessment will be used for summary. Exceptions may be made for incomplete assessments, in which case, more complete assessments may be given priority. Details are provided in a separate programming conventions document.

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

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 date of the end-of-study/early termination visit will be used in the calculation of treatment duration. For the incomplete last dose date of the study drug, the imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

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

5.8 Missing or Incomplete Dates for Concomitant or Post-Treatment 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 the analysis dataset specification document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

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

5.10 Missing Severity Assessment for Adverse Events

If the severity is missing for a treatment-emergent AE, 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 a treatment-emergent AE, 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 summary due to, for example, a character string reported for a numeric variable, an appropriately determined coded value will be used in the summary. The coding algorithms will be detailed in the analysis dataset specification document. The actual values as reported in the database will be presented in data listings.

5.13 Duplicate Subjects

Duplicate-subject screening will be performed throughout the schizophrenia program to identify whether there are individuals who are randomized more than once into the double-blind schizophrenia studies ACP-103-034, -038 or -064. Confirmed duplicates will only be included for summaries under the subject number and antecedent study to which they were first randomized and rolled over from. Data collected under other subject number(s) or antecedent study (studies) will be listed but will not be summarized. Case narratives will be provided for duplicate subjects.

6 SUBJECT DISPOSITION

The number of sites that enrolled at least 1 subject and number of subjects enrolled will be summarized by region (North America, Europe, and rest of world) and overall. Enrolled subjects are subjects who signed informed consent for Study ACP-103-035 and not recorded as "rollover failure" on eCRF.

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

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

7 PROTOCOL DEVIATIONS

Protocol deviations that occurred during the extension study period will be reviewed periodically over the course of the study. The review process, definition of the deviation categories, and the

classification of a deviation as major or minor are detailed in the Protocol Deviation Project Instruction document. Protocol deviations will also be assessed with respect to relationship to COVID-19.

For enrolled subjects, a summary of the number and percentage of subjects with major protocol deviations for each deviation category will be presented in three ways: all protocol deviations, COVID-19 related protocol deviations, and non-COVID-19 related protocol deviations.

Three listings of major 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 open-label Baseline characteristics will be summarized for the Safety Analysis Set using descriptive statistics. Summaries by region (North America, Europe, and rest of world) 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 PANSS total score, Baseline CGI-S, Baseline CGI-SCH-S, Baseline NSA-16 total score and Baseline PSP score. Note that CGI-SCH-S and NSA-16 scores are only collected from subjects who rolled over from Study ACP-103-038 and -064.

Age (the eCRF reported age reflects a subject's age at the open-label Baseline visit date), Baseline PANSS total score, Baseline CGI-S, Baseline CGI-SCH-S, Baseline NSA-16 total score, Baseline PSP score and Baseline WoRQ score will be presented as both continuous and categorical variables. Age categories will be presented as 18 to 35 and >35 years old. Baseline PANSS will be presented as <90 (lower symptom severity) or $\geq$90 (higher symptom severity). Baseline CGI-S and CGI-SCH-S categories will be presented as scores 1 to 7 as well as $\leq$4 (lower disease severity) or $\geq$5 (higher disease severity). Baseline NSA-16 will be presented as $\leq$55 (lower symptom severity) or >55 (higher symptom severity). Baseline PSP will be presented by deciles (e.g., scores of 31 to 40, 41 to 50, etc.).

Schizophrenia disease history will be summarized for Safety 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) (this information is only collected for subjects who rolled over from Study ACP-103-038 and -064)
- Time (years) since first antipsychotic treatment
- Treated with clozapine in the past (yes or no) (this information is only collected for subjects who rolled over from Study ACP-103-034)
- Number of hospitalizations for treatment of schizophrenia (0, 1 to 5, 6 to 10, or >10)
- Current main background antipsychotic medication
- Duration (months) of current main background antipsychotic medication
- Had suicidal ideation or behavior at open-label Baseline (yes or no)
- Had non-suicidal self-injurious behavior at open-label Baseline (yes or no)

Open-label Baseline visit 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 received clozapine, date of last clozapine dose and the dose level will also be listed. 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 23.0 or newer. The subject incidence will be summarized for each system organ class (SOC) and preferred term for the 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

10.1.1 Exposure to open-label pimavanserin

For each subject, the duration of exposure to open-label pimavanserin (open-label last dose date – open-label first dose date + 1), cumulative dose (first multiply the number of doses taken by the dose level for each kit utilized during the open-label study period, then sum the results from all kits), and average daily dose (cumulative dose in mg divided by duration of open-label exposure in days) will be calculated and summarized.

