Janssen Research & Development, LLC

Statistical Analysis Plan

A Prospective, Matched-Control, Randomized, Open-Label, Flexible-Dose, Study in Subjects with Recent-Onset Schizophrenia or Schizophreniform Disorder to Compare Disease Progression and Disease Modification Following Treatment with Paliperidone Palmitate Long-Acting Injection or Oral Antipsychotics

Disease Recovery Evaluation and Modification (DREaM) Study
Phase 3b

Protocol R092670SCH3013

R092670 (paliperidone palmitate)

Status:	Approved
Date:	6 December 2019
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-ERI-201245942, 1.0

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

LIST OF ATTACHMENTS **90**

ABBREVIATIONS

ANCOVA	analysis of covariance
ANOVA	analysis of variance
bpm	Beats per minute
CMH	Cochran-Mantel-Haenszel
CRDPSS	Clinician-Rated Dimensions of Psychosis Symptom Severity scale (DSM-5)
CRF	case report form (paper or electronic as appropriate for this study)
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders (4th edition)
DSM-5	Diagnostic and Statistical Manual of Mental Disorders (5th edition)
ECG	electrocardiogram
eDC	electronic data capture
ER	extended release
ESRS-A	Extrapyramidal Symptom Rating Scale-Abbreviated
GEE	Generalized Estimation Equations
GCP	Good Clinical Practice
ICF	informed consent form
ICM	intracortical myelin
IM	intramuscular
ISST-Plus	InterSePT Scale for Suicidal Thinking-Plus
ITT	intent-to-treat
IVRS	interactive voice response system
IWRS	interactive web response system
LAI	long-acting injectable
LOCF	last-observation-carried-forward
LS	least squares
MATRICS	Measurement and Treatment Research to Improve Cognition in Schizophrenia
MCCB	MATRICS Consensus Cognitive Battery
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model repeated measures
MRI	magnetic resonance imaging
MSQ	Medication Satisfaction Questionnaire
OAP	oral antipsychotic
PPPP	paliperidone palmitate treatment sequence (PP1M followed by PP3M)
PP1M	paliperidone palmitate 1-month injection
PP3M	paliperidone palmitate 3-month injection
PSP	Personal and Social Performance scale
R076477	Sponsor's designation for study drug (paliperidone)
TEAE	treatment-emergent adverse event

1. INTRODUCTION

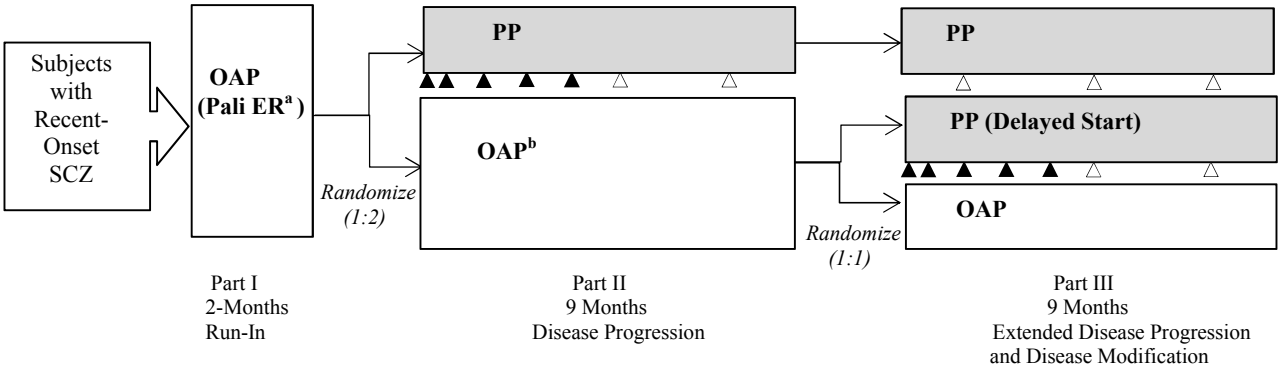
This statistical analysis plan (SAP) describes study objectives, study design, and the sample size calculation based on protocol R092670-SCH-3013 (DREaM). This SAP also contains planned analyses, including definitions of analysis populations, derived variables, and statistical methods for the analysis of efficacy and safety data.

1.1. Trial Objectives

The study includes 3 treatment phases: a 2-month, open-label, flexible-dose, oral run-in phase (Part I), and 2 sequential, 9-month, matched-control, randomized, open-label, active-controlled, flexible-dose treatment phases (Part II, referred to as the 'Disease Progression Phase' and Part III, referred to as the 'Extended Disease Progression and Disease Modification Phase'). Part II and Part III each have their own objectives.

An overview of the study design is provided in Figure 1.

Figure 1: Overview of the Study Design



▲ = PP1M injection
△ = PP3M injection

[a] All subjects will be started on paliperidone ER or oral risperidone. Subjects who find paliperidone ER/oral risperidone intolerable will be withdrawn from the study; subjects who tolerate paliperidone ER/oral risperidone but find it inadequately efficacious may be switched to another protocol-specified OAP at the discretion of the investigator.

[b] Subjects randomized to the OAP treatment group will continue their OAP treatment (ie, paliperidone ER, or other OAP) from Part I.

Note: For optimal management of symptoms/tolerability, subjects receiving PP3M (during either Part II or Part III) may go back to treatment with PP1M (monthly injections of 78, 117, 156 or 234 mg, flexibly dosed) for further dose adjustment or for the duration of the trial with the approval of the medical monitor.

OAP=oral antipsychotic; Pali ER=paliperidone extended-release; PP=paliperidone palmitate; PP1M=paliperidone palmitate 1-month injection; PP3M=paliperidone palmitate 3-month injection; SCZ=schizophrenia or schizophreniform disorder.

Part I Objectives: Run-In

The run-in phase is used to prospectively characterize the population enrolled and match subjects with similar characteristics so that, despite expected dropouts, subjects can be validly re-randomized when the delayed-start treatment is initiated. This 2-month run-in phase will also be used to identify subjects who have a propensity to discontinue.

Part II Objectives: Disease Progression

Subjects who complete the Part I oral run-in phase will enter Part II and will be randomized in a 1:2 ratio to either start PP (paliperidone palmitate) treatment (ie, PP1M followed by PP3M) or to continue their OAP (oral antipsychotic) treatment from Part I. It is expected that most subjects will be receiving paliperidone ER at the start of Part II, but some subjects may be on an alternative OAP. The Part II treatment duration is 9 months.

Primary Objective

- To compare the effectiveness of PP versus OAP treatment (ie, paliperidone ER, oral risperidone, or other OAP) in delaying time to first treatment failure over 9 months' treatment in subjects with recent-onset schizophrenia or schizophreniform disorder.

Key Secondary Objectives:

- To evaluate changes in cognition as measured by the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB) composite score following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To evaluate changes in functioning as measured by the Personal and Social Performance scale (PSP) following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To evaluate changes in brain intracortical myelin (ICM) volume as measured by inversion recovery (IR) and spin echo (SE) magnetic resonance imaging (MRI) in the frontal lobe following 9 months' treatment with PP compared to 9 months' treatment with OAP.

Secondary Objectives:

- To evaluate changes in cognition as measured by the individual domains of the MCCB (ie, working memory, verbal learning, speed of processing, attention/vigilance, visual learning, reasoning and problem solving, and social cognition) following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To evaluate changes in illness severity as measured by the Clinical Global Impression-Severity scale (CGI-S) following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To evaluate changes in severity of psychotic symptoms, as measured by the 8 items of the Clinician-Rated Dimensions of Psychosis Symptom Severity scale (CRDPSS) (ie, delusions, hallucinations, disorganized speech, abnormal psychomotor behavior, negative symptoms [restricted emotional expression or avolition], impaired cognition, depression, and mania) following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To evaluate changes in medication satisfaction as measured by the medication satisfaction questionnaire (MSQ) [patient-reported outcome] following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To assess overall safety of PP.

Exploratory Objectives:

- To assess changes in resting state functioning MRI (fMRI), and changes in cortical thickness, gray matter volume, white matter volume, ventricular volume, intrasulcal CSF volume, and subcortical myelin integrity, as measured by MRI following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To explore the overall healthcare resource utilization use as measured by the Resource Utilization Questionnaire (RUQ) following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To explore differences in satisfaction with goal setting following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- To explore quantitative assessments of daily activities following 9 months' treatment with PP compared to 9 months' treatment with OAP.
- Additional objectives in Part II will include:
    - To evaluate changes in MCCB composite and domain scores by number of treatment failures.
    - To evaluate changes in PSP total and domain scores by number of treatment failures.
    - To examine correlations in change scores between MCCB and PSP total scores; and examine relationships on change scores among ICM volumes, MCCB scores, and PSP Total score.

Part III Objectives: Extended Disease Progression and Disease Modification

Subjects who complete Part II will be eligible to enter Part III. At the start of Part III, subjects treated with OAP during Part II will be re-randomized in a 1:1 ratio to either continue treatment with OAP (OAP-OAP group) or to switch to PP (OAP-PP group). Subjects treated with PP during Part II will continue the same treatment (PP-PP group). The Part III treatment duration is 9 months.

The *Extended Disease Progression* objectives will focus on comparisons between the OAP-OAP and PP-PP groups; the *Disease Modification* objectives will focus on comparisons between the PP-PP and OAP-PP groups (ie, subjects who started treatment with PP early vs. subjects who started PP treatment 9 months later).

Extended Disease Progression Objectives

Primary Objective

- To compare the effectiveness of PP-PP versus OAP-OAP treatment in delaying time to first treatment failure over 18 months' treatment in subjects with recent-onset schizophrenia or schizophreniform disorder.

Key Secondary Objectives:

- To evaluate changes in cognition as measured by the MCCB composite score following 18 months' treatment with PP compared to 18 months' treatment with OAP in subjects with recent-onset schizophrenia or schizophreniform disorder.
- To evaluate changes in functioning as measured by the PSP following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To evaluate changes in brain ICM volume in the frontal lobe following 18 months' treatment with PP compared to 18 months' treatment with OAP.

Secondary Objectives:

- To evaluate changes in cognition as measured by the individual domains of the MCCB following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To evaluate changes CGI-S following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To evaluate changes in severity of psychotic symptoms, as measured by the 8 items of the CRDPSS following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To evaluate changes in medication satisfaction as measured by the MSQ (patient-reported outcome) following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To evaluate differences in time to first treatment failure and subsequent treatment failures over 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To assess overall safety of PP.

Exploratory Objectives:

- To assess changes in resting state fMRI, and changes in cortical thickness, gray matter volume, white matter volume, ventricular volume, intrasulcal CSF volume, and subcortical myelin integrity, as measured by MRI following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To explore the overall healthcare resource utilization use as measured by the RUQ following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To explore differences in satisfaction with goal setting following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- To explore quantitative assessments of daily activities following 18 months' treatment with PP compared to 18 months' treatment with OAP.
- Additional objectives in EDP phase will include:
    - To evaluate changes in MCCB composite and domain scores by number of treatment failures.
    - To evaluate changes in PSP total and domain scores by number of treatment failures.

- To examine correlations in change scores between MCCB and PSP total scores; and examine relationships on change scores among ICM volumes, MCCB scores, and PSP Total score.

Disease Modification Objectives:

Primary Objective

- Using a delayed-start approach, to compare changes in cognition as measured by the MCCB composite score following 9 months' additional PP treatment in subjects originally randomized to PP (PP-PP group) compared to 9 months' delayed-start PP treatment in subjects originally randomized to OAP treatment (OAP-PP group).

Key Secondary Objectives:

- Using a delayed-start approach, to compare changes in functioning as measured by the PSP following 9 months' additional PP treatment in subjects originally randomized to PP (PP-PP group) compared to 9 months' delayed-start PP treatment in subjects originally randomized to OAP treatment (OAP-PP group).
- Using a delayed-start approach, to compare changes in brain ICM volume in the frontal lobe following 9 months' additional PP treatment in subjects originally randomized to PP (PP-PP group) compared to 9 months' delayed-start PP treatment in subjects originally randomized to OAP treatment (OAP-PP group).

Overall Primary Hypothesis

The overall primary hypothesis to be tested in this study is that 9 months' treatment with PP is superior to 9 months' treatment with OAP in delaying time to first treatment failure in subjects with recent-onset schizophrenia or schizophreniform disorder. The primary efficacy null hypothesis is that there is no difference in the distribution of time to first treatment failure in Part II between the PP and OAP treatment groups.

1.2. Trial Design

This is a prospective, matched-control, randomized, open-label, active-controlled, flexible-dose, multi-center study designed to compare the effectiveness of PP versus OAP in delaying time to first treatment failure in subjects with recent-onset schizophrenia or schizophreniform disorder. The study will also evaluate whether long-acting injectable (LAI) treatment with PP can slow disease progression and possibly modify disease course compared to OAP medications, by tracking changes in cognition, functioning, and frontal lobe ICM volume. Other efficacy, safety, and exploratory endpoints will also be assessed.

Approximately 275 men and women between the age of 18 and 35 years, who have a diagnosis of schizophrenia or schizophreniform disorder who experienced their first psychotic episode within 2 years of study entry will be enrolled. The total study duration for each subject will be approximately 86 weeks, including a screening phase (up to 4 weeks), a 2-month oral run-in phase (Part I), and two 9-month treatment phases (Part II and Part III). The study periods are briefly described below:

- **Screening (up to 4 weeks; Day -28 to -1):** Subjects who provide written informed consent will undergo the screening procedures, including a review of the study entry criteria. Any prestudy OAP will be tapered off and must be discontinued by Week 5 of Part I. Tapering of the previous OAP can start at the beginning of the screening period. Tapering and discontinuation should be managed by the investigator, as clinically appropriate.
- **Part I, Oral Run-In Phase (2 months):** After completing the screening period, subjects meeting the inclusion and exclusion criteria will be entered into Part I, a 2-month oral run-in phase. All subjects will initially receive flexible dosing with oral paliperidone ER (1.5-12 mg/day) or oral risperidone (1-6 mg). Treatment with oral paliperidone ER or oral risperidone will enable investigators to establish tolerability prior to randomization to the LAI formulation in Part II. Subjects who find oral paliperidone ER or oral risperidone intolerable will be withdrawn from the study. Subjects who tolerate oral paliperidone ER or oral risperidone but find it inadequately efficacious after treatment for an adequate duration at an adequate dosage, may be switched to another protocol-specified OAP at the discretion of the investigator. Any of the following 7 OAPs are permitted: aripiprazole, haloperidol, olanzapine, paliperidone ER, perphenazine, quetiapine, and risperidone. The following demographic and baseline characteristics and clinical data will be collected during Part I to be used for matching during randomization into Part II and Part III: age, gender, race, prior antipsychotic exposure, substance use history, MCCB composite score, and PSP total score.
- **Part II, Disease Progression Phase (9 months):** Subjects who complete Part I will be eligible to enter Part II. On Day 1 of Part II (Day 57 of Part I), subjects will be randomized in a 1:2 ratio to open-label treatment with either PP or to continued OAP treatment for 9 months. Dynamic central randomization will be performed, based on matching criteria determined in Part I. It is estimated that approximately 225 subjects will be randomized in Part II, ie, approximately 75 subjects will be randomized to the PP treatment group and 150 subjects to the OAP treatment group.
  - Subjects randomly assigned to the OAP treatment group will continue their OAP treatment from Part I. It is expected that most subjects will be receiving oral paliperidone ER or oral risperidone, but some subjects may be receiving an alternative OAP. Investigators are encouraged to continue the original OAP (ie, the OAP at Part II baseline) as monotherapy throughout the remainder of the study but, if clinically indicated, a switch to an alternative OAP or add-on of an additional OAP is permitted after the first randomization visit. Switching or add-on of another OAP due to inadequate efficacy, tolerability, or safety will be assessed as a treatment failure. Subjects with treatment failure will continue participation in the study. Multiple switches (ie, treatment failures) are permitted during the study. The same 7 OAPs identified in Part I are allowed (ie, aripiprazole, haloperidol, olanzapine, paliperidone ER, perphenazine, quetiapine, and risperidone).
  - Subjects randomly assigned to the PP treatment group will receive 5 doses of PP1M followed by PP3M once every 12 weeks. PP1M initiation dosing (first injection 234 mg [150 mg eq.] on Day 1 of Part II and second injection 156 mg [100 mg eq.] on Day 8, both in the deltoid muscle) followed by 3 injections of flexible doses of PP1M (78-234 mg [50-150mg eq.]) on Days 36, 64, and 92, either in the deltoid or gluteal muscle. On Day 120, subjects will receive an injection of PP3M (using a 3.5-fold multiple of the PP1M dose received on Day 92) either in the deltoid or gluteal muscle. On Day 204, subjects will receive a flexible dose of PP3M (273-819 mg [175-525 mg

eq.]) either in the deltoid or gluteal muscle. Investigators are encouraged to use PP1M and PP3M as antipsychotic monotherapy and to adjust the injection dose for management of symptoms/tolerability. If required, supplemental oral paliperidone ER (up to 6 mg/day) or oral risperidone (up to 3 mg/day) may be given. It should be noted that adding oral paliperidone/oral risperidone will not be considered a treatment failure unless supplemental paliperidone ER/oral risperidone is given for a combined total of more than 84 days or if paliperidone ER doses exceed 6 mg/day or oral risperidone doses exceed 3 mg/day (see further details in the Dosage and Administration Section). Adding any other antipsychotic will also be considered a treatment failure. Subjects with treatment failures will continue the study unless the PP injection is discontinued permanently.

- **Part III, Extended Disease Progression and Disease Modification Phase (9 months):** Subjects who complete Part II will be entered into Part III. On Day 1 of Part III (Day 260 of Part II), subjects in the OAP treatment arm will be re-randomized in 1:1 ratio to continued treatment with their OAP (OAP-OAP group) or to PP (OAP-PP or 'Delayed-start PP' arm). The Delayed-start PP arm will receive PP1M and PP3M treatment as described in Part II (ie, 5 doses of PP1M followed by PP3M once every 12 weeks). Subjects previously assigned to treatment with PP in Part II will continue in that treatment group with PP3M injections every 12 weeks for a total of 3 injections (PP-PP group). Randomization will be based on matched criteria identified in Part I. Subjects will be followed for an additional 9 months.

Subjects will be assessed for treatment failure at all visits during Part II and Part III. Other efficacy assessments (MCCB, PSP, CGI-S, CRDPSS, and MSQ) will be performed as specified in the Time and Events Schedule. Safety will be monitored through evaluation of AEs, clinical laboratory parameters, vital signs, body weight, ESRS, ISST-Plus, and physical examination findings. Resource use (measured using the RUQ), goal setting experience, and quantitative assessment of daily activities will be assessed as exploratory endpoints. A pharmacogenomic blood sample will be collected from subjects who consent separately to this component of the study (where local regulations permit). Subject participation in pharmacogenomic research is optional.

Approximately half of the enrolled subjects will also undergo brain MRI scans for assessment of ICM volume and other exploratory MRI endpoints. In addition, approximately 20 healthy control subjects (comparable in age, sex, and race to the subjects with schizophrenia/schizophreniform disorder undergoing MRI scans) will be identified at each MRI center and followed as controls for the MRI machine calibration for the duration of the study without treatment.

Note that subjects who experience a treatment failure and do not withdraw consent or meet the criteria for withdrawal from the study will continue in the study and be followed through to the end of the study.

Subjects who miss scheduled injections or visits are allowed to re-enter the same assigned treatment group of the study.

Treatment Failure Criteria

Subjects will be assessed at each visit during Part II and Part III for the occurrence of treatment failure. Treatment failure is defined as any of the following: 1) Psychiatric hospitalization due to worsening symptoms (including Emergency Room visits ≥23 hours, and not including hospitalization due to social reasons); 2) Any deliberate self-injury, suicidal ideation or behavior, homicidal ideation or violent behavior that is clinically significant and needs immediate intervention as determined by the study physician; 3) New arrest/incarceration (not related to probation or existing warrant); 4) Discontinuation of antipsychotic treatment due to inadequate efficacy as determined by the study physician; 5) Discontinuation of antipsychotic treatment due to safety or tolerability as determined by the study physician; 6) Treatment supplementation with another antipsychotic due to inadequate efficacy as determined by the study physician (note: use of oral paliperidone ER or oral risperidone in the PP treatment group will not be considered a treatment failure unless supplemental treatment with oral paliperidone ER or oral rispedone is longer than 84 days or if paliperidone ER doses exceed 6 mg/day or oral risperidone doses exceed 3 mg/day); 7) Increase in the level of psychiatric services (such as from office visit to day hospitalization) in order to prevent imminent psychiatric hospitalization as determined by the study physician.