Duration of OL pimavanserin 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 <6 weeks (14 to 41 days), 6 to <12 weeks (42 to 83 days), 12 to <20 weeks (84 to 139 days), 20 to <28 weeks (140 to 195 days), 28 to <36 weeks (196 to 251 days), 36 to <44 weeks (252 to 307 days), 44 to <52 weeks (308 to 363 days), and ≥52 weeks (364 days or longer). Kaplan-Meier curves of duration on study drug will also be presented. In addition, for subjects from Study ACP-103-034 and -038, summaries of whether subjects had any dose change (yes vs. no), their highest dose level (20 mg or 34 mg), their lowest dose level (10 mg or 20 mg), and their last dose level (10 mg, 20 mg, or 34 mg) will also be provided.

The pimavanserin dose levels are expressed as free base.

10.1.2 Total exposure to pimavanserin across the double-blind and open-label periods

For subjects who received pimavanserin in the antecedent studies, the total duration of exposure (in days) to pimavanserin will be calculated as the sum of the exposure durations in the double-blind and open-label periods:

[(*date of last double-blind dose*) – (*date of first double-blind dose*) + 1] + [(*date of last open-label dose*) – (*date of first open-label dose*) + 1].

The cumulative dose of pimavanserin will be calculated by first multiplying the number of doses taken by the dose level (in mg) for each kit returned to the site during both the double-blind and the open-label study periods and then summing the results for all kits. The average daily dose of pimavanserin (in mg) will be calculated as the cumulative dose (in mg) divided by total duration of exposure (in days). For subjects who received placebo in the antecedent studies, the total duration of exposure to pimavanserin, the cumulative dose of pimavanserin, and the average daily dose of pimavanserin across the double-blind and open-label periods will be the same as the values calculated for the open-label study period alone. Duration of exposure, cumulative dose, and average daily dose for pimavanserin across the double-blind and open-label periods will be summarized descriptively.

Duration of total exposure to pimavanserin 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 <6 weeks (14 to 41 days), 6 to <12 weeks (42 to 83 days), 12 to <20 weeks (84 to 139 days), 20 to <28 weeks (140 to 195 days), 28 to <36 weeks (196 to 251 days), 36 to <44 weeks (252 to 307 days), 44 to <52 weeks (308 to 363 days), 52 to <60 weeks (364 to 419 days), 60 to <68 weeks (420 to 475 days), 68 to <78 weeks (476 to 545 days), and ≥78 weeks (546 days or longer). In addition, for Study ACP-103-034 and -038, summaries of whether subjects had any dose change (yes vs. no), their highest dose level (20 mg or 34 mg), their lowest dose level (10 mg or 20 mg), and their last dose level (10 mg, 20 mg, or 34 mg) will also be presented.

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

11 CONCOMITANT AND POST-TREATMENT MEDICATION

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

Medications will be coded using WHO Drug Dictionary 2020March or newer version. For concomitant medications, the number and percentage of subjects taking each drug class (ATC Level 3) and medication preferred term will be tabulated for Safety Analysis Set. A subject will be counted only once per drug class or per medication preferred term for the summary.

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

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

12.1 Efficacy Variables

Secondary Endpoints

- Change from Baseline in the Clinical Global Impressions – Severity scale (CGI-S)
- Change from Baseline in the Clinical Global Impression of Schizophrenia Scale – Severity (CGI-SCH-S) (for subjects from Study ACP-103-038 and ACP-103-064 only)

Exploratory Endpoints

- Change from Baseline in the Positive and Negative Syndrome Scale (PANSS) total score, subscores (positive scale, negative scale, and general psychopathology scale), and Marder factor scores (negative symptoms, positive symptoms, disorganized thought, uncontrolled hostility/excitement, and anxiety/depression factors) and individual item scores

  See Section 5.2.3 for the PANSS items included in each Marder factor.
- Change from Baseline in the Negative Symptom Assessment-16 (NSA-16) scale total score, domain scores (communication, emotion/affect, social involvement, motivation, or retardation), and individual item scores (for subjects from Study ACP-103-038 and ACP-103-064 only)
- Change from Baseline in the Personal and Social Performance Scale (PSP) score
- Change from Baseline in 36-item Short Form Health Survey (SF-36) score (for subjects from ACP-103-034 and ACP-103-038 only)

  See Section 5.2.6 for details
- Change from Baseline in 10-item Drug Attitude Inventory (DAI-10) score (for subjects from Study ACP-103-034 and ACP-103-038 only)
- Change from Baseline in Work Readiness Questionnaire (WoRQ) total score (for subjects from Study ACP-103-064 only)
- Change from Baseline in Work Readiness Questionnaire (WoRQ) readiness to work question (item 8) (for subjects from Study ACP-103-064 only)

12.2 Adjustment for Covariates

Not applicable.

12.3 Handling of Missing Data

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

12.4 Multiple Comparisons / Multiplicity

No hypothesis testing is planned.

12.5 Examination of Subgroups

Summaries for below subgroups will be tabulated for the change from Baseline in the CGI-S and CGI-SCH-S scores:

- region (North America, Europe, or rest of world)
- age group (18 to 35 or >35 years old)
- sex (male or female)

13 METHODS OF EFFICACY ANALYSES

13.1 Analysis of Continuous Efficacy Endpoints

Descriptive statistics for all secondary and exploratory efficacy endpoints listed in Section 12.1 will be tabulated by cohorts based on antecedent study and double-blind treatment group at scheduled timepoints. The summaries of the change from Baseline results will be presented by cohort at scheduled timepoints in two ways as specified in Section 5.1. The scheduled timepoints across the double-blind and the open-label study periods for all efficacy endpoints are presented in Table 5 below.

Table 5 Scheduled Assessment Timepoints for Efficacy Endpoints Across Double-Blind (DB) Studies ACP-103-034, -038 and -064 and Open-Label (OL) Study -035

	CGI-S		CGI-SCH-S		PANSS		NSA-16		PSP		SF-36		DAI-10		WoRQ
	034	038 /064	034	038 /064	034	038 /064	034	038 /064	034	038 /064	034	038	034	038	064
DB Baseline	x	N/A	N/A	x	x	x	N/A	x	x	x	x	x	x	x	x
DB W1	x	N/A	N/A		x		N/A								
DB W2	x	N/A	N/A	x	x		N/A	x							
DB W3	x	N/A	N/A		x		N/A								
DB W4	x	N/A	N/A	x	x		N/A	x							
DB W5	x	N/A	N/A		x		N/A								
DB W6 (034 EOS)	x	N/A	N/A		x		N/A		x		x		x		
DB W8		N/A		x				x		x					x
DB W14		N/A		x				x							x
DB W20		N/A		x				x							
DB W26 (038/064 EOS)		x		x		x		x		x		x		x	x
OL W2	x	x	N/A	x			N/A								
OL W6	x	x	N/A	x			N/A								
OL W12	x	x	N/A	x	x	x	N/A		x	x	x	x	x	x	x
OL W20	x	x	N/A	x			N/A								
OL W28	x	x	N/A	x	x	x	N/A	x	x	x	x	x	x	x	x
OL W36	x	x	N/A	x			N/A								
OL W44	x	x	N/A	x			N/A								
OL W52	x	x	N/A	x	x	x	N/A	x	x	x	x	x	x	x	x

13.2 Categorical Analysis

13.2.1 For All Subjects

Two types of responders will be counted by cohort at each scheduled post-Baseline visit:

- ≥ 20% reduction in the corrected PANSS total score (subtracting 30 points from the raw scores; see details below) from Baseline
- ≥ 30% reduction in the corrected PANSS total score (subtracting 30 points from the raw scores; see details below) from Baseline

Note that since PANSS is an interval scale that lacks a natural zero point, the percent change in PANSS total score will be calculated based on corrected scores after subtracting 30 points from the raw scores (Leucht et al., 2007; Leucht et al., 2009; Obermeier et al., 2010; Obermeier et al., 2011). For example, if a subject's Baseline PANSS total score is 80 and Week 52 PANSS total score is 30 (absent of all symptoms), the percent change from Baseline to Week 52 in PANSS total score will be calculated as:

$$[(30 - 30) - (80 - 30)] \div (80 - 30) \times 100\% = -100\%.$$

At any given visit, the proportion of responders will be calculated using 2 different methods:

- Observed cases (subjects with missing PANSS total score are excluded from the calculation): number of responders divided by number of subjects with non-missing PANSS total score for the given visit and cohort.
- Missing as failures (subjects with missing PANSS total score are considered as non-responders in the calculation): number of responders divided by total number of subjects in the Safety Analysis Set for the given cohort.