Any changes in antipsychotic medications (switching, discontinuation, or add-on) must be evaluated against the treatment failure criteria. If any of these changes do not meet the treatment failure criteria, these must be documented and recorded in the eDC.

Subject Population

Subjects with schizophrenia or schizophreniform disorder: Approximately 275 subjects who meet all of the inclusion and none of the exclusion criteria will be enrolled in the study. Refer to the main text for a complete list of inclusion and exclusion criteria.

The major inclusion criteria include the following: men and women aged 18 to 35 years, inclusive, current diagnosis of schizophrenia (295.90) or schizophreniform disorder (295.40) as defined by DSM-5 and confirmed by the Structured Clinical Interview for DSM-5 Disorders (SCID) with a first psychotic episode within the last 24 months, and requiring treatment with an antipsychotic medication or a change in antipsychotic medication due to lack of efficacy, tolerability, safety issues, or investigator/subject preference.

The major exclusion criteria include: positive urine drug screen test for cocaine, amphetamines, opiates, or PCP at screening; current DSM-5 diagnosis of dissociative disorder, bipolar disorder, major depressive disorder, schizoaffective disorder, autistic disorder, or intellectual disabilities; and meets the DSM-5 definition of moderate or severe substance use disorder (except for nicotine) within 2 months prior to screening.

Healthy Control Subjects: Each MRI site will identify and enroll approximately 20 healthy control subjects who will undergo MRI assessments only. These healthy control subjects should be comparable in age, sex, and race to the cohort of subjects with schizophrenia/schizophrenia

disorder undergoing MRI scans. Refer to the main text for a complete list of inclusion and exclusion criteria.

Major exclusion criteria are: evidence of a known psychiatric disorder, neurological disorder (eg, epilepsy) or significant head injury; first degree relative who has schizophrenia, schizophreniform, schizoaffective, or bipolar disorder; or meets the DSM-5 definition of moderate or severe substance use disorder (except for nicotine) within 2 months prior to screening. Subjects who are unable to undergo MRI scan for any reason, including because of body size (unable or difficult to fit in MRI instrument) or MRI contraindicated due to presence of metallic objects (pacemaker, etc.) will also be excluded.

Dosage and Administration

Part I, Oral Run-In Phase

During Part I, all subjects will initially be treated with paliperidone ER (1.5 to 12 mg/day) or oral risperidone (1-6 mg/day). Any prestudy OAP other than risperidone or oral paliperidone ER will be tapered off and must be discontinued by Week 5 of Part I. Subjects already being treated with oral risperidone or oral paliperidone ER should be continued at the dose deemed to be most appropriate by the investigator). Adjustment of the dosage will be done at the investigator's discretion, based on the individual subject's clinical response to and tolerability of the study drug. To be eligible for randomization in Part II, subjects must be able to tolerate a minimum dose of 3 mg of oral paliperidone or 2 mg of oral risperidone for at least 2 weeks prior to entry to Part II. Subjects who find either oral paliperidone ER or oral risperidone intolerable will be withdrawn from the study. Subjects who tolerate oral paliperidone ER/oral risperidone but find it inadequately efficacious after treatment for an adequate duration at an adequate dosage (per clinical judgment), may be switched to another protocol-specified OAP at the discretion of the investigator. The following 7 OAPs are permitted: aripiprazole, haloperidol, olanzapine, oral paliperidone ER, perphenazine, quetiapine, and risperidone.

Part II and Part III, Active-Controlled Treatment Phases

Paliperidone Palmitate (PP1M/PP3M)

All drug injections must be administered by an individual who has received appropriate medical training to administer an IM injection.

Subjects who are randomly assigned to the PP treatment group at the start of Part II will discontinue their OAP treatment from Part I and will be started on PP1M. Subjects will be subsequently switched to PP3M following 5 injections of PP1M. A transition period of a maximum of 5 weeks will be allowed for the previous OAP.

PP1M will be administered IM once-monthly, after the first 2 injections that are given one week apart (Day 1 and Day 8). The first two doses will be administered through a deltoid injection on alternating arms. The subsequent injections can be given either in the deltoid or the gluteal muscle. The first dose of PP1M will be 234 mg given in the deltoid muscle at Day 1 of the treatment phase. The second dose of PP1M will be 156 mg given in the deltoid muscle at Day 8 of the treatment phase. Subsequent doses of PP1M on Days 36, 64, and 92 will be given every 28

(±7) days in either the deltoid or gluteal muscle. The investigator may select from 78, 117, 156, or 234 mg, according to the subjects' clinical needs. Subjects will continue to return to the study site every 4 weeks for injections and for study evaluations.

At the Day 120 visit (±7 days), subjects will start PP3M treatment. The initial PP3M dose will be calculated as 3.5 fold multiple of the final PP1M dose administered on Day 92. Subjects will receive PP3M injections once every 12 weeks (±14 days). Investigators will be permitted to flexibly adjust the dose of PP3M as clinically necessary with the dose options for PP3M being 273, 410, 546, or 819 mg. Injections of PP3M may be administered in either the deltoid muscle or the upper outer portion of the gluteal muscle. The side of each injection (left or right) should be alternated and recorded.

Supplemental OAP Use in the PP Treatment Group: Investigators are strongly encouraged to use PP1M and PP3M as antipsychotic monotherapy and to adjust the injection dose for management of symptoms/tolerability. Subjects in the PP treatment group who are tolerating the medication but experience symptom exacerbation during the study will be allowed to have supplemental antipsychotic medication, ie, oral paliperidone ER (up to 6 mg/day or oral risperidone (up to 3 mg/day, for no longer than a total of 84 days during the total PP treatment period. If the supplemental treatment with oral paliperidone ER/oral risperidone is longer than 84 days or if oral paliperidone ER doses exceed 6 mg/day or oral risperidone doses exceed 3 mg/day, the subject will be considered a treatment failure. Adding any other antipsychotic will also be considered a treatment failure.

A switch to an alternative antipsychotic is not permitted in the PP treatment group. If it is deemed clinically necessary to stop PP1M or PP3M in a subject assigned to PP treatment, the subject will be withdrawn from the study and discontinuation of PP will be recorded as a treatment failure.

Oral Antipsychotic Treatment

Subjects randomly assigned to the OAP treatment group at the start of Part II will continue their OAP treatment from Part I. It is expected that most subjects will be receiving oral paliperidone ER or oral risperidone, but some subjects may be receiving an alternative OAP at entry into Part II.

Investigators are encouraged to continue the same OAP (ie, the OAP at Part II baseline) as monotherapy throughout the remainder of the study, and to adjust the OAP dose for management of symptoms/tolerability at any time. After the initial randomization visit, a switch to a different OAP or the addition of another OAP is allowed in the OAP treatment arm if clinically indicated; however, switching or add-on of another OAP due to inadequate efficacy, tolerability, or safety will be assessed as a treatment failure. The same 7 OAPs specified for Part I (aripiprazole, haloperidol, olanzapine, paliperidone ER, perphenazine, quetiapine, and risperidone) will be permitted. Multiple switches to other protocol-specified OAPs will be allowed during the study. Any change in antipsychotic medication (switching, discontinuation, or add-on) must be evaluated against the treatment failure criteria.

Administration of PP or an alternative LAI antipsychotic is prohibited in subjects assigned to OAP treatment. If an LAI agent is deemed clinically necessary for subjects assigned to the OAP group, their data will be censored as a treatment failure and they will be discontinued from the study.

Subjects should generally be treated within the approved label for all OAPs. Any exceptions should first be discussed with the Medical Monitor.

1.3. Statistical Hypotheses for Trial Objectives

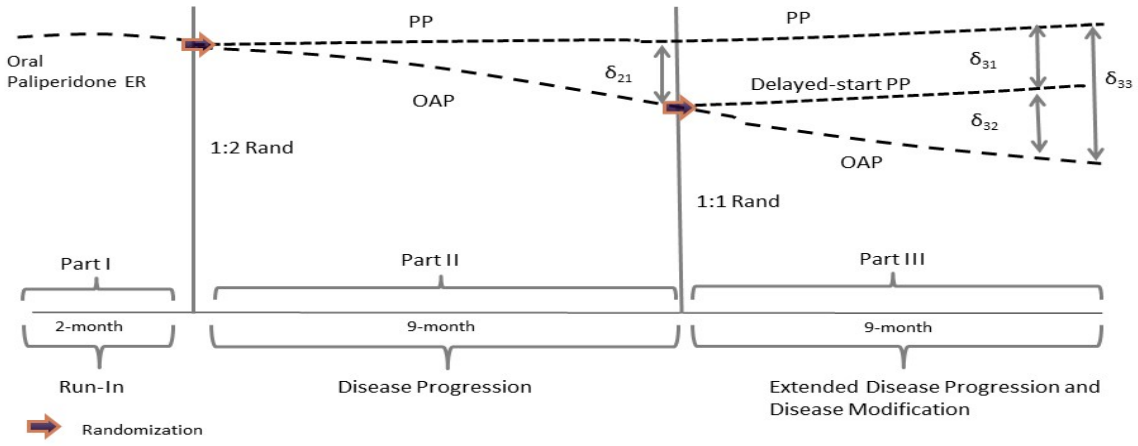
The overall primary hypothesis to be tested in this study is that 9 months' treatment with PP is superior to 9 months' OAP treatment in delaying time to treatment failure in subjects with recent-onset schizophrenia or schizophreniform disorder. This will be assessed at the end of Part II.

Other hypotheses to be evaluated in this study will assess whether LAI treatment with PP can slow down disease progression and possibly modify disease course in recent-onset subjects compared to OAP medications. This will be assessed by tracking changes in cognition (MCCB composite score), functioning (PSP), and brain imaging assessments (ICM volume).

Hypothetical outcomes of the DREaM study are shown schematically in Figure 2. Panels A and B differ with respect to the size of δ_{31} (i.e., the Part 3 difference between delayed-start and patients on PP throughout the trial). Panel A is titled, "Treatment Effect on Disease Progression and Modification"; Panel B is titled, "Delayed-Start Treatment Effect". Numerically higher scores are assumed to correspond to better outcomes. Based on prior experience and the likelihood that patients randomized to OAP will be less compliant with taking their medications, we expect that in Part 2 patients randomly assigned to OAP's will experience a greater degree of disease progression than those receiving PP. In Part 3, if PP has disease modifying effects, the two PP groups (immediate start and delayed-start) will have similar slopes and will be parallel, such that differences that develop after an early start are not overcome and the between groups measures of disease progression do not catch up (Panel A). If early introduction of PP does not have disease-modifying effects, the size of δ_{31} will significantly decrease compared to δ_{21} (the difference between treatment groups at the start of Part 3) (Panel B). The delayed-start treatment benefit is measured by the size of the difference between δ_{31} and δ_{21}.

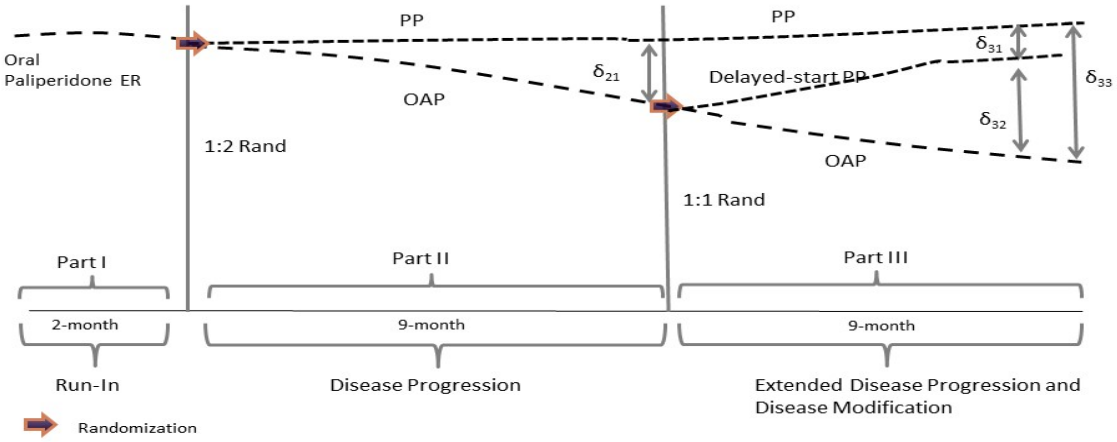
Figure 2: Hypothetical outcomes of DREaM, assuming higher scores correspond to better outcomes

A. Treatment Effect on Disease Progression and Modification



δ_{21}:Treatment effect on disease progression; δ_{31}:Lead treatment effect; δ_{32}:Delayed-start treatment effect on disease progression; δ_{33}:Overall effect of treatment.

B. Delayed-Start Treatment Effect



δ_{21}:Treatment effect on disease progression; δ_{31}:Lead treatment effect; δ_{32}:Delayed-start treatment effect on disease progression; δ_{33}:Overall effect of treatment.

1.3.1. Part II: Disease Progression

The overall primary hypothesis in this study is that 9 months' treatment of subjects with recent-onset schizophrenia or schizophreniform disorder with PP is superior to 9 months' treatment with OAP in delaying time to first treatment failure. The primary efficacy null hypothesis is that there is no difference in the distribution of time to first treatment failure in Part II between the PP and OAP treatment groups.

The key secondary hypotheses to be tested in Part II of this study are to demonstrate that 9 months' treatment of subjects with recent-onset schizophrenia or schizophreniform disorder with PP is superior to 9 months' treatment with OAP in:

- improving or maintaining cognition (as measured by the change in MCCB composite score from baseline)
  - *The key secondary efficacy null hypothesis is that there is no difference in mean change from Part II baseline to the Part II end point in the MCCB composite score between PP and OAP*
- maintaining functioning (as measured by time to 7-point worsening in the PSP total score from baseline)
  - *The key secondary efficacy null hypothesis is that there is no difference in time to 7-point worsening in PSP total score in Part II between PP and OAP*
- increasing or preserving brain ICM volume of the frontal lobe as compared to baseline
  - *The key secondary efficacy null hypothesis is that there is no difference in mean change in ICM volume of the frontal lobe from Part II baseline to the Part II end point between PP and OAP*

1.3.2. Part III: Extended Disease Progression and Disease Modification

1.3.2.1. Extended Disease Progression

Let it be assumed that the quantity δ_{33} in **Error! Reference source not found.** represents the cumulative effect on disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment. The Extended Disease Progression hypotheses to be tested in Part III of this study are to demonstrate that 18 months' treatment of subjects with recent-onset schizophrenia or schizophreniform disorder with PP is superior to 18 months' treatment with OAP in:

- the overall primary hypothesis to be tested in this phase of the study is that 18 months' treatment with PP-PP is superior to 18 months' OAP-OAP treatment in delaying time to treatment failure in subjects with recent-onset schizophrenia or schizophreniform disorder.
- improving or maintaining cognition (as measured by the change in MCCB composite score from baseline)
  - *The key secondary efficacy null hypothesis is that there is no difference in mean change from Part II baseline to the Part III end point in the MCCB composite score between PP and OAP. In other words, H_0: $\delta_{33} = 0$ for MCCB composite score.*
- maintaining functioning (as measured by time to 7-point worsening in the PSP total score from baseline)
  - *The key secondary efficacy null hypothesis is that there is no difference in time to 7-point worsening in PSP total score from Part II baseline to the Part III end point*

- increasing or preserving brain ICM volume of the frontal lobe as compared to baseline

  – *The key secondary efficacy null hypothesis is that there is no difference in mean change in ICM volume of the frontal lobe from Part II baseline to the Part III end point between PP and OAP*

1.3.2.2. Disease Modification

In order to demonstrate disease modification, the following results are required:

- At the end point of Part II: Subjects treated with PP for 9 months (early-start group) must demonstrate better outcomes on MCCB than those treated with OAP for 9 months. Similar results for changes in functioning (PSP) and changes in brain anatomy (ICM) will be used to support this finding (ie, treatment effect on disease progression).

- At the end point of Part III: Subjects from Part II who have been treated with PP for 18 months (early-start group) must continue to show better outcomes on MCCB compared with subjects treated with OAP for 18 months (a differential treatment effect is still evident), and compared with subjects treated with OAP for 9 months followed by PP for 9 months (the lead treatment effect remains significant after 9 months); ie, that late initiation of treatment with PP does not allow for achievement of the same level of cognition after 9 months. [(See corresponding null hypothesis below (H_0: $\delta_{31} - \gamma . \delta_{21} \leq 0$)]. Similar results for changes in PSP and ICM will be used to support this finding for disease modification.

For each of the endpoints (MCCB, PSP, and ICM), the following hypotheses will be tested as a function of estimated changes in the Part II endpoint using the observed scores. Let δ_{21} denote the estimated mean differences between treatment groups at the end of Part II for a given endpoint, treatment effect on disease progression. Below, the parameter $\mu_{PP\text{-}PP}$ corresponds to the mean score for a given endpoint following 9 months' of additional PP3M treatment in subjects originally randomized to PP and $\mu_{OAP\text{-}PP}$ corresponds to mean score for 9 months' of delayed-start PP treatment in subjects originally randomized to OAP treatment. Let δ_{31} denote the estimated difference between $\mu_{PP\text{-}PP}$ and $\mu_{OAP\text{-}PP}$, lead treatment effect at the end of Part III. The corresponding null hypotheses for a given endpoint will be listed as:

> H_0: $\delta_{31} - \gamma . \delta_{21} \leq 0$, where γ is a fixed number between 0 and 1 indicating the ratio of treatment effect in Part III compared to the Part II observed differences.

The quantity γ will be determined after Part II database lock and prior to finalization of Part III Statistical Analysis Plan. For γ, the quantities 0.75 and 0.50 will be considered principally. The γ could be different for a different endpoint. The quantity γ is called conditional noninferiority margin.

In addition, for each of the endpoints, δ_{32} (delayed-start treatment effect on disease progression) will be tested in Part III using Part III baseline scores. Below, the parameter $\mu_{OAP\text{-}PP}$ corresponds to mean score for 9 months' of delayed-start PP treatment in subjects originally randomized to OAP treatment and $\mu_{OAP\text{-}OAP}$ corresponds to the mean score for a given endpoint following

9 months' of additional OAP treatment in subjects originally randomized to OAP. The corresponding null hypotheses for a given endpoint will be listed as:

$$H_0: \delta_{32} = \mu_{OAP\text{-}PP} - \mu_{OAP\text{-}OAP} = 0.$$

The quantity δ_{33} represents the cumulative effect on disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment. At the conclusion of Part III, the database will again be locked and data analyzed to assess response in variables δ_{31}, δ_{32}, and δ_{33} (overall effect of treatment). Positive results (δ_{33} is significant [a differential treatment effect is still evident] **and** δ_{31} is significant [the lead effect remains significant after 9 months]) will be used to support a conclusion for disease modification.

1.4. Sample Size Determination

The overall primary efficacy null hypothesis is that there is no difference in the distribution of time to treatment failure between the PP and OAP in the treatment of subjects with recent-onset schizophrenia or schizophreniform disorder in Part II. Treatment differences will be compared using a log-rank test. It is assumed that treatment failure rate in Part II is approximately 40% for the OAP group and 20% for the PP group at month 9 with a corresponding hazard ratio of 0.44. It is also assumed that the hazard rates of treatment failure for the two groups are proportional. Additional assumptions made to calculate the expected number of subjects that need to be randomized to obtain the required number of treatment failures are:

- In both treatment groups, 10% of the randomized subjects will be lost-to-follow-up.
- Uniform accrual rate during the 15 month accrual period.

With these assumptions, it is planned to randomize 225 (75 in PP and 150 in OAP group) subjects in a 1:2 ratio to receive either PP or OAP to obtain at least 62 treatment failures to show that PP is significantly different from OAP at the 2 sided significance level of 0.05, with 80% power to detect a hazard ratio of 0.44 using a log rank test.

Blinded surveillance of the total number of events in Part II will be performed during the study to assess the appropriateness of the assumptions. The number of subjects enrolled and the number of subjects who discontinue before entering Part II will be closely monitored.

Assuming 20% attrition rate during the 2-month Run-in period (Part I), the total number of subjects to be enrolled in Part I will be approximately 275.

One of the key secondary efficacy null hypotheses in Part II is that there is no difference in mean change from Part II baseline to the Part II end point in the MCCB composite score between PP and OAP in the treatment of subjects with recent-onset schizophrenia or schizophreniform disorder. The corresponding efficacy variable is the change in the MCCB composite score. This variable will be analyzed in Part II using mixed-effects repeated measures analysis of covariance (ANCOVA) model, described later. Similarly, the matching null hypothesis in Part III is that

there is no difference in mean change from Part II baseline to the Part III end point in the MCCB composite score between PP and OAP in the treatment of subjects with recent-onset schizophrenia or schizophreniform disorder.