13.2.2 For Subjects Rolled Over from Study ACP-103-038 and -064

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

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 52 NSA-16 total score is 16 (absent of all symptoms), the percent change from Baseline to Week 52 in NSA-16 total score will be calculated as:

$$[(16 - 16) - (60 - 16)] \div (60 - 16) \times 100\% = -100\%.$$

At any given visit, the proportion of responders will be calculated using 2 different methods:

- Observed cases (subjects with missing NSA-16 total score are excluded from the calculation): number of responders divided by number of subjects with non-missing NSA-16 total score for the given visit and cohort.
- Missing as failures (subjects with missing NSA-16 total score are considered as non-responders in the calculation): number of responders divided by total number of subjects in the Safety Analysis Set for the given cohort.

13.2.3 For Subjects Rolled Over from Study ACP-103-064

WoRQ readiness to work question (item 8) will also be summarized in shift tables by visit.

14 SAFETY ANALYSES

The safety summaries will be presented for the Safety Analysis Set using descriptive statistics. Safety variables include AEs, clinical laboratory variables, vital signs, body weight, BMI, physical examinations, ECG, C-SSRS, AIMS, BARS, and SAS variables. Safety variables will also be summarized by cohorts based on antecedent study and double-blind treatment group. For each continuous measure in clinical laboratory variables, vital signs, and electrocardiogram, change from Baseline results will be presented in two ways as specified in Section 5.1.

14.1 Adverse Events

All Adverse events (AEs) will be coded using MedDRA version 23.0 or newer.

An AE (classified by preferred term) will be considered a treatment-emergent AE (TEAE) if it started on or after the open-label first study dose date and no later than the open-label last study dose date + 30 days.

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

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

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.

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 the open-label first dose, date of the open-label last dose, dose level at AE onset, severity, frequency, outcome, relationship to study drug, action taken with study drug, and required therapy. When a date is presented, the study day associated with the date will also be displayed. Separate listings will be presented for subjects with SAEs, subjects with AEs leading to discontinuation and subjects who died (if any). In these listings, an indicator for TEAEs will also be included. 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 Baseline, Week 28 and Week 52/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.

- Hematology tests include the following:
  - Complete blood count (CBC) including:
    - White blood cell (WBC) count
    - Complete differential (relative and absolute)
    - Hematocrit (Hct), hemoglobin (Hgb), red blood cells (RBC), platelets
    - Reticulocyte
- Serum chemistry tests include the following:
  - Sodium (Na), potassium (K), chloride (Cl), phosphorus (P), calcium (Ca), carbon dioxide (CO_2), blood urea nitrogen (BUN), creatinine (Cr), uric acid
  - Alanine aminotransferase (ALT/SGPT), aspartate aminotransferase (AST/SGOT), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), total bilirubin (TBIL), lactate dehydrogenase (LDH), glucose
  - Albumin (ALB), total protein
  - Creatine kinase (CK)/creatine phosphokinase (CPK)
  - Prolactin
  - Lipid panel
    - Total cholesterol, high-density lipoprotein (HDL)-cholesterol, triglycerides, low-density lipoprotein(LDL)-cholesterol, cholesterol/HDL ratio, non-HDL cholesterol

  Note: 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
- Urinalysis tests include the following:
  - Occult blood, leukocyte esterase, protein, glucose, ketones, specific gravity, pH
  - Urine drug screen tests may be performed at any time throughout the 52-week Treatment Period, at the discretion of the Investigator. Urine drug screen will include testing for the following substances: tetrahydrocannabinol (THC), benzodiazepines, barbiturates, cocaine, amphetamine, methamphetamine, Ecstasy, opiates, methadone, oxycodone, buprenorphine, and phencyclidine. A positive UDS for benzodiazepines will be evaluated by the Investigator in the context of allowed anxiolytics

- A urine pregnancy test will be performed for all women at Baseline and at each scheduled visit. Baseline pregnancy test results must be confirmed to be negative before a subject is administered any study drug
- Additionally, home urine pregnancy tests should be performed at Weeks 10, 16, 24, 32, 40, and 48 for women of childbearing potential; sites should ascertain the results of these home tests by telephone.
- Any positive urine pregnancy test at home or the site must be followed by a confirmatory serum pregnancy test.