This is the first study to assess the efficacy of PP on cognition in subjects with recent-onset schizophrenia or schizophreniform disorder. A small study comparing the clinical efficacy of the long-acting injectable formulation of risperidone (RLAI) to the oral formulation of risperidone (oral Ris) in the early course of schizophrenia in UCLA Aftercare Research program showed an effect size of 0.46 between the treatment groups in the MCCB composite score (change from baseline) at Month 6 end point. We consider an effect size of 0.3 and above to be clinically meaningful for cognition in a therapeutic area where there are no known effective treatments.

Assuming an effect size of approximately 0.40 for the difference in the mean change from Part II baseline to Part II end point in the MCCB composite score between PP and OAP, with 1:2 randomization ratio, sample sizes of 75 subjects in the PP group and 150 subjects in the OAP group are necessary to provide 80% power based on a two group t-test with a 0.05 two-sided significance level. This computation is also consistent with the time to treatment failure calculations ignoring correlations among these endpoints.

It is assumed that approximately 20% of subjects who entered Part II will not be transitioned to the Part III (Extended Disease Progression and Disease Modification) period in each treatment group. Approximately, a total of 180 subjects will be available at the Part III baseline. Subjects in the OAP treatment arm will be re-randomized in a 1:1 ratio to continued treatment with their oral treatment or to PP (delayed-start PP arm). Subjects previously assigned to treatment with PP in Part II will continue in that treatment in Part III. Thus, approximately 60 subjects will be available for each treatment group at the Part III baseline. Assuming approximately 5% of subjects will be missing either the Part III baseline or all post-baseline MCCB assessments, a sample size of 57 in each group will have 80% power to detect an effect size of 0.53 between PP and OAP in the difference in the mean change from Part II baseline in the MCCB composite score using a two group t-test with a 0.05 two-sided significance level in Part III. The primary objective in Part III is to compare changes in cognition following 18 months' treatment with PP compared to 18 months' treatment with OAP.

1.5. Randomization and Blinding

Procedures for Randomization and Stratification

There will be 2 randomizations in this study, once at the start of Part II and once at the start of Part III. Randomization will be used to minimize bias in the assignment of subjects to treatment groups, to increase the likelihood that known and unknown subject attributes (eg, demographic and baseline characteristics) are evenly balanced across treatment groups, and to enhance the validity of statistical comparisons across treatment groups.

Matching will be performed using dynamic central randomization. In Part II, subjects will be assigned to receive either PP or OAP in a 1:2 ratio based on an algorithm implemented in the

interactive voice response system (IVRS) or interactive web response system (IWRS) before the study. It is estimated that approximately 225 subjects will be randomized in Part II; therefore approximately 75 subjects will be randomized to the PP treatment group and 150 subjects to the OAP treatment group. Dynamic central randomization minimizes the imbalance in the distribution of the number of subjects across treatment groups within the levels of each individual stratification factor: age, race, gender, duration of previous antipsychotic usage prior to screening, baseline MCCB composite score, baseline PSP score, substance use history, MRI participation, and site. Based on the algorithm, the IVRS/IWRS will assign a unique treatment code, which will dictate the treatment assignment. The requestor must use his or her own user identification and personal identification number when contacting the IVRS/IWRS, and will then give the relevant subject details to uniquely identify the subject. To eliminate the predictability of randomization for the next subject, treatment assignment probabilities will also be utilized.

Subjects who complete Part II will be entered into Part III. On Day 1 of Part III (Day 260 of Part II), subjects in the OAP treatment arm will be re-randomized in a 1:1 ratio to continued treatment with their oral treatment or to PP (Delayed-start PP arm). The Delayed-start PP arm will receive PP1M and PP3M treatment as described in Part II (ie, 5 doses of PP1M followed by PP3M once every 12 weeks). Subjects previously assigned to treatment with PP in Part II will continue in that treatment group with PP3M injections every 12 weeks for a total of 3 injections. Randomization will be based on matched criteria identified in Part I. Similar to the Part II randomization, dynamic central randomization will be implemented for Part III. Stratification factors will include age, race, gender, duration of previous antipsychotic usage prior to screening, baseline PSP score at Part I, baseline MCCB score at Part I, substance abuse history, and site.

Subjects who re-enter the study after missing scheduled injections or visits must stay in the same assigned treatment group.

Blinding

No blinding of investigators or subjects will be used in this study.

Centralized blinded evaluations will be utilized for the assessment of imaging endpoints (eg, ICM volume) to reduce potential bias during data collection and evaluation. For the imaging endpoints, the assessment of the MRI images will be performed by a rater who is blinded to drug-treatment information and to clinical and demographic characteristics of the subjects.

MCCB assessments will be performed by raters at each study site and there is no blinding for the local MCCB raters. All MCCB data collected at each study site will be sent to central raters for review. Assessments will each be scored by central raters who are blinded to treatment information. Central raters will also review the neurocognitive data quality and correct any errors. The central raters will enter the final scores into the MCCB Scoring program to derive T-scores and composites scores and enter the raw and derived MCCB scores to the eDC.

2. GENERAL ANALYSIS DEFINITIONS

2.1. Analysis Phases

There are 4 analysis phases defined in this study, see Time and Event Schedule:

- Screening: Up to 4 weeks, Day -28 to Day -1.
- Part I, Oral Run-In Phase: 2 Months
- Part II, Disease Progression Phase: 9 Months
- Part III, Extended Disease Progression and Disease Modification Phase: 9 Months. Extended Disease Progression part and Disease Modification part differs with respect to duration of examination period.

Each analysis phase has its own analysis phase start and end dates.

TIME AND EVENTS SCHEDULE: SCREENING

Treatment Phase	Screening
Visit:	**1**
Day (Part I)	-28 to -1
Screening/Administrative	
Informed consent	X
Informed consent for optional genetic research samples	X
Inclusion/exclusion criteria	X
Medical and psychiatric history	X
SCID	X
Pre-morbid IQ estimate (WTAR)[a]	X
Urine Drug Screen	X
Informed consent for the subject's designated individual	X
Efficacy Assessments	
MCCB	X
PSP	X
CRDPSS (DSM-5)	X
CGI-S	X
MSQ	X
Safety Assessments	
Physical examination	X
ECG	X
Clinical laboratory tests	X
Urine Pregnancy test	X
ESRS-A	X
ISST-Plus Short form[b]	X
Vital signs (blood pressure, pulse rate) and weight	X
Adverse Events	X
Concomitant medication	X

[a] An alternative test will be defined by the sponsor if the study is conducted in countries other than the United States.

[b] If suicidality is identified (ie, 'yes' is answered to questions 1.0, 2.2, 2.3, or 2.4 of the ISST-Plus Short Form) the full ISST-Plus must be administered in its entirety.

KEY: CGI-S=Clinical Global Impression-Severity scale; CRDPSS=Clinician-Rated Dimensions of Psychosis Symptom Severity; DSM-5=Diagnostic and Statistical Manual of Mental Disorders, 5th edition; ECG=electrocardiogram; ESRS-A=Extrapyramidal Symptom Rating Scale-Abbreviated; IQ=intelligence quotient; ISST-Plus=InterSePT Scale for Suicidal Thinking-Plus; MCCB=MATRICS Consensus Cognitive Battery; MSQ=Medication Satisfaction Questionnaire; PSP=Personal and Social Performance scale; SCID=Structured Clinical Interview for DSM-5 Disorders; UV=unscheduled visit; WTAR=Wechsler Test of Adult Reading

TIME AND EVENTS SCHEDULE: PART I, PART II, AND PART III

Treatment Phase:	Part I[a] (Oral Run-in)					Part II (Disease Progression)										Part III (Extended Disease Progression and Disease Modification)											
Visit:	2	3	4	5	6	7[p]	8	9	10	11	12	13[s]	14	15	16[s]	17[p]	18	19[o]	20	21	22	23[s]	24	25	26[s]	27/EOS[p]	UV[m]
Day (Part I)	1	8	15	29	43	57																					
Day (Part II)						1	8	36	64	92	120	148	176	204	232	260											
Day (Part III)																1	8	36	64	92	120	148	176	204	232	260	
Visit window		±4	±4	±4	±4	±7	±4	±7	±7	±7	±7	±7	±7	±7[n]	±7	±7	±4	±7[n]	±7	±7	±7[n]	±7	±7	±7[n]	±7		
Administrative																											
Inclusion/exclusion criteria	X																										
Urine Drug Screen						X										X									X	
Randomization						X[b]										X[c]											
Contact designated individual[d]				X		X		X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	
Blood sample collection for DNA[r]	X																										
Study Drug Administration/Prescription[e]																											
Administer study drug (PP)[f]	O	O	O	O	O	1M	1M	1M	1M	1M	3M	-	-	3M	-	-	-[s]	3M	-	-	3M	-	-	3M	-	-	
or provide prescription (O)[g]						O	O[s]	O	O	O	O	O	O	O	O	1M	1M	1M	1M	1M	3M	-	-	3M	-	-	
																O	O[s]	O	O	O	O	O	O	O	O	-	
Efficacy Assessments[h]																											
Assessment for treatment failure							X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X[h]
MCCB[i]				X		X				X			X			X				X			X			X	X[h]
PSP				X		X				X			X			X				X			X			X	X[h]
CRDPSS (DSM-5)	X			X		X				X			X			X				X			X			X	X[h]
CGI-S	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X[h]
MSQ	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X[h]
Imaging[h,t]																											
MRI						X				X						X				X						X	X[h]
Exploratory Assessments																											
Goal setting and review[j]	X			X		X		X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	
Assessment of daily activities	X					X				X			X			X				X			X			X	
RUQ	X					X				X			X			X				X			X			X	
Safety Assessments																											
Physical examination																										X	
EC						X										X										X	
Clinical laboratory tests						X										X										X	
Urine Pregnancy test						X										X										X	
ESRS-A						X				X			X			X				X			X			X	
ISST-Plus Short form[k]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Vital signs[l] and weight	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Adverse Events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medication	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

[a] During Part I, all subjects will initially receive flexible doses of paliperidone ER (1.5-12 mg/day) or 1-6 mg/day of oral risperidone. Subjects who find paliperidone ER or oral risperidone intolerable will be withdrawn from the study. Subjects who tolerate paliperidone ER or oral risperidone but find it inadequately efficacious may be switched to another protocol-specified OAP. Any of the following 7 OAPs will be permitted during Part I: aripiprazole, haloperidol, olanzapine, paliperidone ER, perphenazine, quetiapine, and risperidone.

[b] At the start of Part II, subjects will be randomized in a 1:2 ratio to start PP injections or to continue OAP treatment from Part I.

[c] At the start of Part III, subjects who were randomized to OAP in Part II will be re-randomized in a 1:1 ratio to continue OAP from Part II (OAP-OAP group) or switch to PP injections (OAP-PP group). Subjects randomized to PP in Part II will continue the same treatment (PP-PP group).

[d] The subject's designated individual should be contacted to check on the subject's wellbeing, such as subject's general health, common daily activities and progress in personal or health goals. Contact may be by telephone if the designated individual does not accompany the subject at the scheduled visit.

[e] All procedures should be performed prior to study drug administration/prescription, except for MRI, which does not need be performed on the day of a scheduled visit, but should be completed within ±7 days.

[f] Subjects who are randomly assigned to start PP treatment will receive PP1M injections on Day 1, Day 8 (±4 days), Day 36 (±7 days), Day 64 (±7 days), and Day 92 (±7 days). The first dose of PP3M will be given on Day 120 (±7 days) and then every 12 weeks (±14 days) thereafter. If these windows are exceeded, contact the medical monitor. Refer to Section **Error! Reference source not found.** of the Protocol for details on PP dosing as well as re-initiation of PP in cases of missed doses.

[g] Subjects who are randomly assigned to the OAP treatment group will continue their OAP treatment from Part I. Subjects will be provided a voucher to present at a local pharmacy to receive their assigned study drug. Following the initial randomization visit, a switch to a different OAP or the addition of another OAP is allowed in the OAP treatment group if clinically indicated; however, switching or add-on of another OAP due to inadequate efficacy, safety, or tolerability will be assessed as a treatment failure. Refer to Section **Error! Reference source not found.** of the Protocol for definition of Treatment Failure. Any of the following 7 agents will be permitted during Part II and Part III: aripiprazole, haloperidol, olanzapine, paliperidone ER, perphenazine, quetiapine, and risperidone. Multiple switches to alternative OAPs (ie, multiple treatment failures) are allowed during the study.

[h] All efficacy assessments and MRI scans (in the subgroup of subjects undergoing MRI assessment) must be completed at the first occurrence of treatment failure, or as soon as possible, even they are not scheduled to be done for the visit or for the unscheduled visit. At subsequent treatment failures, PSP, CGI-S, CRDPSS, and MSQ assessments should be performed.

[i] For every individual subject, MCCB should be administered at the same time of day (within a 2-hour window) for every assessment.

[j] Patient Happiness Assessment and Goal Setting Preparation will be performed at Day 1 Part I only; Patient Goal Setting Documentation will be completed on Day 1 of Part I, II, and III; Assessment of Patient Goal Attainment will be assessed by the subject every month.

[k] If suicidality is identified (ie, 'yes' is answered to questions 1.0, 2.2, 2.3, or 2.4 of the ISST-Plus Short Form) the full ISST-Plus must be administered in its entirety.

[l] Blood pressure and pulse rate.

[m] Unscheduled visits should be performed, as necessary for the following reasons:

- Unscheduled visits that are clinically indicated: Unscheduled visits should be performed as necessary in the judgment of the physician for appropriate clinical care, including reasons of safety or tolerability.
- Unscheduled visits for investigational purposes: Any subject who experiences a protocol-defined treatment failure event after randomization to Part II (Visit 7) should undergo an unscheduled visit. This applies to every treatment failure episode; not just the initial one. If an unscheduled visit can't be performed before the next scheduled protocol visit, applicable unscheduled visit procedures should be performed during the scheduled visit even when they are not normally part of the scheduled visit. Note that subjects who experience a treatment failure event should continue in the study, unless they meet study withdrawal criteria.

[n] At these visits (Visit 15, 19, 22, and 25), the treatment window for PP3M dosing will be ±14 days, except for the first PP3M injection, which will be ±7 days.

[o] In order to synchronize visit dates for all 3 treatment arms, Visit 19 occurs 92 days (instead of 84 days, ie, 12 weeks) after the previous PP3M injection. However, it is still within the ±14-day window.

[p] Note that completion and discontinuation will be recorded in the CRF for each treatment phase.

[r] The pharmacogenomic (DNA) sample should be collected at the specified time point, however if necessary it may be collected at a later time point without constituting a protocol deviation.

[s] These visits can be conducted by telephone.

[t] Approximately half the subjects will undergo MRI scans at selected study sites. MRI assessments are optional to subjects. MRI scans will also be performed in the healthy control group; the schedule for MRI assessments in healthy control subjects will be provided in a separate MRI manual.

KEY: CGI-S=Clinical Global Impression-Severity scale; CRDPSS=Clinician-Rated Dimensions of Psychosis Symptom Severity; ECG=electrocardiogram; EOS=end of study; ESRS-A=Extrapyramidal Symptom Rating Scale-Abbreviated; ISST-Plus= InterSePT Scale for Suicidal Thinking-Plus; MCCB=MATRICS Consensus Cognitive Battery; MRI=magnetic resonance imaging; MSQ=Medication Satisfaction Questionnaire; O=oral antipsychotic; 1M=paliperidone palmitate 1-month formulation; 3M=paliperidone palmitate 3-month formulation; PP=paliperidone palmitate; PSP=Personal and Social Performance scale; RUQ=Resource Use Questionnaire; UV=unscheduled visit

2.1.1. Study Reference Start and End Dates

The trial reference start date will be the Part I Start Date. The trial reference end date will be the date of the last trial-related procedure. Specifically, it will be the maximum of the last study medication date, or last visit date (scheduled or unscheduled), or End-of-Treatment/Early Withdrawal visit date. At each scheduled visit, all the assessments need to be completed before the study drug administration.

2.1.2. Analysis Phase Start and End Dates

Screening Phase

The start date of the screening phase is the informed consent date. The screening phase end date is the Part I start date.

Part I, Oral Run-In Phase

The oral run-in phase will begin on the day (referred to as, 'Part I Start Date') of Visit 2. The oral run-in phase end date (denoted as 'Part I End Date') is the date of Part II randomization (for OAP subjects) or the first PP injection date in Part II for subjects who enter the Part II; otherwise, subjects who withdraw during the Part I, the Part I End Date is the same as the Reference End Date.

Part II, Disease Progression Phase

The Part II will begin on the day (referred to as, 'Part II Start Date') of the first randomization (for OAP subjects) or the first paliperidone palmitate injection date, which is Day 1 in Part II.

For subjects who complete/discontinue from the Part II, the Part II end date (denoted as 'Part II End Date') is the maximum of the date of last visit in Part II. Specifically, for those subjects who enter the Part III and randomized, the end date will be the date of Part III randomization. Those subjects who are assigned to PP in Part II and continue to be in PP, the end date will be the date of Visit 17. For subjects who withdraw during the Part II, the Part II End Date is the same as the Reference End Date.

Part III, Extended Disease Progression and Disease Modification Phase

The Part III will begin on the day (referred to as, 'Part III Start Date') of the second randomization for OAP subjects from Part II. For those subjects who are assigned to PP in Part II, the Part III Start Date will be the earliest Visit 17 date.

For subjects who complete/discontinue from the Part III, the Part III end date (denoted as 'Part III End Date') is the maximum of the date of last visit in Part III. The Part III End Date is the same as the Reference End Date.

The extended disease progression phase will start with Part II Start Date and will continue into Part III. The objective of this phase to examine differences (i.e., δ_{33} for efficacy endpoints) between patients given OAPs vs. those given PP up to 18 months. The end date for the extended disease progression phase will be same as the Part III end date.

2.1.3. Study Day and Relative Day

Study day is calculated relative to the reference start date for the study. Relative day is calculated relative to the analysis phase start date of the analysis phase in which the data are captured. A minus (-) sign indicates days prior to the start of study or prior to the start of the analysis phase.

Study day for an event on or after the start of the study is calculated as:

event date- reference start date+1.

Study day for an event prior to the start of the study is calculated as:

event date-reference start date

Relative day for an event on or after the analysis phase start date is calculated as:

event date – analysis phase start date + 1.

Relative day for an event prior to the analysis phase start date is calculated as:

event date – analysis phase start date.

There is no study day 0 or relative day 0.

2.2. Baseline and End Point

Baseline is defined for each parameter/assessment.

Part I: The 'Baseline Part I' value for the Part I is defined as the last assessment on or before the Visit 2 date. Note that the baseline is defined for each parameter of interest whereas the Part I Start Date is defined at subject level and remains the same for all parameters of interest. The baseline scores may consist of information from the Screening phase.

Part II: The last observation prior to or on the start date of Part II is denoted as, 'Baseline Part II'. It is defined for all efficacy variables as well as for the safety variables. Baseline Part II will also be listed as baseline for the Extended Disease Progression phase.

Part III: The last observation prior to or on the start date of Part III is denoted as, 'Baseline Part III'. It is defined for all efficacy variables as well as for the safety variables.

For each variable measured over time, the 'End Point Part I' value is defined as the last postbaseline assessment value during the run-in Phase. This value will be the same as the Baseline Part II value for subjects who continue into the Part II phase.

The 'End Point Part II' value is defined as the last postbaseline assessment value during the disease progression phase. This value will be the same as the Baseline Part III value for subjects who continue into the Part III phase.

The 'End Point Part III' value is defined as the last postbaseline assessment value during the Part III phase. The End Point Part III will also serve for the end point for the Disease Progression Phase.

A subject will be included in the end point analysis if they have both a baseline assessment and a post-baseline assessment.

2.3. Flagging Treatment Failures for Primary Analysis

All treatment failures that occur within 1 day from the Part II period discontinuation date will be used in the primary analysis in Part II. Treatment failures will also be captured in Part III. Treatment failures will also be identified during the combined Part II and III periods which is referred to as Extended Disease Progression phase using event and censoring dates. Timing and reasons for only repeated treatment failures will be determined by the Treatment Failure Board. Corresponding analysis set from the Treatment Failure Board will only be used in repeated events analysis in Parts II and EDP phases.