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.

Clinical laboratory values reported as continuous values for hematology, chemistry and urinalysis will be summarized using descriptive statistics at Baseline and scheduled post-Baseline visits. The change from Baseline values will also be summarized at scheduled post-Baseline visits. The overall minimum and maximum post-Baseline observed and change from Baseline values will also be summarized (for open-label study period only). For hemoglobin, hematocrit and uric acid, the above summaries will be presented for each gender as well as for both genders combined. For urinalysis with categorical results, the number and percentage of subjects will be tabulated by category at Baseline and scheduled post-Baseline visits, and the denominator is the number of subjects with non-missing values for the given parameter, visit and cohort.

Laboratory values will also be summarized in shift tables to determine the number and percentage of subjects with values classified as below (low), within (normal), or above (high) normal ranges at scheduled post-Baseline visits relative to the same classification at the Baseline visit. The shifts from Baseline to overall post-Baseline minimum and overall post-Baseline maximum will also be presented (for open-label study period only). 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 cohort. For the summaries of shift from Baseline 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 cohort. For hemoglobin, hematocrit and uric acid, the shift summaries will be presented for each gender as well as for both genders combined.

Number and percentage of subjects with potentially clinically important (PCI) laboratory values at scheduled post-Baseline visits and overall post-Baseline (for open-label study period only) will be summarized by Baseline status (all or within normal range) for selected parameters. PCI criteria are listed in Table 6 and Table 7. 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 value for the given parameter, visit and cohort, and the denominator is the number of subjects with non-missing values for the given parameter, visit and cohort. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI value for the given parameter and cohort, and the denominator is the number of subjects with at least 1 post-Baseline value for the given parameter and cohort. For hemoglobin, hematocrit, and uric acid, the count and percentage of subjects with PCI values will be presented for each gender as well as for both genders combined. A listing of all PCI values in study ACP-103-035 will be provided. This listing will include all observations from study ACP-103-035 for those subjects and parameters for which at least one PCI value (including Baseline) was observed.

Table 6 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 7 Criteria for Potentially Clinically Important Laboratory Values - Urinalysis

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

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

14.3 Vital Signs

Vital signs are assessed at Baseline, Weeks 2, 6, 12, 20, 28, 36, 44 and Week 52/ET visits.

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

Vital sign values will be considered PCI if they meet both the observed value criteria and the change from Baseline criteria listed in Table 8. The number and percentage of subjects with post-Baseline vital signs that are PCI will be summarized at scheduled post-Baseline visits and for overall post-Baseline (for open-label study period only). 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 value for the given parameter, visit and cohort, and the denominator is the number of subjects with non-missing values for the given parameter, visit and cohort. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI value for the given parameter and cohort, and the denominator is the number of subjects with at least 1 post-Baseline value for the given parameter and cohort. A listing of all PCI values in study ACP-103-035 will be provided. This listing will include all observations from study ACP-103-035 for those subjects and parameters for which at least one PCI value (including Baseline) was observed.

Table 8 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)

In countries outside of Czech Republic, ECG is performed at Baseline, Weeks 2, 28 and Week 52/ET visits. At the clinical sites in Czech Republic, ECG is performed at all visits. For ECG data summaries, the common scheduled visits across all clinical sites will be used (i.e. Baseline, Weeks 2, 28 and Week 52/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, including the cardiologist's interpretation.

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

QTc intervals include QTcB (Bazett's formula) and QTcF (Fridericia's formula). QTcF will also be categorized into the following categories (msec) and the number and percentage of subjects in each category will be summarized at scheduled visits and overall post-baseline maximum (for open-label study period only):

- 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 (for open-label study period only) will also be summarized (i.e. if a

subject has one or more 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 (for open-label study period only) 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 cohort. 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 cohort.