2.4. Visit Windows

As subjects may participate in more than one period in this trial, little knowledge is gained by presenting patient data with a treatment group in the listings. Thus, the following coding has been devised to compactly and efficiently describe a patient's treatment history in the trial. Although data summaries will be carried out for Part I, the treatment history information starts from Part II randomization to aid treatment group comparisons. Those subjects who are listed with codes OAPOAP and PPPP will by definition also be part of the Extended Disease Progression phase.

Code	Treatment History
PP	Entered Part II, randomized to paliperidone palmitate,
OAP	Entered Part II, randomized to oral antipsychotics
OAPPP	Entered Part III, randomized to paliperidone palmitate
OAPOAP	Entered Part III, randomized to oral antipsychotics
PPPP	Entered Part III, no second randomization

Because subjects do not always adhere to the protocol visit schedule, the following rules are applied to assign actual visits to protocol visits. Listed below are the visit windows for analysis and the target days for each visit. These windows are distinct from the visit windows specified in the Time and Event (T&E) Schedule in the protocol to inform the conduct of the study. The reference day for each phase is the corresponding first day. For those subjects who are randomized in Part III (PPPP treatment group), there will be another visit window schedule using the Part II Start Date (Part II Day 1) as the reference day. This will allow us compare OAPOAP and PPPP treatment groups starting from Part II through Part III treatment phase.

If a subject has 2 or more scheduled or unscheduled visits in one visit window, the visit closest to the target day will be used as the protocol visit for that visit window. If 2 actual visits are equidistant from the target day within a visit window, the later visit is used. If a visit window has no scheduled visits but does have unscheduled visits, then the unscheduled visit closest to the scheduled visit will be used.

All assignments will be made in chronological order. Once a visit is assigned to a visit window, it will no longer be used for a later time point except for the end point.

The last non-missing post-baseline value will be carried forward and will be defined as the last observation carried forward (LOCF) endpoint for all efficacy and safety parameters within each study phase. It is important that the carry-forward procedure only be applied within a phase and not across multiple phases. A subject will be included in the LOCF efficacy analysis if they have both a baseline assessment (except where a baseline assessment is not done) and a post-baseline assessment. Hence baseline assessment will not be carried forward. The LOCF assessments will be created for each post-baseline visit. If there are multiple visits within the last available time window, the last visit will be carried forward assuming the last visit took place within the reference end date. The imputed LOCF time points will be labeled 'XX LOCF'. For OAPOAP and PPPP treatment group comparisons, 18-month study (Extended Disease Progression phase), the absolute last visit scores will be listed as 'XX LOCF' where XX refers to visit label. The objective with the Extended Disease Progression phase is to show differences between those subjects who have been treated with PP for up to 18 months (early start group) compared with patients treated with OAP for up to18months. Extended Disease Progression phase will be labeled as EDP for the rest of the document. Listed below are the visit windows and the target days (if applicable) for each visit defined in the protocol for all phases (**Error! Reference source not found.**).

Table 1a: Time Intervals for Visits for MCCB, PSP					
Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Screening	1	-28 to -1	Screening	<1	-28 to -1
Part I		1	Baseline Part I[a]	<= 1	1
	5	29	Day 29 (Part I)	2 - 42	29
	7	57	Day 57 (Part I)	43 to End of Part I	57
		Part I Final Visit	End Point Part I	2 to End of Part I	
Part II	7	1	Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
		Part II Final Visit	End Point Part II	2 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	21	92	Day 92 (Part III)	2-134	92
	24	176	Day 176 (Part III)	135-218	176
	27	260	Day 260 (Part III)	219 to End of Part III	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[b]	7	1	Baseline EDP=Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
	21	352	Day 92 (Part III)	2-134	92
	24	436	Day 176 (Part III)	135-218	176
	27	520	Day 260 (Part III)	219 to End of Part III	260
		EDP Final Visit	End Point EDP Phase= End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit.

[a] Screening visit window or Visit 2 assessment will be listed as Part I Baseline.

[b] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

Table 1b: Time Intervals for Visits for CRDPSS					
Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Screening	1	-28 to -1	Screening	<1	-28 to -1
Part I	2	1	Baseline Part I	<= 1	1
	5	29	Day 29 (Part I)	2 - 42	29
	7	57	Day 57 (Part I)	43 to End of Part I	57
		Part I Final Visit	End Point Part I	2 to End of Part I	
Part II	7	1	Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
		Part II Final Visit	End Point Part II	2 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	21	92	Day 92 (Part III)	2-134	92
	24	176	Day 176 (Part III)	135-218	176
	27	260	Day 260 (Part III)	219 to End of Part III	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[a]	7	1	Baseline EDP=Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
	21	352	Day 92 (Part III)	2-134	92
	24	436	Day 176 (Part III)	135-218	176
	27	520	Day 260 (Part III)	219 to End of Part III	260
		EDP Final Visit	End Point EDP Phase= End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit.

[a] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

Table 1c: Time Intervals for Visits for CGI-S, MSQ					
Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Screening	1	-28 to -1	Screening	<1	-28 to -1
Part I	2	1	Baseline Part I	<= 1	1
	3	8	Day 8 (Part I)	2 - 12	8
	4	15	Day 15 (Part I)	13 – 22	15
	5	29	Day 29 (Part I)	23 – 36	29
	6	43	Day 43 (Part I)	37 – 50	43
	7	57	Day 57 (Part I)	51 to End of Part I	57
		Part I Final Visit	End Point Part I	2 to End of Part I	
Part II	7	1	Baseline Part II	1	1
	9	36	Day 36 (Part II)	2 – 50	36
	9+i	36+28*i (i=1..7)	Day 36+28*i (Part II)	50+28*(i-1)+1 to 50+28*i	36+28*i
	17	260	Week 260 (Part II)	247 to End of Part II	260
		Part II Final Visit	End Point Part II	2 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	19	36	Day 36 (Part III)	2-50	36
	19+i	36+28*i (i=1..7)	Day 36+28*i (Part III)	50+28*(i-1)+1 to 50+28*i	Day 36+28*i
	27	260	Week 260 (Part III)	247 to End of Part III	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[a]	7	1	Baseline EDP=Baseline Part II	1	1
	9	36	Day 36 (Part II)	2 – 50	36
	9+i	36+28*i (i=1..8)	Day 36+28*i (Part II)	50+28*(i-1)+1 to 50+28*i	36+28*i
	19	296	Day 36 (Part III)	2-50	36
	19+i	296+28*i (i=1..7)	Day 36+28*i (Part III)	50+28*(i-1)+1 to 50+28*i	Day 36+28*i
	27	520	Week 260 (Part III)	247 to End of Part III	260
		EDP Final Visit	End Point EDP Phase= End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit.

[a] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

Table 1d: Time Intervals for Visits for MRI

Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Part II	7	1	Baseline Part II	<=42	1
	11	92	Day 92 (Part II)	43 – 176	92
	17	260	Week 260 (Part II)	177 to End of Part II	260
		Part II Final Visit	End Point Part II	43 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	21	92	Day 92 (Part III)	2-176	92
	27	260	Week 260 (Part III)	177 to End of Part II	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[a]	7	1	Baseline EDP=Baseline Part II	<=42	1
	11	92	Day 92 (Part II)	43 – 176	92
	17	260	Week 260 (Part III)	177 to End of Part III	260
	21	352	Day 92 (Part III)	2-176	92
	27	520	Week 260 (Part III)	177 to End of Part III	260
		EDP Final Visit	End Point EDP Phase= End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit. The last post-baseline assessment in Part II will be used as Part III baseline.

[a] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

Table 1e: Time Intervals for Visits for Goal Setting and Review

Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Part I	2	1	Baseline Part I	<= 1	1
	5	29	Day 29 (Part I)	2 – 42	29
	7	57	Day 57 (Part I)	43 to End of Part I	57
		Part I Final Visit	End Point Part I	2 to End of Part I	
Part II	7	1	Baseline Part II	1	1
	9	36	Day 36 (Part II)	2 – 50	36
	9+i	36+28*i (i=1..7)	Day 36+28*i (Part II)	50+28*(i-1)+1 to 50+28*i	36+28*i
	17	260	Week 260 (Part II)	247 to End of Part II	260
		Part II Final Visit	End Point Part II	2 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	19	36	Day 36 (Part III)	2-50	36
	19+i	36+28*i (i=1..7)	Day 36+28*i (Part III)	50+28*(i-1)+1 to 50+28*i	Day 36+28*i
	27	260	Week 260 (Part III)	247 to End of Part III	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[a]	7	1	Baseline EDP=Baseline Part II	1	1
	9	36	Day 36 (Part II)	2 – 50	36
	9+i	36+28*i (i=1..8)	Day 36+28*i (Part II)	50+28*(i-1)+1 to 50+28*i	36+28*i
	19	296	Day 36 (Part III)	2-50	36
	19+i	296+28*i (i=1..7)	Day 36+28*i (Part III)	50+28*(i-1)+1 to 50+28*i	Day 36+28*i
	27	520	Week 260 (Part III)	247 to End of Part III	260
		EDP Final Visit	End Point EDP Phase= End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit.

[a] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

Table 1f: Time Intervals for Visits for Daily Activities, RUQ

Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Part I	2	1	Baseline Part I	<= 1	1
	7	57	Day 57 (Part I)	2 to End of Part I	57
		Part I Final Visit	End Point Part I	2 to End of Part I	
Part II	7	1	Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
		Part II Final Visit	End Point Part II	2 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	21	92	Day 92 (Part III)	2-134	92
	24	176	Day 176 (Part III)	135-218	176
	27	260	Day 260 (Part III)	219 to End of Part III	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[a]	7	1	Baseline EDP=Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
	21	352	Day 92 (Part III)	2-134	92
	24	436	Day 176 (Part III)	135-218	176
	27	520	Day 260 (Part III)	219 to End of Part III	260
		EDP Final Visit	End Point EDP = End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit.

[a] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

Table 1g: Time Intervals for Visits for ESR-A					
Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Screening	1	-28 to -1	Screening	<1	-28 to -1
Part I		1	Baseline Part I[a]	<= 1	1
	7	57	Day 57 (Part I)	2 to End of Part I	57
		Part I Final Visit	End Point Part I	2 to End of Part I	
Part II	7	1	Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
		Part II Final Visit	End Point Part II	2 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	21	92	Day 92 (Part III)	2-134	92
	24	176	Day 176 (Part III)	135-218	176
	27	260	Day 260 (Part III)	219 to End of Part III	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[b]	7	1	Baseline EDP=Baseline Part II	1	1
	11	92	Day 92 (Part II)	2-134	92
	14	176	Day 176 (Part II)	135-218	176
	17	260	Day 260 (Part II)	219 to End of Part II	260
	21	352	Day 92 (Part III)	2-134	92
	24	436	Day 176 (Part III)	135-218	176
	27	520	Day 260 (Part III)	219 to End of Part III	260
		EDP Final Visit	End Point EDP Phase= End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit.

[a] Screening visit window or Visit 2 assessment will be listed as Part I Baseline.

[b] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

Table 1h: Time Intervals for Visits for ISST Plus, Vital Signs, and Weight					
Analysis Phase	Scheduled Visit Number	Scheduled Day	Time Interval (label on output)	Time Interval (Day)	Target Time Point from start of each phase (Day)
Screening	1	-28 to -1	Screening	<1	-28 to -1
Part I	2	1	Baseline Part I	<= 1	1
	3	8	Day 8 (Part I)	2 - 12	8
	4	15	Day 15 (Part I)	13 – 22	15
	5	29	Day 29 (Part I)	23 – 36	29
	6	43	Day 43 (Part I)	37 – 50	43
	7	57	Day 57 (Part I)	51 to End of Part I	57
		Part I Final Visit	End Point Part I	2 to End of Part I	
Part II	7	1	Baseline Part II	1	1
	8	8	Day 8 (Part II)	2 - 22	8
	9	36	Day 36 (Part II)	23 – 50	36
	9+i	36+28*i (i=1..7)	Day 36+28*i (Part II)	50+28*(i-1)+1 to 50+28*i	36+28*i
	17	260	Week 260 (Part II)	247 to End of Part II	260
		Part II Final Visit	End Point Part II	2 to End of Part II	
Part III	17	1	Baseline Part III	1	1
	18	8	Day 8 (Part III)	2 - 22	8
	19	36	Day 36 (Part III)	23 - 50	36
	19+i	36+28*i (i=1..7)	Day 36+28*i (Part III)	50+28*(i-1)+1 to 50+28*i	Day 36+28*i
	27	260	Week 260 (Part III)	247 to End of Part III	260
		Part III Final Visit	End Point Part III	2 to End of Part III	
EDP[a]	7	1	Baseline EDP=Baseline Part II	1	1
	8	8	Day 8 (Part II)	2 - 22	8
	9	36	Day 36 (Part II)	23 – 50	36
	9+i	36+28*i (i=1..8)	Day 36+28*i (Part II)	50+28*(i-1)+1 to 50+28*i	36+28*i
	18	268	Day 8 (Part III)	2 - 22	8
	19	296	Day 36 (Part III)	23 - 50	36
	19+i	296+28*i (i=1..7)	Day 36+28*i (Part III)	50+28*(i-1)+1 to 50+28*i	Day 36+28*i
	27	520	Week 260 (Part III)	247 to End of Part III	260
		EDP Final Visit	End Point EDP Phase= End Point Part III	2 to End of Part III	

Note: The time intervals and target time point days are relative to the respective phase. Final post-baseline assessment during the preceding phase will be used as the value for the current phase baseline visit.

[b] Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP.

LABs

Laboratory assessments will be slotted to visit windows according to sample date rather than the visit date. Clinical laboratories will be summarized at the following visits: Baseline Part I, End Point Part I, Baseline Part II, End Point Part II, Baseline Part III, and End Point Part III. Baseline for the Part I will be captured during the screening period.

Combination of Part II and Part III scores will be utilized for the Extended Disease Progression phase. Extended Disease Progression phase starts from Part II and extends into Part III for only those subjects who are on PPPP and OAPOAP. Part II baseline will serve as EDP Baseline and Part III end point will serve as EDP end point.

Observations from local labs with missing units will be omitted.

2.5. Analysis Sets

Efficacy analyses will be based on the intent-to-treat (ITT) and explanatory intent-to-treat (eITT) analysis sets. All safety analyses will be based on ITT analysis set. As there are different objectives for different periods of this study, separate ITT and eITT populations will be identified for the Part I (Run-in), Part II (Disease Progression), Part III (Disease Modification) and EDP (Extended Disease Progression) periods. The eITT analysis set will not be included in Part I. The analysis sets are defined below.

Part I (Run-in) Intent-to-Treat Analysis Set

Efficacy and safety summaries for the Part I period will be based on the Part I ITT analysis set, which will include all subjects who receive at least one dose of study medication (or any portion of the dose) in Part I, regardless of their compliance with the protocol.

Part II (Disease Progression) Analysis Sets

Efficacy and safety summaries for the Part II period will be based on the Part II ITT analysis set, which will include all randomized subjects who receive at least one dose of randomized study medication (or any portion of the dose) in Part II, regardless of their compliance with the protocol.

An explanatory ITT (eITT) analysis set will also be defined in Part II. The eITT analysis set will consist of all ITT subjects, as well as their study assessments for the time period between the date of randomization and the eITT end point. The eITT end point for subjects who had a treatment failure will be defined as 1st treatment failure date +7 days. For those subjects who did not have treatment failure, the eITT end point will be same as the Part II end point. Changes in MCCB, PSP, ICM volumes, and CGI-S will also be analyzed using eITT analysis set.

The MRI ITT analysis set in Part II will include those subjects who are part of the Part II ITT analysis set and had a confirmed MRI scan anytime during Part II.

Part III Analysis Sets

Efficacy and safety summaries for the Part III period will be based on the Part III ITT analysis set, which will include all randomized subjects who receive at least one dose of study medication (or any portion of the dose) in Part III, regardless of their compliance with the protocol. On Day 1 of Part III, subjects in the OAP treatment arm will be re-randomized in a 1:1 ratio to continued treatment with their oral treatment or to PP (delayed-start PP arm). Subjects previously assigned to treatment with PP in Part II will continue on that treatment; however, in order to be included in the ITT population, this group must have at least one post-baseline MCCB or PSP assessment.

An explanatory ITT (eITT) analysis set will be defined in Part III. The eITT analysis set will consist of all ITT subjects, as well as their study assessments for the time period between the date of randomization and the eITT end point. The eITT end point for subjects who had a treatment failure will be defined as 1st treatment failure date +7 days. For those subjects who did not have treatment failure, the eITT end point will be same as the Part III end point. Changes in MCCB, PSP, ICM volumes, and CGI-S will also be analyzed using eITT analysis set.

The MRI ITT analysis set in Part III will include those subjects who are part of the Part III ITT analysis set and had a confirmed MRI scan anytime during Part III.

EDP (Extended Disease Progression) Analysis Sets

Efficacy and safety summaries for the EDP period will be based on the Part III ITT analysis set, which will include all randomized subjects who receive at least one dose of study medication (or any portion of the dose) in Part III, regardless of their compliance with the protocol.

An explanatory ITT (eITT) analysis set for EDP phase will be same as eITT for Part III.

The MRI ITT analysis set for the EDP period will be based on Part III MRI ITT analysis set.

2.6. Definition of Subgroups

Analyses will be provided for the overall primary endpoint, time to treatment failure, by the following subgroups.

- Sex
- Race (White, Black, Other)
- Age Group in years (<20 years, 20 to 24 years, >24 years)
- Prior Antipsychotic Exposure (<6 months, 6 to 12 months, >12 months)
- MCCB Composite Score using the Part I – Visit 5 assessment (<25, 25-35, >35)
- PSP Score using the Part I – Visit 5 assessment (<=70, >70)

- Substance Abuse History (0 years, >0 to 2 years, > 2 years)
- Current Primary DSM-5 Diagnosis (schizophrenia and schizophreniform)
- Subjects who continued into Part III and those who did not.

2.7. Imputation Rules for Missing Dates

2.7.1. Missing AE Dates

A conservative approach will be used to handle any missing portion of dates for adverse events. The rules for handling the incomplete AE onset dates will be as follows:

(1) The missing day of the month will be estimated as follows: If the month and year are known, compare the month with the month of Part I start date, Part II start date, and Part III start date.

If the month of AE is equal to the month of Part I start date then the estimated day is the start day of Part I.

If the month of AE is equal to the month of Part II start date then the estimated date is the start day of Part II.

If the month of AE is equal to the month of Part III start date then the estimated date for AE will be the start day of the Part III.

If the month of AE is different (greater than or less than) from the month of Part I, or Part II, or Part III then the estimated day for AE is the 1st day of the month.

(2) If both the day and the month are missing: NO estimation is to be performed.

For incomplete AE recovery dates, the rules are:

1. The missing day of the month will be estimated as follows: If the month and year are known then the missing day will be imputed as the last day of that month.
2. If both the day and the month are missing then the imputed date will be the date of the reference end date (provided the AE start date is not missing and after the Part I start date).
3. If the AE end date is missing however and the AE outcome says the AE is not recovered/not resolved then the AE end date = reference end date.

If the imputed AE date is greater than the reference end date then the imputed AE will be cutoff at the reference end date.

Treatment-emergent flags will be created for identifying AEs that occurred in each phases of the study. Treatment-emergent adverse events (TEAEs) and serious treatment-emergent adverse events (SAEs) will be defined as adverse events (AEs) with start dates on or after the start of a

period and prior to 30 days after the permanent discontinuation of the study medication. An ongoing AE diagnosed prior to the start date of treatment in a period will be included in that period if there is an increase in its severity or change in relationship to study medication on or after the beginning of the period. Adverse events with completely missing onset dates will not be included in the summarizations. In addition to TEAE flags based on AE start dates, another flag will be created to identify those AEs that last more than one period.

2.7.2. Missing Concomitant Medication Dates

If the start or end date of the drug intake is missing, the following rules will be applied. If only the day is missing (month and year are not missing) then the first day of the month will be used as the start day and the last day of the month will be used as an end day. If month or year is missing, then no imputation will be made. Exceptions are: when the drug intake is marked as "prior" or "continuing", the end date of the respective period is the last day of study medication. The first day of the year will be used if only the year is present for a start date. This convention is similar to prior medications date imputation rule.

Concomitant medications will be defined as medications with a start date equal to or greater than the first date of each period or with a start date prior to the first date of each period that are either ongoing or have a stop date greater than the first date of each period or a stop date that is missing. In this way, concomitant medications will be carried forward from previous periods based on their stop date and ongoing status. Medications started on the Visit 27 date (trial termination visit) or after the Visit 27 date will not be considered concomitant since subjects are allowed to use any medication after trial completion.

If a partial date is reported, it is assumed the medication (or therapy) was taken in all phases that overlap with the partial date. If both start and end dates are missing but this concomitant medication was taken both prior to the study entry and still ongoing at study end, it is assumed medication was taken in all phases.

2.7.3. Baseline Hospitalization Dates

The duration (days) of the index hospitalization will be calculated as (hospitalization end date from RUQ page – hospitalization start date +1). The current hospitalization duration will be computed matching the RUQ Hospitalization Start Date to the Date of Current Hospitalization Admission in the Psychiatric History page. The corresponding STOP date will be identified from the RUQ page. Index hospitalization start and stop dates can be identified from the RUQ page.

If the hospitalization start/stop date is completely missing, or the year is missing, then no imputation will be performed for the prior hospitalizations. If the start/stop dates of hospitalization are partially missing the following rule will apply:

Hospitalization start date:

- If only the day is missing, it is imputed using the first day of the month.
- If only the month is missing, it is imputed using January.

- If only both the day and month are missing, it is imputed using January 1.

Hospitalization stop date:

The earlier of the date before the current hospitalization admission and the following imputed date:

- If only the day is missing, it is imputed using the last day of the month.
- If only the month is missing, it is imputed using December.
- If only both the day and the month are missing, it is imputed using December 31.

2.7.4. Baseline Characteristics and Prior Medications

Partial dates will be imputed as follows for analysis of prior medication dose administration, demographic and background characteristics:

- The first day of the month will be used if only the month and year are present for a start date. The same scheme will be applied for the date of first schizophrenia disorder diagnosis, the onset date of first psychotic episode, prior medications, etc.
- The first day of the year will be used if only the year is present for a start date. The same scheme will be applied for the date of first schizophrenia disorder diagnosis, the onset date of first psychotic episode, prior medications, etc.
- The last day of the month will be used if only the month and year are present for a stop date and if the month is prior to month of Day 1 (Part I). If the month of a stop date and month of Day 1 (Part I) is the same, then the screening date will be used as the stop date.

The last day of the year will be used if only the year is present for a stop date and if year is prior to Day 1 (Part I) year. If the year of a stop date and year of Day 1 (Part I) is the same, then the screening date will be used as the stop date.

2.7.5. Post Part I Baseline Hospitalization, Emergency Room Visit Dates

The post Part I baseline hospitalization related dates should not be missing. In case of partial dates, the following rules will be applied. However, prior to implementing this rule, all the available data correction tactics must be considered to capture the missing day field.

Hospitalization start date or emergency room visit:

- If the trial medication started during that month, then set the start date equal to the trial medication date.
- If the trial medication started prior to that month, then set the start date equal to the 1st day of that month.

- If the study medication has not been administered during that month, then set the start date equal to the 1st day of that month.
- If the month is split over phases, then set the start date equal to the trial medication date.
- If the day and month of the start date are missing, then leave date as missing.

Hospitalization end date:

- If the trial medication stopped during that month, then set the end date equal to the stop date of the trial medication.
- If the trial medication stopped after that month, then set the end date equal to the last day of that month.
- If the trial medication stopped prior to that month, then set the end date equal to the 1st day of that month.
- If the day and month of the end date are missing, then leave end date as missing.

3. RULES FOR PRESENTATIONS

The overall primary efficacy end point is the time between subject randomization and the first documentation of a treatment failure during the Part II phase. Treatment failures will also be examined in Part III and during the 18-month EDP study period. The primary efficacy null hypothesis is that there is no difference in distribution of time to treatment failure between paliperidone palmitate and oral antipsychotics. There will be no interim analysis.

Summary statistics will be presented to give a general description of the patients studied and an overview of the efficacy and safety results. Data from all sites will be combined in the computation of these descriptive summaries. Categorical variables will be summarized by presenting the number and percentage of patients in each category. All percentages will be presented to one decimal point. Continuous variables will be summarized using n, mean, standard deviation (SD), median, minimum, and maximum values.

All statistical tests will be two-sided. All p-values will be rounded to three decimal places; p-values less than 0.001 will be presented as <0.001 in all tables; p-values greater than 0.999 will be presented as 1.000. All group comparisons from analysis of variance (ANOVA) and analysis of covariance (ANCOVA) models will be based on Type III sums of squares. All confidence intervals will be two-sided with 95% coverage. The primary analysis will be evaluated for statistical significance at the 0.05 level.

All data from the Part I phase will be categorized and presented in three main columns: ALL (all Part I subjects), those patients who entered Part II, and those patients who did not enter Part II. There will be no statistical testing to compare these groups.

All data from the Part II phase will be categorized and presented using two columns: those patients who are randomized to paliperidone palmitate (PP) and those patients who are randomized to oral antipsychotics (OAP).

All data from the Part III phase will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those patients who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

All data from 18-month EDP phase of the study will be categorized and presented using two columns: those patients who are NOT re-randomized (PPPP) in Part III, and those patients who are re-randomized to oral antipsychotics (OAPOAP) in Part III. The differences at Month 18 represent the cumulative effect on disease progression.

Computations for all of the results will be performed using the SAS® Version 9.2 or higher computer software package or R.

4. SUBJECT INFORMATION

4.1. Discontinuation/Completion Information

The numbers and percentage of subjects entering the Part I, Part II, Part III, and EPD phase will be summarized for the ITT Analysis Set. Those subjects including healthy controls undergoing MRI assessments will also be summarized.

The number and percentage of subjects who completed the trial, and those who prematurely discontinued from the trial along with the reason for discontinuation, will be summarized by period and treatment group. In addition, the number of subjects who withdrew by visit will be provided. For Part II, additional summaries will be generated for the subgroups of patients in the PP treatment group based on patterns of dose modifications/transitions.

Time to all cause discontinuation in Part II, Part III, and EDP phase (expressed as number of days from randomization or Day 1 to termination) will be analyzed using Kaplan-Meier estimated time-to-event curves, and compared among treatment groups using log-rank tests. Time to discontinuation excluding treatment failure will also be derived for Part II, Part III, and EDP phase.

Subject enrollment by country and study sites will also be listed.

4.2. Demographics and Baseline Characteristics

Demographics baseline characteristics, psychiatric history, medical history, and nicotine and substance use will be presented for the ITT populations in all periods.

Demographic and clinical characteristics, including age (years), age category (<20 years, 20- 24 years, >24 years), gender, race, ethnicity, height (cm), weight (kg) and body mass index (BMI) (kg/m^2), and BMI category (normal: BMI< 25, overweight: 25≤ BMI< 30, and obese: BMI≥ 30),

education, employment status, and living situation at study entry will be summarized descriptively by study phase and treatment group. Baseline values of efficacy variables, including MCCB, PSP, CRDPSS, CGI-S, and MSQ score will be summarized. Randomization factors, specifically, age group in years (<20 years, 20 to 24 years, >24 years), prior antipsychotic exposure (<6 months, 6 to 12 months, >12 months), MCCB composite score using the Part I – Visit 5 assessment (<25, 25-35, >35), PSP Score using the Part I – Visit 5 assessment (<=70, >70), and substance abuse history (0 years, >0 to 2 years, > 2 years) summaries will also be provided for Parts II and III, and EDP phase.

Psychiatric history, including current primary diagnosis, duration of illness (≤ 5 years vs. >5 years - using date of first schizophrenia or schizophreniform diagnosis), duration since first psychiatric episode (years), age (years) at first psychiatric episode, age (years) at first psychiatric diagnosis, age (years) at first pharmacological treatment for psychiatric symptoms, prior psychiatric diagnosis, setting of current psychiatric treatment, duration (days) of current hospitalization, age at first psychiatric hospitalization, number of total psychiatric hospitalizations (both continuous and categorical summaries, 0, 1, 2, 3, and >3 times), time since last psychiatric hospitalization discharge, any suicide attempts, number of suicide attempts, time since last suicide attempt, and birth history will be summarized.

In Part II, demographic and clinical characteristics and psychiatric history summaries will be repeated for the following subgroups: 1-) Current Primary DSM-5 Diagnosis (schizophrenia and schizophreniform); 2-) Subjects who continued into Part III and those who did not.

Previous and current use of tobacco and previous substance use will be summarized. If an imputed date for last substance use is later than the screening date, the imputed date will be set equal to the screening date when calculating the time since last substance use. Post-baseline tobacco related summaries will also be presented.

Medical history will be summarized for each study period using corresponding ITT analysis sets. The number of abnormalities will be displayed for medical histories that are both not currently active and currently active.

Prior medications will be defined as medications with start dates prior to the start of study medication or medications with prior flag. Prior medications will be coded to ATC level 4 and generic terms using the World Health Organization (WHO) Drug Dictionary, version 2017[1] or higher version. Psychotropic medications will be further classified based on sponsor-defined medication classes (see Appendix Table 2) and sponsor-defined preferred terms that combine like medications. Prior non-psychotropic and psychotropic medications will be summarized separately. Prior non-psychotropic medications will be summarized by ATC Level 4 and generic term using frequencies and percentages by treatment group. Prior psychotropic medications will be presented by sponsor-defined class and sponsor-defined preferred term. Additionally, the number of subjects taking 0, 1, 2, or >2 medications will be presented.

Psychotropic medications taken at baseline only for Part I, defined as medications taken (present) within 7 days prior to the screening date and one day following the screening date will also be summarized.

Hospital admission and discharge information will be summarized for the ITT population using the RUQ page. It will include the duration of the hospital stay for the index episode and the cumulative duration of rehospitalizations.

4.3. Extent of Treatment Exposure

4.3.1. Injections

The number and percent of subjects who receive 1, 2, 3, etc. injections of PP will be listed. The total exposure (days) to paliperidone palmitate will also be summarized. This summary will be presented separately for each of the study periods. The number and percent of subjects at each dose level will be summarized. For each month during the Parts II and III, and during EDP phase, a frequency distribution will be presented showing the number of subjects who receive the injection in the deltoid or gluteal muscle. The treatment exposure, mean dose, mode dose, and final dose will be presented both continuously and categorically for each period separately.

4.3.2. Oral Antipsychotics

The number and percentage of subjects who receive study drug will be summarized for each study period. Total OAP exposure from study phases Part I, Part II, Part III, and EDP will also be presented. The total duration of exposure will be based on the relative day and study day fields which were described in earlier sections. Frequency distribution of the duration will be presented separately for each study phase. Mode and last dose levels will also be summarized. Cumulative frequency distribution will also be presented. Additionally, the average dose, mean modal dose, and the last dose of OAPs will be summarized.

4.4. Protocol Deviations

The frequency and percent of protocol violations excluding screen failures will be presented for each of the study phases separately using corresponding ITT populations. Important protocol violations will be identified by a clinical review of the data prior to data base lock. Protocol violations include, but are not limited to, taking a medication other than those allowed by the protocol, failure to satisfy inclusion/exclusion criteria or stabilization criteria, wrong treatment given, and efficacy scales are performed by non-qualified personnel.

4.5. Prior and Concomitant Medications

Concomitant medications will be defined as medications with a start date equal to or greater than the first study phase date of each phase or with a start date prior to the first study date of each phase that are either ongoing or have a stop date greater than the first study date of each phase or a stop date that is missing. In this way, concomitant medications will be carried forward from previous phases based on their stop date and ongoing status. Medications started on the Visit 27 date (trial termination visit) or after the Visit 27 date will not be considered concomitant since

subjects are allowed to use any medication after trial completion. Concomitant medications will be coded to ATC class and generic name using the World Health Organization (WHO) Drug Dictionary. Concomitant medication use will be summarized by number of medications (None (if any), 1, 2, ≥3) as well as by the ATC class and generic name using frequencies and percentages for each of the ITT populations. Psychotropic medications will be further classified based on sponsor-defined medication classes (see Appendix Table 2) and sponsor-defined categorization of preferred terms that combine like medications and will be summarized. The number and percentage of subjects taking psychotropic medications due to an adverse event will also be included. Patients with multiple occurrences of a medication in an ATC class and/or generic name will only be counted once within each ATC class and/or generic name.

4.5.1. Prior Psychotropic Medications

The number and percentage of subjects who received psychotropic medications prior to the screening date will be presented by the generic term category within the psychotropic group. Each medication will be categorized into one of the following psychotropic groups: Acetylcholinesterase inhibitors, Antidepressants, Anti-extrapyramidal symptoms (EPS), Antihistamines, Atypical antipsychotics, Benzodiazepines, Beta blockers, Depot antipsychotics, Mood stabilizers and antiepileptics, Non-benzodiazepines hypnotics and anxiolytics, Stimulants, and Typical antipsychotics. These summaries will be presented for the Parts I, II, III, and EDP phase ITT population analysis sets.

The number and percentage of subjects who received prior antiparkinsonian medications (anticholinergics [anti-EPS] or antihistamines) will be provided by the generic term category within the psychotropic group.

4.5.2. Concomitant Benzodiazepines (Sedatives/Hypnotics/Anxiolytics)

The number and percentage of subjects who received benzodiazepines during the Parts II and III, and EDP phase will be provided by generic term category for the ITT analysis sets.

For each subject, the total duration of each benzodiazepine therapy will be calculated. Descriptive statistics of the duration (days) of benzodiazepine therapy use during the study periods will be presented for the ITT analysis sets. Duration of concomitant medication is defined as stop date – start date +1, where stop and start dates of the concomitant medication during the respective phases are defined below.

For the summary of duration of concomitant medication during the Part II phase, if the start date of concomitant medication is prior to the Part II start date, the Part II start date will be used as concomitant medication start date for duration calculation. If the end date of concomitant medication is after the Part II end date or indicated as continuing, then the Part II end date will be used as the concomitant medication end date for duration calculation.

For the summary of duration of concomitant medication during the Part III phase, if the start date of concomitant medication is prior to the Part II start date, then the Part III start date will be used as concomitant medication start date for duration calculation. If the end date of concomitant

medication is after the Part III end date or indicated as continuing, then the Part III end date will be used as the concomitant end date in calculating duration.

For the summary of duration of concomitant medication during the EDP phase, Part II start date and Part III end dates will be used to compute duration.

4.5.3. Concomitant Medications Other Than Benzodiazepines

The number and percentage of subjects who receive concomitant therapies other than benzodiazepines during the study will be provided by generic term category for the ITT analysis sets. Those concomitant medications, other than benzodiazepines, received by at least 5% of the subjects in either study treatment group will be presented for the Part II, III, and EDP phase ITT analysis sets.

4.5.4. Concomitant Antiparkinsonian Medications

The number and percentage of subjects who received concomitant antiparkinsonian medications (anticholinergics [anti-EPS] or antihistamines) will be provided by the generic term category within the psychotropic group for each of the study phase using the ITT analysis sets.

5. EFFICACY

5.1. Level of Significance

The overall primary hypothesis to be tested in this study is that 9 months' treatment with PP is superior to 9 months' OAP treatment in delaying time to treatment failure in subjects with recent-onset schizophrenia or schizophreniform disorder. This will be assessed at the end of Part II.

Other hypotheses to be evaluated in this study will assess whether LAI treatment with PP can slow down disease progression and possibly modify disease course in recent-onset subjects compared to OAP medications. This will be assessed by tracking changes in cognition (MCCB composite score), functioning (PSP), and brain imaging assessments (ICM volume).

The hypothesized effect of PP versus OAP on disease progression and disease modification is summarized in **Error! Reference source not found.**. Disease progression and disease modification will be assessed for MCCB composite score, PSP total score, and for ICM volume.

Part II (Disease Progression) will enroll all subjects who complete Part I. Subjects will be randomized in a 1:2 ratio to either start treatment with PP or to continue with their OAP treatment from Part I. The early treatment effect identified after early start at the conclusion of Part II (δ_{21}) (Figure 2) represents the treatment effect after 9 months resulting from continuous treatment with LAI compared with oral treatment. It seeks to identify differential treatment effects between the two alternative approaches on disease progression at the end of Part II. In particular, in addition to differences in time to first treatment failure, differences in cognition, functioning, and ICM volume will be assessed. At the conclusion of Part II, the database will be soft-locked and data analyzed to determine if there is an early treatment effect (δ_{21}= treatment effect on disease progression) demonstrating superiority of PP relative to oral treatment.

At the onset of Part III, subjects in the OAP arm will be re-randomized to continued treatment with their prior oral treatment or to PP. Subjects will be followed for an additional 9 months. Three effects will be assessed for a given endpoint: δ_{31}, δ_{32}, and δ_{33}. The quantity δ_{31} represents the lead treatment effect after an early start with 18 months of PP treatment and shows the lead effect remaining in the early start group over the delayed-start effect after a 9-month treatment duration. The quantity δ_{32} represents the delayed-start effect on disease progression for PP compared to continuing OAP after 9 months of additional treatment. The quantity δ_{33} represents the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment. At the conclusion of Part III, the database will be locked and data analyzed to assess response in variables δ_{31}, δ_{32}, and δ_{33}. Positive results (δ_{33} is significant [a differential treatment effect is still evident] **and** δ_{31} is significant [the lead effect remains significant after 9 months]) will be used to support a conclusion of disease modification using MCCB composite score in addition to examining difference between δ_{31} and δ_{21}. Similar observations for changes in PSP and ICM will be assessed.

In Part II, the overall Type I error rate for testing PP (PP1M/PP3M) versus OAP for both the primary efficacy endpoint and key secondary efficacy endpoints will be controlled at the 2-sided 0.05 significance level using a combination of fixed sequence gatekeeper approach and Holm's step-down procedure[2]. Time to first treatment failure will be tested first, followed by change in key secondary endpoints (MCCB, PSP, and ICM). Time to first treatment failure will be examined first using the log-rank test statistics, then using a Cox's proportional hazards model to describe results based on hazard ratio. If the null hypothesis corresponding to time to first treatment failure is rejected, then the key secondary endpoints will be tested at the 5% level using Holm's step down procedure, thus maintaining an overall Type I error rate of 5%. If the primary null hypothesis is not rejected, testing of change in key secondary endpoints will still be performed, but no unqualified statements about the statistical significance regarding change will be made. In Holm's procedure, p-values from analyses of MCCB, PSP, and ICM will be ordered from lowest to highest. Let $p(1)<p(2)<p(3)$ to be ordered p-values corresponding null hypotheses H(1), H(2), and H(3). In Step 1, if $p(1) < 0.05/3$ then corresponding null hypothesis will be rejected and testing will be examined in second step; otherwise, none of the hypotheses will be rejected and we will stop testing. In Step 2, if $p(2) < 0.05/2$ then corresponding null hypothesis will be rejected and move onto third step; otherwise, H(2) and H(3) will not be rejected and we will stop testing. In Step 3, p(3) will be tested at 0.05 level. All other exploratory hypotheses will be tested at the 2-sided 0.05 significance level without adjustments for multiplicity.

For the examination of extended disease progression and disease modification, there will be no adjustments for multiplicity. During these phases, we seek to show that subjects with recent-onset schizophrenia or schizophreniform disorder who have a delayed start in their treatment with PP fail to catch up with subjects who start 9 months earlier in the course of their disease. We hypothesize that changes on measures of a pathophysiology biomarker (ICM), clinical symptoms (CGI), cognition (MCCB) and functioning (PSP) should be relatively better in subjects who initiate treatment earlier.

5.2. Part II, Primary Efficacy Endpoint

The primary efficacy endpoint is the time from Day 1 Part II (randomization) to the first treatment failure during the Part II phase.

Subjects will be evaluated at the time points indicated in the Time and Events Schedule for the occurrence of any of the events below that are identified as treatment failures. Only treatment failure events occurring after randomization in Part II (up to the end of Part III) will be assessed. The investigator will determine whether the event meets one of the protocol definitions, and will provide documentation of dates and times associated with the event. Treatment failure will be defined as any of the following:

- Psychiatric hospitalization due to worsening symptoms (including Emergency Room visits ≥23 hours, and not including hospitalization due to social reasons);
- Any deliberate self-injury, suicidal ideation or behavior, homicidal ideation or violent behavior that is clinically significant and needs immediate intervention as determined by the study physician;
- New arrest/incarceration (not related to probation or existing warrant);
- Discontinuation of antipsychotic treatment due to inadequate efficacy as determined by the study physician;
- Discontinuation of antipsychotic treatment due to safety or tolerability as determined by the study physician;
- Treatment supplementation with another antipsychotic due to inadequate efficacy as determined by the study physician (note: use of oral paliperidone ER in the PP treatment group will not be considered a treatment failure unless supplemental use of paliperidone ER or oral risperidone exceeds the dose levels or treatment durations specified in study protocol - Section **Error! Reference source not found.**); and
- Increase in the level of psychiatric services (such as from office visit to day hospitalization) in order to prevent imminent psychiatric hospitalization as determined by the study physician.