Electrocardiogram values will be considered PCI if they meet or exceed the upper limit values listed in Table 9. The number and percentage of subjects with PCI values will be summarized at scheduled post-Baseline visits and for overall post-Baseline (for open-label study period only). 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 cohort, and the denominator is the number of subjects with non-missing values for the given parameter, visit and cohort. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least 1 post-Baseline PCI value for the given parameter and cohort, and the denominator is the number of subjects with at least 1 post-Baseline value for the given parameter and cohort. A listing of all PCI values in study ACP-103-035 will be provided. This listing will include all observations from study ACP-103-035 for those subjects and parameters for which at least one PCI value (including Baseline) was observed.

Table 9 Criteria for Potentially Clinically Important ECG Values

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

14.5 Physical Examination

Physical examination is performed at Baseline and Week 52/ET visits. Note that pelvic and/or urogenital examination may be deferred, unless the Investigator deems this to be clinically indicated.

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

14.6 Other Safety Variables

14.6.1 Suicidality

The C-SSRS is assessed at Baseline, Weeks 2, 6, 12, 20, 28, 36, 44 and Week 52/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.

The event counts and the number and percentage of subjects reporting any post-Baseline non-suicidal self-injurious behavior will also be tabulated.

For calculating the percentages, the denominator will be the number of subjects in the Safety Analysis Set within each cohort.

14.6.2 Extrapyramidal Symptom Measures

14.6.2.1 Abnormal Involuntary Movement Scale (AIMS)

The AIMS is assessed at Baseline, Weeks 2, 12, 28, 44 and Week 52/ET visits.

The AIMS total score will be summarized using descriptive statistics at Baseline and scheduled post-Baseline visits. The change from Baseline scores will also be summarized at scheduled post-Baseline visits.

In addition, the number and percentage of subjects with dyskinesia will be summarized at scheduled visits and overall post-Baseline (for open-label study period only). 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 time frame, all assessments will be evaluated using the above criteria. The tabulations will be presented for subjects who have at least 1 AIMS assessment as well as for a subset of these subjects who do not have 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 Baseline, Weeks 2, 12, 28, 44 and Week 52/ET visits.

The BARS total score and the Global Clinical Assessment of Akathisia score will be summarized using descriptive statistics at Baseline and scheduled post-Baseline visits. The change from Baseline scores will also be summarized at scheduled post-Baseline visits.

In addition, the number and percentage of subjects with akathisia will be summarized at scheduled visits and overall post-Baseline (for open-label study period only). Akathisia is defined as having a Global Clinical Assessment of Akathisia score ≥2. If there are multiple assessments performed within the same time frame, all assessments will be evaluated using the above criteria. The tabulations will be presented for subjects who have at least 1 BARS assessment as well as for a subset of these subjects who do not have 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 Baseline, Weeks 2, 12, 28, 44 and Week 52/ET visits.

The SAS total score will be summarized using descriptive statistics at Baseline and scheduled post-Baseline visits. The change from Baseline scores will also be summarized at scheduled post-Baseline visits.

The number and percentage of subjects with Parkinsonism will be summarized at scheduled visits and overall post-Baseline (for open-label study period only). Parkinsonism is defined as having a SAS total score >3. If there are multiple assessments performed within the same time frame, all assessments will be evaluated using the above criteria. The tabulations will be presented 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

Not applicable.

16 INTERIM ANALYSIS

An interim analysis will be conducted prior to study completion.

17 DATA MONITORING/REVIEW COMMITTEE

Safety data are monitored throughout the study and aggregate safety reports are produced and reviewed approximately quarterly. In addition, safety data from this study will be presented periodically to the Data and Safety Monitoring Board (DSMB) for Studies ACP-103-034, -038 and -064 to help with the interpretation of benefit versus risk in those studies.

18 COMPUTER METHODS

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

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

19 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

None.

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.

Obermeier M, et al. Should the PANSS be rescaled? *Schizophrenia Bulletin.* 2010; 36(3): 455-460.

Obermeier M, et al. Is the PANSS used correctly? A systematic review. *BMC Psychiatry.* 2011; 11:113.

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

21 APPENDICES

21.1 Summary of Version Changes

Version No:	Document History Description of Update	Author(s)	Version Date
1.0	Original version	PPD	26 JAN 2024