Any changes in antipsychotic medications (switching, discontinuation, or add-on) must be evaluated against the treatment failure criteria. If any of these changes do not meet the treatment failure criteria, these must be documented and recorded in the eDC.

At the time of **first** treatment failure, or as soon as possible thereafter, all efficacy assessments (MCCB, PSP, CGI-S, CRDPSS, and MSQ) and MRI imaging (in the subgroup of subjects

undergoing MRI assessment) must be performed. At the time of any **subsequent** treatment failures (or as soon as possible thereafter), PSP, CGI-S, CRDPSS, and MSQ assessments should be performed.

Note that subjects who experience a treatment failure and do not withdraw consent or meet the criteria for withdrawal from the study will continue in the study and be followed through to the end of the study.

Subject status during Part II	Time to **First** Treatment Failure/Censoring	Censoring indicator
Randomized subjects who experienced treatment failure during Part II	(Date of treatment failure – Part II start date) + 1	No
Randomized subjects who remained treatment failure-free at the end of the Part II	(End of Part II date – Part II start date) + 1	Yes
Early withdrawal/discontinued during the Part II without treatment failure	(Date of early withdrawal – Part II start date) + 1	Yes

5.2.1. Primary Estimand

The primary efficacy estimand, the main clinical quantity of interest to be estimated in the study, is defined by the following components:

The *treatment* condition in Part II is:

- Paliperidone Palmitate (PP) treatment group including PP1M for a minimum 4 months (five to eight injections) followed by PP3M treatment.
- Oral Antipsychotics (OAP) treatment group includes oral paliperidone or risperidone or another protocol-specified oral AP: aripiprazole, haloperidol, perphenazine, quetiapine, or olanzapine.

The *population* is restricted to subjects with recent-onset schizophrenia or schizophreniform disorder who were randomized in Part II and took any study medication regardless of compliance with study protocol.

The *variable* is time to first treatment failure during the disease progression phase (Part II).

The *intercurrent events and corresponding strategy are the following:*

- Treatment discontinuation – Hypothetical Strategy: After treatment discontinuation in Part II, assume similar efficacy for subjects who discontinued as those subjects from the same treatment group who did not discontinue the treatment.

The *population-level summary* is the Kaplan-Meier estimate of the survival function. The primary analysis will be based on the Part II ITT analysis set.

5.2.2. Analysis Methods

Under the primary estimand, only treatment failures occurring during the Part II phase prior to treatment discontinuation will be counted as events in the primary analysis. Treatment differences will be compared using a log-rank test. The cumulative distribution function of the time to treatment failure will be estimated by the Kaplan-Meier method [3,4]. The 95% CIs for the median treatment failure rates, as well as the failure rates at 3 months, 6 months, and at 9 months will be provided. Standard Error (SE) estimates will be computed using Greenwood's formula. The estimate of the hazard ratio and its 95% CI will be provided by treatment group based on the Cox proportional hazards model[5]. The reasons for first treatment failure and subsequent treatment failures will be summarized at each visit and end point.

If the log-rank test result is statistically significant (p-value <0.05) in favor of PP, the null hypothesis would be rejected, and we would conclude that PP is superior to OAP in delaying time to treatment failure. Hazard ratio and 95% CI will be used to describe reduction in risk treatment failure.

The main analysis relies on the assumption of ignorable censoring. Therefore, sensitivity analyses will be performed to stress-test the robustness of results to deviations from ignorable censoring for the Part II ITT analysis population. Specifically, it is assumed that subjects on PP who discontinue prematurely from the Part II phase have a higher event hazard starting from the discontinuation time, compared with similar subjects who remain in this phase. The higher event hazard is determined by single sensitivity parameter Delta, representing the ratio of subject-specific hazard at any given time point *t* following discontinuation compared to that same subject's hazard at the same time *t* if he or she had remained in the Part II. A Kaplan-Meier multiple imputation (KMMI) non-parametric approach will be used for the imputation of treatment failure events as described by Taylor et al. (2002)[6] and Lipkovich et al. (2016)[7]. The number of multiple imputations (MI) will be set to 1000 and a seed equal to 234 will be used for MI. A sequence of Delta values will be used for all subjects with non-administrative censoring from the PP group (i.e. subjects censored due to other reasons than the end of Part II phase cut-off), starting with 1 (ignorable censoring) and increasing by 1 until the point until the tipping point, where the results are no longer significant. For the OAP group, the sensitivity parameter Delta will be set to one, i.e. maintaining the ignorable censoring assumption.

An additional analysis will also be carried out for time to early discontinuation of study medication during the Part II for any reason including treatment failure and for any reason not

including treatment failure. This analysis will be similar in methodology to the primary efficacy analysis.

5.2.3. Model Diagnostics

To assess the appropriateness of the proportional hazards assumption, a log-log survival plot of the Kaplan-Meier estimates will be generated. If the proportional hazards assumption is correct, this plot should present approximately parallel lines corresponding to the two treatment groups. Cumulative sums of Schoenfeld residuals[8] over time may also be used to assess the proportional hazards assumption.

5.2.4. Subgroup Analyses

Additional analysis of the time to treatment failure will be also be evaluated by age groups, race, sex, baseline BMI group, and randomization factors. Separate Cox's proportional hazard models will be used to individually assess the effect of these covariates (including treatment and one covariate at a time).

In Part II, primary time to treatment failure summaries will be repeated for the following subgroups: 1-) Current Primary DSM-5 Diagnosis (schizophrenia and schizophreniform); 2-) Subjects who continued into Part III and those who did not.

5.2.5. Time to Treatment Failure by Reason

The reasons for treatment failure will be summarized at each visit and endpoint. Exploratory time to event endpoints of interest include:

- Time to 1st psychiatric hospitalization.
- Time to 1st any deliberate self-injury, suicidal ideation or behavior, homicidal ideation or violent behavior.
- Time to 1st new arrest/incarceration.
- Time to 1st discontinuation of antipsychotic treatment due to inadequate efficacy.
- Time to 1st discontinuation of antipsychotic treatment due to safety or tolerability.
- Time to 1st treatment supplementation with another antipsychotic due to inadequate efficacy.
- Time to 1st increase in the level of psychiatric services.

The same analysis methods described for the primary efficacy endpoint will be used to analyze the above exploratory time to event endpoints. No multiplicity adjustment will be made to these tests.

5.2.6. Multiple Events

Multiple treatment failures from the same subject will be analyzed as recurrent events. The cumulative mean function (CMF)[9] of treatment failure is a function of time and will be defined as the expected number of treatment failures per subject in a given time interval since randomization. A proportional rates/means model including a term for treatment will be used to compare the CMFs of treatment failure between the two treatment groups. Multiple treatment failure related data points will be captured from the Treatment Failure Board data set.

5.3. Part II, Key Secondary Endpoints

The key secondary hypotheses to be tested in Part II of this study are to demonstrate that 9 months' treatment of subjects with recent-onset schizophrenia or schizophreniform disorder with PP is superior to 9 months' treatment with OAP in:

- improving or maintaining cognition (as measured by the change in MCCB composite score from baseline)
  - *The key secondary efficacy null hypothesis is that there is no difference in mean change from Part II baseline to the Part II end point in the MCCB composite score between PP and OAP*
- maintaining functioning (as measured by time to at least 7-point worsening in the PSP total score from baseline)
  - *The key secondary efficacy null hypothesis is that there is no difference in time to 7-point worsening in PSP total score in Part II between PP and OAP*
- increasing or preserving brain ICM volume of the frontal lobe as compared to baseline
  - *The key secondary efficacy null hypothesis is that there is no difference in mean change in ICM volume of the frontal lobe from Part II baseline to the Part II end point between PP and OAP*

At the end of Part II, changes in MCCB composite score, PSP, and ICM volume will be analyzed to determine if there is an early treatment effect demonstrating disease progression of PP relative to OAP. The early treatment effect is illustrated by δ_{21} in **Error! Reference source not found.**.

In Part II, the overall Type I error rate for testing PP versus OAP for both the primary efficacy endpoint and key secondary efficacy endpoints will be controlled at the 2-sided 0.05 significance level using a combination of fixed sequence gatekeeper approach and Holm's step-down procedure. Time to first treatment failure will be tested first, followed by change in key secondary endpoints (MCCB, PSP, and ICM). Time to first treatment failure will be examined

first using the log-rank test statistics. If the null hypothesis corresponding to time to first treatment failure is rejected, then the key secondary endpoints will be tested at the 5% level using Holm's step down procedure, thus maintaining an overall Type I error rate of 5%. If the primary null hypothesis is not rejected, testing of change in key secondary endpoints will still be performed, but no unqualified statements about the statistical significance regarding change will be made. In Holm's procedure, p-values from analyses of MCCB, PSP, and ICM will be ordered from lowest to highest. Let $p(1)<p(2)<p(3)$ to be order p-values corresponding null hypotheses H(1), H(2), and H(3). In Step 1, if $p(1) < 0.05/3$ then corresponding null hypothesis will be rejected and testing will be examined in second step; otherwise, none of the hypotheses will be rejected and we will stop testing. In Step 2, if $p(2) < 0.05/2$ then corresponding null hypothesis will be rejected and move onto third step; otherwise, H(2) and H(3) will not be rejected and we will stop testing. In Step 3, p(3) will be tested at 0.05 level.

5.3.1. MCCB – MATRICS Consensus Cognitive Battery

5.3.1.1. Definition

The MATRICS Consensus Cognitive Battery (MCCB) was developed to provide a relatively brief evaluation of key cognitive domains relevant to schizophrenia and related disorders and is recommended as the standard outcome measure for clinical trials of cognition-enhancing drugs for schizophrenia. The MCCB includes 10 tests that measure 7 cognitive domains.

MCCB Tests

Cognitive Domain	Test	Description
Speed of processing	*Brief Assessment of Cognition in Schizophrenia* (BACS): Symbol-Coding	Timed paper-and-pencil test in which respondent uses a key to write digits that correspond to nonsense symbols
	Category Fluency: Animal Naming	Oral test in which respondent names as many animals as she/he can in 1 minute
	Trail Making Test: Part A	Timed paper-and-pencil test in which respondent draws a line to connect consecutively numbered circles placed irregularly on a sheet of paper
Attention/Vigilance	*Continuous Performance Test*-Identical Pairs (CPT-IP)	Computer-administered measure of sustained attention in which respondent presses a response button to consecutive matching numbers
Working memory (nonverbal)	*Wechsler Memory Scale*®-3rd Ed. (WMS®-III): Spatial Span	Using a board on which 10 cubes are irregularly spaced, respondent taps cubes in same (or reverse) sequence as test administrator
(verbal)	*Letter-Number Span*	Orally administered test in which respondent mentally reorders strings of number and letters and repeats them to administrator
Verbal learning	*Hopkins Verbal Learning Test-Revised*™ (HVLT-R™)	Orally administered test in which a list of 12 words from three taxonomic categories is presented and the respondent is asked to recall as many as possible after each of three learning trials
Visual Learning	*Brief Visuospatial Memory Test-Revised (BVMT-R™)*	A test that involves reproducing six geometric figures from memory
Reasoning and problem solving	*Neuropsychological Assessment Battery*®	Seven timed paper-and-pencil mazes of increasing difficulty that measure

	(NAB®): Mazes	foresight and planning
Social cognition	*Mayer-Salovey-Caruso Emotional Intelligence Test* (MSCEIT™): Managing Emotions	Paper-and-pencil multiple-choice test that assesses how people manage their emotions

The MCCB will be administered by a qualified rater at the time points indicated in the Time and Events Schedule. The MCCB assessment must also be performed at the time of first occurrence of treatment failure.

For every individual subject, MCCB should be administered at the same time of day (within a 2-hour window) for every assessment. If possible, the same person should administer the tests at each occasion.

All MCCB data will be sent to central raters for review. Assessments will each be scored by central raters who are blinded to treatment information. Central raters will also review the neurocognitive data quality and correct any errors. The central raters will enter the final scores into the MCCB Scoring program to derive T-scores and composites scores. The central raters will enter the raw and derived MCCB scores to the eDC.

5.3.1.2. Analysis Methods

Part II

The change from baseline in MCCB composite score will be analyzed using a mixed model repeated measures (MMRM) ANCOVA model[10,11,12,13]. The analysis will be based on observed data, ie, data collected at each time point without carrying forward previous values. The data points from unscheduled visits will also be included. The response variable will be the change in MCCB composite score. The model will include Part II baseline MCCB composite score as a fixed-effect covariate; treatment (PP and OAP) and country and time as fixed-effect (categorical) factors, and the interaction between time and treatment. Using this model, treatment effects at the Month 9 end point will be estimated based on differences between least squares (LS) means. Accompanying 95% CIs for the LS mean differences between PP and OAP will be presented. An unstructured matrix will be used for the covariance of the within-subject repeated measures as a base case.

The purpose of this analysis is to examine treatment differences at Month 9 using LS means from the MMRM model. The treatment effect measures the deviation from the hypothesis of equality of means among treatments, “averaged” over the treatment duration. The time effect is a measure of deviation from the hypothesis of constancy of mean response over time for all treatment groups combined after aggregating over subjects. The treatment-by-time interaction tests the hypothesis of parallel response profiles over time in the treatment groups. A significant

treatment-by-time interaction means that changes in response over time differ among treatments; in other words, there is a difference among the treatment groups, but the magnitude of the difference varies over time. The interaction will remain in the model, regardless of significance, in order to obtain an estimate of the treatment effect at the Month 9 time point. In the ITT analysis set, the model-based test for the mean treatment group difference will also be examined.

To assess normality, diagnostic plots such as normal quantile-quantile plots of residuals will be created. If a high degree of non-normality is suspected, remedies such as rank-based methods will be considered.

The eITT analysis set will also be used to test changes at end point. The eITT analysis set consists of MCCB score assessed at or prior to the eITT end point.

Similar analysis will also be repeated using the observed scores (not the change) and domain scores.

In addition, the actual values and change from baseline scores will be summarized for both observed data as well as LOCF data by treatment group. The treatment group differences will be analyzed using an ANCOVA model. The model will include treatment and country as fixed effect design factors, and baseline MCCB composite score as a covariate. Using this model, treatment effects will be estimated based on differences between least squares (LS) means. Accompanying 95% CIs (not adjusted for multiplicity) for the LS mean differences will be presented. Within treatment group difference for change from baseline will be evaluated using a paired t-test. Plots of the LS-Mean (± standard error (SE)) change from baseline in MCCB total score over time will also be presented for both the observed and LOCF data.

Primary MCCB composite score summaries will be repeated for the following subgroups: 1-) Current Primary DSM-5 Diagnosis (schizophrenia and schizophreniform); 2-) Subjects who continued into Part III and those who did not.

5.3.2. PSP - Personal and Social Performance Scale

5.3.2.1. Definition

The PSP scale is a clinician-rated instrument that assesses the degree of dysfunction a subject exhibits in the past month within 4 domains of behavior: a) socially useful activities, b) personal and social relationships, c) self-care, and d) disturbing and aggressive behavior. The results of the assessment are converted to a numerical score following the PSP scoring guideline. A score between 71 and 100 indicates a mild degree of dysfunction; a score between 31 and 70 indicates varying degrees of difficulty, and a subject with a score of ≤30 has functioning so poor that he or she requires intensive supervision.

The PSP scale is to be administered by a qualified rater at the time points indicated in the Time and Events Schedule. The PSP assessment must also be conducted at the time of first occurrence of treatment failure, and should also be performed at the time of subsequent occurrences of treatment failure. If possible, the same person should administer this scale at each occasion.

5.3.2.2. Analysis Methods

Part II

The analysis of time to at least 7-point worsening in PSP total score will be similar in methodology to the primary efficacy analysis.

Subject status during Part II	Time to **First** >= 7-point worsening in PSP Total Score	Censoring indicator
Randomized subjects who experienced at least 7-point worsening during Part II	(Date of event – Part II start date) + 1	No
Randomized subjects who remained event free at the end of the Part II	(End of Part II date – Part II start date) + 1	Yes
Early withdrawal/discontinued during the Part II without an event	(Date of early withdrawal – Part II start date) + 1	Yes

Mean changes in PSP total score will be analyzed using the same method as the MCCB analysis except the baseline MCCB score will be replaced by the corresponding PSP baseline score in MMRM and ANCOVA models. Primary PSP summaries will be repeated for the following subgroups: 1-) Current Primary DSM-5 Diagnosis (schizophrenia and schizophreniform); 2-) Subjects who continued into Part III and those who did not.

Time to 10-point worsening in PSP total score will also be analyzed.

Frequency counts, percentages, and cumulative percentages of subjects reporting each PSP deciles, 3 PSP categories (Poor, Variable, Good), and domain levels will be summarized for both observed data and LOCF data by treatment group. In addition to 3 PSP categories, subjects who achieve a PSP score $\geq$ 71 vs. <70 will be identified and summarized. The incidence of PSP responders (PSP score $\geq$ 71 vs. <70) will be compared between active treatment group and placebo. Differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH)[14] test using the modified ridit scores, stratifying on site.

At each post-baseline assessment and LOCF, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline PSP categories. Within group differences will be evaluated using the McNemar's test.

Each PSP domain was assessed on a 6-point severity scale of dysfunction: 1 = absent, 2 = mild, 3 = manifest, 4 = marked, 5 = severe, and 6 = very severe. We define no impairment to mild impairment as a PSP domain score of absent or mild dysfunction; moderate to severe impairment is defined as manifest, marked, severe, or very severe dysfunction. For these binomial summaries, differences between treatment groups will be evaluated based on the Cochran-

Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site. At each post-baseline assessment and LOCF, a shift from baseline table will summarize the changes relative to baseline PSP categories. Within group differences will be evaluated using the McNemar's test.

5.3.3. MRI Brain Imaging Assessments

5.3.3.1. Definition

Selected sites will be performing MRI brain imaging. Approximately half of the enrolled subjects will undergo MRI scans. MRI scans are optional to subjects. All images will be sent to a central site for analysis. The imaging raters at the central site will be blind to clinical characteristics, treatment assignment, and demographic characteristics of subjects.

Brain imaging assessments will be performed in subjects with schizophrenia/schizophreniform disorder and in healthy control subjects. Subjects with schizophrenia or schizophreniform disorder who are not able to undergo scans (eg, unable to fit or difficulty fitting in MRI instrument, or MRI contraindicated due to presence of metallic objects [pacemaker, etc.]) will be excluded from participating in MRI scans, but will still be eligible to participate in all other aspects of this study. Healthy control subjects who are not able to undergo MRI scans will not be eligible to participate in this study.

If the initial MRI is clinically abnormal and shows presence of a severe brain abnormality that would preclude analyses (eg, large hemangioma), the individual will be excluded from undergoing further MRI scans. If this occurs in a healthy control subject, the subject will be withdrawn from the study. If the initial scan is abnormal in a subject with schizophrenia/schizophreniform disorder, the subject will be excluded from further MRI assessments but will continue in all other aspects of the study.

A separate MRI manual will be provided to the relevant sites with detailed information regarding the MRI procedures. In brief, brain ICM volume will be measured by inversion recovery (IR) and spin echo (SE) MRI sequences focused on the frontal lobe. Cortical thickness, gray matter and white matter volumes will be measured by 3D MPRAGE MRI. Ventricular volume and intrasulcal CSF will be measured by SE MRI sequences. The subcortical myelin will be measured by MRI sequences optimized for DTI. The resting state fMRI will also be measured.

Brain imaging in subjects with schizophrenia or schizophreniform disorder will be performed at the time points indicated in the Time and Events Schedule. Brain imaging must also be conducted at the time of first occurrence of treatment failure or as soon as possible thereafter.

MRI assessments in healthy control subjects will be described in the MRI manual.

The MRI data set will include both raw and adjusted measures. Adjusted Intracortical Myelin Fraction (ICM_div_ICV_adj) will be considered as primary variable of interest.

Variable Name	Description	Notes
ICV_cc	Intracranial volume	Data units are cubic centimeters (cm3)
Total_cc	Total brain tissue volume (WM+GM)	Data units are cubic centimeters (cm3)
IR_cc	WM volume measured on IR image	Data units are cubic centimeters (cm3)
PD_cc	WM volume measured on PD image	Data units are cubic centimeters (cm3)
ICM_old_cc	Intracortical myelin volume (old version)	Data units are cubic centimeters (cm3)
ICM_cc	Intracortical myelin volume	Data units are cubic centimeters (cm3)
Total_div_ICV	Total brain tissue fraction (WM+GM)	Data as fraction of intracranial volume
IR_div_ICV	WM fraction measured on IR image	Data as fraction of intracranial volume
PD_div_ICV	WM fraction measured on PD image	Data as fraction of intracranial volume
ICM_old_div_ICV	Intracortical myelin fraction (old version)	Data as fraction of intracranial volume
ICM_div_ICV	Intracortical myelin fraction	Data as fraction of intracranial volume
ICV_cc_adj	Adjusted Intracranial volume	Volume after adjusting using HC group (e.g. age, machine drift)

5.3.3.2. Analysis Methods

Part II

The number and percentage of subjects who completed the trial, prematurely discontinued from the trial along with the reason for discontinuation, will be summarized based on the Part II MRI ITT analysis set. Demographics baseline characteristics and psychiatric history will also be presented for the MRI ITT Population. Correlations between the MRI variables and MCCB composite score including domain scores will be listed.

The actual values and change from baseline for each variable will be summarized for both observed data as well as the LOCF data. Box plots will be generated at each time point for observed and change scores separately by treatment group.

Both the change from baseline and observed scores in MRI variables will be analyzed using a mixed model repeated measures (MMRM) ANCOVA model. These models will be repeated for both raw scores and adjusted scores. The analysis will be based on observed data, ie, data collected at each time point without carrying forward previous values. The data points from unscheduled visits will also be included. The response variable will be the change or observed score in MRI scores. The model will include corresponding Part II baseline as a fixed-effect covariate; treatment (PP and OAP) and country and time as fixed-effect (categorical) factors, and the interaction between time and treatment. Using this model, treatment effects at the Month 9 end point will be estimated based on differences between least squares (LS) means. Accompanying 95% CIs for the LS mean differences between PP and OAP will be presented. An unstructured matrix will be used for the covariance of the within-subject repeated measures as a base case.

To assess normality, diagnostic plots such as normal quantile-quantile plots of residuals will be created. If a high degree of non-normality is suspected, remedies such as rank-based methods will be considered.

In addition, the actual values and change from baseline scores will be summarized for both observed data as well as LOCF data by treatment group. The treatment group differences will be analyzed using an ANCOVA model. The model will include treatment and country as fixed effect design factors, and corresponding baseline MRI score as a covariate. Using this model, treatment effects will be estimated based on differences between least squares (LS) means. Accompanying 95% CIs (not adjusted for multiplicity) for the LS mean differences will be presented. Within treatment group difference for change from baseline will be evaluated using a paired t-test. Plots of the LS-Mean (± standard error (SE)) change from baseline in MRI variables over time will also be presented.

The eITT analysis set will also be used to test changes at end point for the primary MRI variable. The eITT analysis set consists of MRI variables assessed at or prior to the eITT end point.

5.4. Secondary Endpoints

No multiplicity adjustments will be made for secondary endpoints. All p-values are considered nominal.

5.4.1. Treatment Failure

5.4.1.1. Treatment Failure in EDP Phase

Subjects will be evaluated at the time points indicated in the Time and Events Schedule for the occurrence of treatment failures. Only treatment failure events occurring after randomization in Part II until the end of Part III will be assessed.

Note that subjects who experience a treatment failure and do not withdraw consent or meet the criteria for withdrawal from the study will continue in the study and be followed through to the end of the study.

Subject status during EDP Phase	Time to **First** Treatment Failure/Censoring	Censoring indicator
Randomized subjects who experienced treatment failure during EDP Phase	(Date of treatment failure – Part II start date) + 1	No
Randomized subjects who remained treatment failure-free at the end of the Part III	(End of Part III date – Part II start date) + 1	Yes
Early withdrawal/discontinued during the Part III without treatment failure	(Date of early withdrawal in Part III – Part II start date) + 1	Yes

Treatment differences will be compared using a log-rank test. The cumulative distribution function of the time to treatment failure will be estimated by the Kaplan-Meier method. The 95% CIs for the median treatment failure rates, as well as the failure rates at months 3, 6, 9, 12, 15, and 18 will be provided. Standard Error (SE) estimates will be computed using Greenwood's formula. Group differences will be evaluated using the log-rank test. The estimate of the hazard ratio and its 95% CI will be provided by treatment group based on the Cox proportional hazards model. The reasons for first treatment failure and subsequent treatment failures will be summarized at each visit and end point.

To assess the appropriateness of the proportional hazards assumption, a log-log survival plot of the Kaplan-Meier estimates will be generated. If the proportional hazards assumption is correct, this plot should present approximately parallel lines corresponding to the two treatment groups. Cumulative sums of Schoenfeld residuals over time may also be used to assess the proportional hazards assumption.

Additional analysis of the time to treatment failure will be also be evaluated by age groups, race, sex, and baseline BMI group. Separate Cox's proportional hazard models will be used to individually assess the effect of these covariates (including treatment and one covariate at a time).

The reasons for treatment failure will be summarized at each visit and endpoint. These time to event endpoints of interest include:

- Time to 1st psychiatric hospitalization.
- Time to 1st any deliberate self-injury, suicidal ideation or behavior, homicidal ideation or violent behavior.
- Time to 1st new arrest/incarceration.
- Time to 1st discontinuation of antipsychotic treatment due to inadequate efficacy.
- Time to 1st discontinuation of antipsychotic treatment due to safety or tolerability.
- Time to 1st treatment supplementation with another antipsychotic due to inadequate efficacy.
- Time to 1st increase in the level of psychiatric services.

For each of the reasons for treatment failure, the Kaplan-Meier estimates, log-rank test, hazard ratios will be provided to differentiate treatment groups.

Multiple treatment failures from the same subject will be analyzed as recurrent events. The cumulative mean function (CMF) of treatment failure is a function of time and will be defined as the expected number of treatment failures per subject in a given time interval since randomization. A proportional rates/means model including a term for treatment will be used to compare the CMFs of treatment failure between the two treatment groups. Multiple treatment failure related data points will be captured from the Treatment Failure Board data set.

5.4.1.2. Treatment Failure in Part III

Subjects will be evaluated at the time points indicated in the Time and Events Schedule for the occurrence of treatment failures. Only treatment failure events occurring after Part III start date will be assessed.

Note that subjects who experience a treatment failure and do not withdraw consent or meet the criteria for withdrawal from the study will continue in the study and be followed through to the end of the study.

Subject status during Part III	Time to **First** Treatment Failure/Censoring	Censoring indicator
Subjects who experienced treatment failure during Part III	(Date of treatment failure – Part III start date) + 1	No
Subjects who remained treatment failure-free at the end of the Part III	(End of Part III date – Part III start date) + 1	Yes
Early withdrawal/discontinued during the Part III without treatment failure	(Date of early withdrawal in Part III – Part III start date) + 1	Yes

In Part III, there are three treatment groups, PPPP, OAPPP, and OAPOAP. Treatment differences among three groups will be compared using a log-rank test. The cumulative distribution function of the time to treatment failure will be estimated by the Kaplan-Meier method. The 95% CIs for the median treatment failure rates, as well as the failure rates at months 3, 6, and 9 will be provided. Standard Error (SE) estimates will be computed using Greenwood's formula. Group differences will be evaluated using the log-rank test. The estimate of the hazard ratio and its 95% CI will be provided by treatment group based on the Cox proportional hazards model. The reasons for first treatment failure and subsequent treatment failures will be summarized at each visit and end point.

To assess the appropriateness of the proportional hazards assumption, a log-log survival plot of the Kaplan-Meier estimates will be generated. If the proportional hazards assumption is correct, this plot should present approximately parallel lines corresponding to the two treatment groups. Cumulative sums of Schoenfeld residuals over time may also be used to assess the proportional hazards assumption.

Additional analysis of the time to treatment failure will be also be evaluated by age groups, race, sex, and baseline BMI group. Separate Cox's proportional hazard models will be used to individually assess the effect of these covariates (including treatment and one covariate at a time).

The reasons for treatment failure will be summarized at each visit and endpoint. These time to event endpoints of interest include:

- Time to 1st psychiatric hospitalization.
- Time to 1st any deliberate self-injury, suicidal ideation or behavior, homicidal ideation or violent behavior.
- Time to 1st new arrest/incarceration.
- Time to 1st discontinuation of antipsychotic treatment due to inadequate efficacy.
- Time to 1st discontinuation of antipsychotic treatment due to safety or tolerability.

- Time to 1st treatment supplementation with another antipsychotic due to inadequate efficacy.

- Time to 1st increase in the level of psychiatric services.

For each of the reasons for treatment failure, the Kaplan-Meier estimates, log-rank test, hazard ratios will be provided to differentiate treatment groups.

Multiple treatment failures from the same subject will be analyzed as recurrent events. The cumulative mean function (CMF) of treatment failure is a function of time and will be defined as the expected number of treatment failures per subject in a given time interval since randomization. A proportional rates/means model including a term for treatment will be used to compare the CMFs of treatment failure between the two treatment groups. Multiple treatment failure related data points will be captured from the Treatment Failure Board data set.

5.4.2. MCCB – MATRICS Consensus Cognitive Battery

5.4.2.1. Definition

Definition is listed in Section 5.3.3.1.

5.4.2.2. Analysis Methods

Part I

The actual values and change from baseline in MCCB composite score and domain scores will be summarized for both observed data as well as LOCF data. Within treatment group difference for change from baseline will be evaluated using a paired t-test. The MCCB composite score will also be categorized into three groups as <25, 25-35, >35. Frequency counts, percentages, and cumulative percentages of subjects reporting each MCCB composite score categories will be summarized. Shifts from baseline to each visit including LOCF by categories will also be displayed.

Part II

Part II analysis is listed as a segment of the Key Secondary Efficacy analysis section, see Section 5.3.

Treatment group differences will also be evaluated by number of treatment failures in Part II.

EDP Phase

The change from baseline in MCCB composite score and domain scores will be analyzed using a mixed model repeated measures (MMRM) ANCOVA model. The analysis will be based on observed data, ie, data collected at each time point without carrying forward previous values. The data points from unscheduled visits will also be included. The response variable will be the change in MCCB composite score or domain scores. The model will include Part II baseline MCCB composite score as a fixed-effect covariate; treatment (PPPP and OAPOAP) and country

and time as fixed-effect (categorical) factors, and the interaction between time and treatment. Using this model, treatment effects at each visit and at Month 18 end point will be estimated based on differences between least squares (LS) means. Accompanying 95% CIs for the LS mean differences between PPPP and OAPOAP will be presented. An unstructured matrix will be used for the covariance of the within-subject repeated measures as a base case.

The eITT analysis set will also be used to test changes at end point. The eITT analysis set consists of MCCB score assessed at or prior to the eITT end point.

Similar analysis will also be repeated using the observed scores (not the change) and domain scores.

In addition, the actual values and change from baseline scores will be summarized for both observed data as well as LOCF data by treatment group. The treatment group differences will be analyzed using an ANCOVA model. The model will include treatment and country as fixed effect design factors, and baseline MCCB composite score or domain score as a covariate. Using this model, treatment effects will be estimated based on differences between least squares (LS) means. Accompanying 95% CIs (not adjusted for multiplicity) for the LS mean differences will be presented. Within treatment group difference for change from baseline will be evaluated using a paired t-test. Plots of the LS-Mean (± standard error (SE)) change from baseline in MCCB total score over time will also be presented for both the observed and LOCF data. Treatment group differences will also be evaluated by number of treatment failures in EDP phase.

Part III

All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

Part III analyses of MCCB composite and domain scores will be analyzed using the Part II analyses strategies.

For the MCCB composite score, the following hypotheses will also be tested as a function of estimated changes in the Part II endpoint using the observed scores to demonstrate disease modification. This analysis assumes at minimum that subjects treated with PP for 9 months (early-start group) demonstrated better outcomes on MCCB composite score than those treated with OAP for 9 months at the end of Part II. Let δ_{21} denote the estimated mean differences between treatment groups at the end of Part II, treatment effect on disease progression. Let δ_{31} denote the estimated mean difference between PPPP and OAPOP groups at the end of Part III. The following hypotheses will be tested using observed scores:

H_0: $\delta_{31} - 0.5.\ \delta_{21} \leq 0$ and H_0: $\delta_{31} - 0.75.\ \delta_{21} \leq 0$ where the fractions 0.5 and 0.75 are the conditional noninferiority margins which indicate the ratio of treatment effect in Part III compared to the Part II observed differences.

These hypotheses will be tested using a mixed model repeated measures (MMRM) ANCOVA models. The analysis will be based on observed data. The model will include corresponding Part II baseline as a fixed-effect covariate; treatment and country and time as fixed-effect (categorical) factors, and the interaction between time and treatment. Using this model, treatment effects at the Month 9 and Mont 18 end point will be estimated based on differences between least squares (LS) means. An unstructured matrix will be used for the covariance of the within-subject repeated measures.

5.4.3. PSP - Personal and Social Performance Scale

5.4.3.1. Definition

Definition is listed in Section 5.3.2.1.

5.4.3.2. Analysis Methods

Part I

The actual values and change from baseline in PSP score will be summarized for both observed data as well as LOCF data. Within treatment group difference for change from baseline will be evaluated using a paired t-test. Frequency counts, percentages, and cumulative percentages of subjects reporting each PSP deciles and domain level will be summarized. Additionally, at each post-baseline assessment and LOCF data, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline for the PSP 10-point categories. Worsening in the PSP change from baseline score is defined as a decrease by one 10-point scale or more, and an improvement in the PSP change from baseline score is defined as an increase by one 10-point scale or more. No testing will be done for this shift summary. The PSP score will also be categorized into three groups as Poor (PSP score ≤30), Variable (>30 PSP score ≤70), and Good (PSP score >70). Frequency counts, percentages, and cumulative percentages of subjects reporting each PSP categories will be summarized. Shifts from baseline to each visit including LOCF by PSP categories will also be listed.

Part II

Part II analysis is listed as a segment of the Key Secondary Efficacy analysis section, see Section 5.3.

Treatment group differences will also be evaluated by number of treatment failures in Part II.

EDP Phase

The analysis of time to at least 7-point worsening in PSP total score will be similar in methodology to the primary efficacy analysis.

Subject status during EDP Phase	Time to **First** >= 7-point worsening in PSP Total Score	Censoring indicator
Randomized subjects who	(Date of event – Part II start date) + 1	No

experienced at least 7-point worsening during EDP phase		
Randomized subjects who remained event free at the end of the Part III	(End of Part III date – Part II start date) + 1	Yes
Early withdrawal/discontinued during the Part III without an event	(Date of early withdrawal in Part III – Part II start date) + 1	Yes

Mean changes in PSP total score will be analyzed using the same method as the MCCB analysis except the baseline MCCB score will be replaced by the corresponding PSP baseline score. Treatment group differences will also be evaluated by number of treatment failures in EDP phase.

Time to 10-point worsening in PSP total score will also be analyzed.

Frequency counts, percentages, and cumulative percentages of subjects reporting each PSP deciles, 3 PSP categories (Poor, Variable, Good), and domain levels will be summarized for both observed data and LOCF data by treatment group. In addition to 3 PSP categories, subjects who achieve a PSP score ≥ 71 vs. <70 will be identified and summarized. The incidence of PSP responders (PSP score ≥ 71 vs. <70) will be compared between active treatment group and placebo. Differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site.

At each post-baseline assessment and LOCF, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline PSP categories. Within group differences will be evaluated using the McNemar's test.

Each PSP domain was assessed on a 6-point severity scale of dysfunction: 1 = absent, 2 = mild, 3 = manifest, 4 = marked, 5 = severe, and 6 = very severe. We define no impairment to mild impairment as a PSP domain score of absent or mild dysfunction; moderate to severe impairment is defined as manifest, marked, severe, or very severe dysfunction. For these binomial summaries, differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site. At each post-baseline assessment and LOCF, a shift from baseline table will summarize the changes relative to baseline PSP categories. Within group differences will be evaluated using the McNemar's test.

Part III

All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

The analysis of time to at least 7-point worsening in PSP total score will be similar in methodology to the primary efficacy analysis.

Subject status during Part III	Time to **First** >= 7-point worsening in PSP	Censoring indicator

	Total Score	
Subjects who experienced at least 7-point worsening during Part III	(Date of event – Part III start date) + 1	No
Subjects who remained event free at the end of the Part III	(End of Part III date – Part III start date) + 1	Yes
Early withdrawal/discontinued during the Part III without an event	(Date of early withdrawal in Part III – Part III start date) + 1	Yes

Mean changes in PSP total score will be analyzed using the same method as the MCCB analysis except the baseline MCCB score will be replaced by the corresponding PSP baseline score.

Time to 10-point worsening in PSP total score will also be analyzed.

Frequency counts, percentages, and cumulative percentages of subjects reporting each PSP deciles, 3 PSP categories (Poor, Variable, Good), and domain levels will be summarized for both observed data and LOCF data by treatment group. In addition to 3 PSP categories, subjects who achieve a PSP score $\geq$ 71 vs. <70 will be identified and summarized. The incidence of PSP responders (PSP score $\geq$ 71 vs. <70) will be compared between active treatment group and placebo. Differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site.

At each post-baseline assessment and LOCF, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline PSP categories. Within group differences will be evaluated using the McNemar's test.

Each PSP domain was assessed on a 6-point severity scale of dysfunction: 1 = absent, 2 = mild, 3 = manifest, 4 = marked, 5 = severe, and 6 = very severe. We define no impairment to mild impairment as a PSP domain score of absent or mild dysfunction; moderate to severe impairment is defined as manifest, marked, severe, or very severe dysfunction. For these binomial summaries, differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site. At each post-baseline assessment and LOCF, a shift from baseline table will summarize the changes relative to baseline PSP categories. Within group differences will be evaluated using the McNemar's test.

For the PSP total score, the following hypotheses will also be tested as a function of estimated changes in the Part II endpoint using the observed scores to demonstrate disease modification. This analysis assumes at minimum that subjects treated with PP for 9 months (early-start group) demonstrated better outcomes on PSP total score than those treated with OAP for 9 months at the end of Part II. Let δ_{21} denote the estimated mean differences between treatment groups at the end of Part II, treatment effect on disease progression. Let δ_{31} denote the estimated mean difference between PPPP and OAPOP groups at the end of Part III. The following hypotheses will be tested using observed scores:

H_0: $\delta_{31} - 0.5.\ \delta_{21} \leq 0$ and H_0: $\delta_{31} - 0.75.\ \delta_{21} \leq 0$ where the fractions 0.5 and 0.75 are the conditional noninferiority margins which indicate the ratio of treatment effect in Part III compared to the Part II observed differences.

These hypotheses will be tested using a mixed model repeated measures (MMRM) ANCOVA models. The analysis will be based on observed data. The model will include corresponding Part II baseline as a fixed-effect covariate; treatment and country and time as fixed-effect (categorical) factors, and the interaction between time and treatment. Using this model, treatment effects at the Month 9 and Mont 18 end point will be estimated based on differences between least squares (LS) means. An unstructured matrix will be used for the covariance of the within-subject repeated measures.

5.4.4. MRI Brain Imaging Assessments

5.4.4.1. Definition

Definition is listed in Section 5.3.3.1.

5.4.4.2. Analysis Methods

Part I

There is no imaging data collected in Part I.

Part II

Part II analysis is listed as a segment of the Key Secondary Efficacy analysis section, see Section 5.3.

EDP Phase

Changes in ICM volume will be analyzed using the same method as the MCCB analysis except for the baseline MCCB score which will be replaced by the corresponding ICM baseline score. Treatment groups will be listed as PPPP and OAPOAP. The objective is to examine the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment.

Part III

All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

Part III analyses of MRI scores will be analyzed using the Part II analyses strategies.

For the Adjusted Intracortical Myelin Fraction (ICM_div_ICV_adj) score, the following hypotheses will also be tested as a function of estimated changes in the Part II endpoint using the observed scores to demonstrate disease modification. This analysis assumes at minimum that subjects treated with PP for 9 months (early-start group) demonstrated better outcomes on Adjusted Intracortical Myelin Fraction score than those treated with OAP for 9 months at the end of Part II. Let δ_{21} denote the estimated mean differences between treatment groups at the end of

Part II, treatment effect on disease progression. Let δ_{31} denote the estimated mean difference between PPPP and OAPOP groups at the end of Part III. The following hypotheses will be tested using observed scores:

H_0: $\delta_{31} - 0.5.\ \delta_{21} \leq 0$ and H_0: $\delta_{31} - 0.75.\ \delta_{21} \leq 0$ where the fractions 0.5 and 0.75 are the conditional noninferiority margins which indicate the ratio of treatment effect in Part III compared to the Part II observed differences.

These hypotheses will be tested using a mixed model repeated measures (MMRM) ANCOVA models. The analysis will be based on observed data. The model will include corresponding Part II baseline as a fixed-effect covariate; treatment and country and time as fixed-effect (categorical) factors, and the interaction between time and treatment. Using this model, treatment effects at the Month 9 and Mont 18 end point will be estimated based on differences between least squares (LS) means. An unstructured matrix will be used for the covariance of the within-subject repeated measures.

5.4.5. CGI-S - Clinical Global Impression of Severity Scale

5.4.5.1. Definition

The CGI-S rating scale is used to rate the severity of a subject's overall clinical condition on a 7-point scale ranging from 1 (not ill) to 7 (among the most extremely ill). This scale permits a global evaluation of the subject's condition at a given time.

This scale will be administered by a qualified rater at the time points indicated in the Time and Events Schedule. The CGI-S assessment must also be conducted at the time of first occurrence of treatment failure and should also be performed at the time of subsequent occurrences of treatment failure. If possible, the same person should administer this scale at each occasion.

5.4.5.2. Analysis Methods

Part I

The actual values and change from baseline in CGI-S score will be summarized for both observed data as well as LOCF data. Within treatment group difference for change from baseline will be evaluated using a paired t-test. Frequency counts, percentages, and cumulative percentages of subjects reporting each CGI-S level will be summarized for both observed data as well as LOCF data.

Part II

The actual values and change from baseline scores of CGI-S will be summarized for both observed data as well as LOCF data by treatment group. The treatment group differences will be analyzed using an ANCOVA model. The model will include treatment and country as fixed effect design factors, and the baseline CGI-S score as a covariate. Using this model, treatment effects will be estimated based on differences between least squares (LS) means. Accompanying 95% CIs (not adjusted for multiplicity) for the LS mean differences will be presented. Within

treatment group difference for change from baseline will be evaluated using a paired t-test. Plots of the LS-Mean (± standard error (SE)) change from baseline in CGI-S score over time will also be presented for the observed data. The assumptions of the ANOVA / ANCOVA models--errors are normally distributed with equal variances--will be tested by examining the model residuals with histograms, normal probability plot and plots of residuals versus fitted values. If there is strong evidence that the assumptions are not satisfied then the p-value from Friedman's Chi-Square Test will also be presented for the ANOVA case, and the p-value from the stratified Mantel-Haenszel test for the ANCOVA case, using the methodologies described in Stokes, Davis, and Koch (1997)[15]. The ordinal observed repeated measures data will also be evaluated using Generalized Estimation Equations (GEE)[15].

Frequency counts, percentages, and cumulative percentages of subjects reporting CGI-S level will be summarized for both observed data as well as LOCF data by treatment group. Differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site.

At each post-baseline assessment and LOCF data, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline for the CGI-S. Worsening in the CGI-S change from baseline score is defined as a decrease by one grade or more, and an improvement in the CGI-S change from baseline score is defined as an increase by one grade or more.

EDP Phase

CGI-S score will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP and OAPOAP. The objective is to examine the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment.

Part III

CGI-S score will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP, OAPPP, and OAPOAP. All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

5.4.6. CRDPSS – Clinician-Rated Dimensions of Psychosis Symptom Severity

5.4.6.1. Definition

The CRDPSS is an 8-item measure that assesses the severity of mental health symptoms that are important across psychotic disorders, including delusions, hallucinations, disorganized speech, abnormal psychomotor behavior, negative symptoms (ie, restricted emotional expression or avolition), impaired cognition, depression, and mania. The severity of these symptoms can

predict important aspects of the illness, such as the degree of cognitive and/or neurobiological deficits. This measure, developed by the APA, is intended to capture meaningful variation in the severity of symptoms, which may help with treatment planning, prognostic decision-making, and research on pathophysiological mechanisms. Each item asks the clinician to rate the severity of each symptom as experienced by the individual during the past 7 days.

Each item on the measure is rated on a 5-point scale (0=none; 1=equivocal; 2=present, but mild; 3=present and moderate; and 4=present and severe) with a symptom-specific definition of each rating level. The clinician may review all of the individual's available information and, based on clinical judgment, select the level that most accurately describes the severity of the individual's condition. The clinician then indicates the score for each item in the "Score" column provided. The response on each item should be interpreted independently when assessing the severity of the psychotic disorder.

The CRDPSS will be completed by a qualified rater at the time points indicated in the Time and Events Schedule. The CRDPSS assessment must also be conducted at the time of first occurrence of treatment failure and should also be performed at the time of subsequent occurrences of treatment failure.

5.4.6.2. Analysis Methods

Part I

Frequency counts, percentages, and cumulative percentages of subjects reporting each CRDPSS level will be summarized for both observed data as well as LOCF data.

Part II

Frequency counts, percentages, and cumulative percentages of subjects reporting each CRDPSS level will be summarized for both observed data as well as LOCF data by treatment group. Differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site.

At each post-baseline assessment and LOCF data, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline for the each CRDSS item. Worsening in the CRDSS item change from baseline score is defined as a decrease by one grade or more, and an improvement in the CRDSS item change from baseline score is defined as an increase by one grade or more. No testing will be done for this shift summary.

EDP Phase

CRDPSS items will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP and OAPOAP. The objective is to examine the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment.

Part III

CRDPSS items will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP, OAPPP, and OAPOAP. All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

5.4.7. MSQ – Medication Satisfaction Questionnaire

5.4.7.1. Definition

Medication satisfaction will be assessed using the MSQ. The MSQ is a self-administered single-item questionnaire with responses on a 7-point Likert-scale as follows: 1=extremely dissatisfied, 2=very dissatisfied, 3=somewhat dissatisfied, 4=neither satisfied nor dissatisfied, 5=somewhat satisfied, 6=very satisfied, 7=extremely satisfied. The MSQ has demonstrated acceptable reliability and validity,[Error! Reference source not found.] making this single-item questionnaire appropriate and easy to use in clinical research. A 1-point change on the MSQ may be considered clinically meaningful.

The MSQ will be completed by the subject at the time points indicated in the Time and Events Schedule. The MSQ must also be conducted at the time of first occurrence of treatment failure, and should also be performed at the time of subsequent occurrences of treatment failure.

5.4.7.2. Analysis Methods

Part I

The actual values and change from baseline in MSQ score will be summarized for both observed data as well as LOCF data. Within treatment group difference for change from baseline will be evaluated using a paired t-test. Frequency counts, percentages, and cumulative percentages of subjects reporting each MSQ level will be summarized for both observed data as well as LOCF data.

Part II

The actual values and change from baseline scores of MSQ score will be summarized for both observed data as well as LOCF data by treatment group. The treatment group differences will be analyzed using an ANCOVA model. The model will include treatment and country as fixed effect design factors, and the baseline MSQ score as a covariate. Within treatment group difference for change from baseline will be evaluated using a paired t-test. The assumptions of the ANOVA / ANCOVA models--errors are normally distributed with equal variances--will be tested. If there is strong evidence that the assumptions are not satisfied then the p-value from Friedman's Chi-Square Test will also be presented for the ANOVA case, and the p-value from the stratified Mantel-Haenszel test for the ANCOVA case. The ordinal repeated measures data will also be evaluated using Generalized Estimation Equations (GEE).

Frequency counts, percentages, and cumulative percentages of subjects reporting each MSQ levels will be summarized for both observed data and LOCF data by treatment group.

Differences between treatment groups will be evaluated based on the Cochran-Mantel-Haenszel mean score (CMH) test using the modified ridit scores, stratifying on site. At each post-baseline assessment and LOCF, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline MSQ categories. Within group differences for shifts will be evaluated using the McNemar's test.

Responses for MSQ will also be dichotomized as Satisfied and Dissatisfied when the observed response on the MSQ scale was between 1 and 4 and between 5 and 7, respectively. The proportion of subjects who were either satisfied or dissatisfied on the MSQ scale will be calculated at each visit and compared between treatment groups using a CMH test. Within group differences for satisfaction vs. dissatisfaction will be evaluated using the McNemar's test.

EDP Phase

MSQ score will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP and OAPOAP. The objective is to examine the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment.

Part III

MSQ score will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP, OAPPP, and OAPOAP. All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

5.5. Part II, EDP Phase Exploratory Assessments

Exploratory evaluations include assessment of medical resource use (based on the RUQ), and goal setting and daily activity evaluations.

5.5.1. RUQ - Resource Utilization Questionnaire

5.5.1.1. Definition

The RUQ will be used in the study as an exploratory tool to assess utilization of resources, such as number of hospitalization days (refers to ≥1 overnight stay), emergency room visits without hospitalization, day or night clinic stays, and outpatient treatment, as well as daily living conditions and productivity of the subject. Healthcare resource utilization will be assessed through regular questioning of subject's resource utilization during visits (via the RUQ) and will be objectively verified through medical records and emergency/crisis center documents obtained by investigative site staff. Work and living status will be assessed at baseline and at regular intervals throughout the study.

5.5.1.2. Analysis Methods

Part I

Descriptive statistics, counts, and percentages will be presented at baseline and at end point. The NA values will not be considered in the examination of RUQ assessments for all categorical and non-ordinal measure data summaries. The calculated percentages will be determined using only patients who provided an applicable answer to the referred questions.

All cause and psychiatric-related hospitalizations will be summarized by reason, by the total number hospitalizations (none, once, twice, and >twice), and overall. For emergency room visits, the number and percent of subjects with ER visits will be summarized by reasons of visit, by the total number of visits (none, once, twice and >twice), and by overall. Data points for day clinic or day treatment programs, outpatient services, daily living, and productivity will also be summarized.

Part II

Descriptive statistics, counts, and percentages will be presented at each assessment time point, end point and each LOCF visit. The NA values will not be considered in the examination of RUQ assessments for all categorical and non-ordinal measure data summaries. The calculated percentages will be determined using only patients who provided an applicable answer to the referred questions.

All cause and psychiatric-related hospitalizations will be summarized by reason, by the total number hospitalizations (none, once, twice, and >twice), and by overall. Hospitalizations due to social reasons will be excluded. The cumulative distribution function of time from Day 1 to first hospitalization for psychiatric reasons will be estimated by the Kaplan-Meier method. The null hypothesis that there is no difference between treatment groups in the time to hospitalization over 9 months will be tested using a log-rank test. In addition, descriptive analysis will also be conducted by hospital types and by the type of ward or unit during hospitalization. Since date of admission and discharge are collected for hospitalization, length of stay (LOS) will be derived and summarized descriptively.

For emergency room visits, the number and percent of subjects with ER visits will be summarized overall, by reasons of visit and by the total number of visits (none, once, twice and >twice) per treatment group. The cumulative distribution function of time from Day 1 to first ER visit for psychiatric reasons will be estimated by the Kaplan-Meier method. The null hypothesis that there is no difference between treatment groups in the time to ER distributions over 9 months will be tested using a log-rank test.

Data points for day clinic or day treatment programs, outpatient services, daily living, and productivity will also be summarized by treatment group.

EDP Phase

RUQ data points will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP and OAPOAP. The objective is to examine the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment.

Part III

RUQ data will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP, OAPPP, and OAPOAP. All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP). Baseline summary for prior hospitalizations and ER visit will go back to 12 month prior to Part III start date.

5.5.2. Goal Setting and Daily Activity Evaluations

5.5.2.1. Definition

Subjects will set up to 3 personal goals and 3 health goals at the beginning of each treatment phase, which will be reviewed with the subject and by the treating clinician.

The patient goals are not binding, but progress towards attaining these goals will be followed throughout the study at the time points indicated in the Time and Events Schedule.

Patient Happiness Assessment and Goal Setting Preparation

The Patient Happiness Assessment and Goal Setting Preparation documents will be reviewed and completed by each subject at the Part I Baseline visit (Visit 2).

Patient Goal Setting Documentation

The treating clinician will review the Goal Setting Preparation document with the subject and agree with up to 3 personal goals and up to 3 health goals for the subject at the start of each treatment phase. These goals will be documented on paper and shared with the subject and the subject's designated individual. The original document will be filed at the investigative (study) site as part of the subject's study source record. The patient goals are not binding.

Patient Goals Attainment

Progress towards attaining the patient goals will be assessed by the subject at the time points indicated in the Time and Events Schedule. Satisfaction regarding the subject's progress towards meeting the goals will be rated on a Likert scale ranging from 'Extremely Dissatisfied' to 'Extremely Satisfied'.

Quantitative Assessment of Daily Activities

The Quantitative Assessment of Daily Activities is a patient reported outcome that documents the time a subject spends in a broad array of common daily activities. Categories include sleep and rest, self-care, work, recreation and social activities. The period covered is the prior week. Refer to the Time and Event Schedule for frequency of assessment.

5.5.2.2. Analysis Methods

At each assessment time point and LOCF data, descriptive statistics will be presented for each of the domains for all the phases.

6. SAFETY

Safety variables to be analyzed include treatment-emergent adverse events, serious adverse events, ESRS-A scores, laboratory parameters, vital signs, physical examination reports, and ISST-Plus. The ITT analysis sets will be used for analyses performed on safety parameters unless otherwise specified. Any statistical tests performed to explore the data will be used only to highlight any comparisons that may warrant further consideration.

6.1. Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) (version 18.1 or above)[16] will be used to classify AEs by system organ class and preferred term. Treatment-emergent adverse events (TEAEs) that occurred in each study phase will be summarized by system organ class and preferred term.

A treatment-emergent adverse event (TEAE) is an event that is new in onset or increased in severity following treatment initiation. An event that starts prior to and ends after the initiation of study drug (Day 1 of each study phase) is to be considered treatment-emergent only if the severity increases after the start of medication. Treatment-emergent flags will be defined separately for Part I, Part II, EDP, and Part III.

Treatment-emergent AE in the Part I are defined as the adverse events with onset date on or after the first study date (onset date >= Part I Start Date and <= Part I End Date) or increase in severity during the same period. An event that starts prior to the Part I Start Date and ends afterwards will be considered treatment emergent in the Part I Phase only if the severity increases on or after the Part I Start Date.

Treatment-emergent AE in the Part II are defined as the adverse events with onset date on or after the first study date (onset date >= Part II Start Date and <= Part II End Date) or increase in severity during the same period. An event that starts prior to the Part II Start Date and ends afterwards will be considered treatment emergent in the Part II Phase only if the severity increases on or after the Part II Start Date.

Treatment-emergent AE in the Part III are defined as the adverse events with onset date on or after the first study date (onset date >= Part III Start Date and <= Part III End Date) or increase in severity during the same period. An event that starts prior to the Part III Start Date and ends afterwards will be considered treatment emergent in the Part III Phase only if the severity increases on or after the Part III Start Date.

Treatment-emergent AE in the EDP phase are defined as the adverse events with onset date on or after the first study date (onset date >= Part II Start Date and <= Part III End Date) or increase in severity during the same period. An event that starts prior to the Part II Start Date and ends afterwards will be considered treatment emergent in the EDP only if the severity increases on or after the Part II Start Date. This phase simply contains information from both Parts II and III.

AEs will be included as TEAEs 30 days after the permanent discontinuation of the study medication.

The number of subjects with at least one TEAE as well as the number of subjects and total number of events for each preferred term and SOC will be summarized regardless of severity and relationship to study medication. For each subject, multiple reports of events that map to a common MedDRA preferred term and SOC will be condensed into a single AE for incidence counts, while each occurrence will be counted for the number of events. Within each summary, TEAEs will be sorted alphabetically.

Severity for AEs is classified in the case report form (CRF) as ‘mild’, ‘moderate’, or ‘severe’. Incidence of TEAEs will be summarized by preferred term, SOC, and maximum severity for each treatment group. If a subject has more than one AE within a preferred term and SOC, the subject will be counted at most once by the maximum severity. Adverse events with missing severity will be classified as ‘not specified’ and will be considered the least severe in cases where a subject has the same event with a specified severity.

Relationship to study medication for AEs is classified in the CRF as ‘not related’, ‘doubtful’, ‘possible’, ‘probable’, or ‘very likely’. Incidence of TEAEs will be summarized by the preferred term, SOC, and maximum relationship for each treatment group. If a subject has more than one AE within a preferred term and SOC, the subject will be counted at most once by the maximum relationship to study medication. Adverse events with missing relationship will be classified as ‘not specified’ and will be considered the least related in cases where a subject has the same event with a specified relationship.

An additional summary will display incidence of TEAEs by ‘not related’ and ‘related’. The responses of ‘not related’ and ‘doubtful’ will be pooled to create the ‘not related’ to study medication group category. The following three responses (‘possible’, ‘probable’, or ‘very likely’) will be pooled to create the ‘related’ to study medication group category. If an event has missing relationship it will be classified as ‘not specified’ and will not be considered ‘related’ or ‘not related’.

Action taken for AEs is classified in the CRF as 'dose increased', 'dose not changed', 'drug interrupted, 'drug withdrawn', and 'not applicable'. Incidence of TEAEs will be summarized by the preferred term, SOC, and action taken for each treatment group. If a subject has more than one AE within a preferred term and SOC, the subject will be counted at most once by the worst outcome. Adverse events with missing action taken will be classified as 'not specified'.

The incidence of TEAEs will be presented for subjects who discontinue the study for an AE, as well as for subjects who experience a SAE. These AE summaries will be summarized in each study phase separately.

Listings of subjects who discontinue the study for an AE and subjects who experience a SAE will be presented.

Summaries of treatment-emergent adverse events for categories of clinical interest will be provided as discussed in the following sections. The preferred terms for each category are given in Attachment 1.

6.1.1. EPS-Related Adverse Events

Treatment-emergent AEs that are related to extrapyramidal symptoms (EPS) will be summarized. The EPS AEs will be categorized into 5 subgroups (tremor, dystonia, hyperkinesia, parkinsonism, and dyskinesia) that include the following MedDRA preferred terms:

Tremor (preferred terms: Tremor, Essential tremor, Intention tremor)

Dystonia (preferred terms: Oculogyration, Oculogyric crisis, Trismus, Tongue spasm, Tongue paralysis, Cervical spasm, Emprosthotonus, Myotonia, Pleurothotonus, Risus sardonicus, Muscle spasms, Blepharospasm, Dystonia, Opisthotonus, Torticollis, Facial spasm, Muscle contracture)

Hyperkinesia (preferred terms: Akathisia, Hyperkinesia, Periodic limb movement disorder, Restless legs syndrome, Restlessness)

Parkinsonism (preferred terms: Hypertonia, Bradykinesia, Cogwheel rigidity, Drooling, Musculoskeletal stiffness, Akinesia, Hypokinesia, Nuchal rigidity, Parkinsonian gait, Parkinsonian rest tremor, Parkinsonism, Muscle rigidity, Muscle tightness, Glabellar reflex abnormal, On and off phenomenon, Parkinson's disease, Parkinsonian crisis, Extrapyramidal disorder, Masked facies)

Dyskinesia (preferred terms: Dyskinesia, Muscle contractions involuntary, Movement disorder, Muscle twitching, Athetosis, Chorea, Choreoathetosis, Tardive dyskinesia, Myoclonus, Protrusion tongue, Rabbit syndrome, Buccoglossal syndrome).

The incidence for each EPS subgroup will be calculated, and no overall incidence for the EPS-related AEs will be presented.

6.1.2. Diabetes Mellitus and Hyperglycaemia-Related Adverse Events

Treatment-emergent adverse events that may be associated with diabetes mellitus and hyperglycaemia will be summarized. MedDRA preferred terms related to diabetes mellitus and hyperglycaemia are defined as follows:

Acquired lipoatrophic diabetes, Diabetic hepatopathy, Fulminant type 1 diabetes mellitus, Hyperglycaemic seizure, Hyperglycaemic unconsciousness, Type 3 diabetes mellitus, Blood 1,5-anhydroglucitol decreased, Blood glucose increased, Diabetes complicating pregnancy, Diabetes mellitus, Diabetes mellitus inadequate control, Diabetes with hyperosmolarity, Diabetic coma, Diabetic hyperglycaemic coma, Diabetic hyperosmolar coma, Diabetic ketoacidosis, Diabetic ketoacidotic hyperglycaemic coma, Fructosamine increased, Gestational diabetes, Glucose tolerance impaired, Glucose tolerance impaired in pregnancy, Glucose urine present, Glycosuria, Glycosuria during pregnancy, Glycosylated haemoglobin increased, Hyperglycaemia, Hyperglycaemic hyperosmolar nonketotic syndrome, Impaired fasting glucose, Insulin resistance, Insulin resistance syndrome, Insulin resistant diabetes, Insulin-requiring type 2 diabetes mellitus, Ketoacidosis, Ketonuria, Ketosis, Latent autoimmune diabetes in adults, Metabolic syndrome, Neonatal diabetes mellitus, Pancreatogenous diabetes, Type 1 diabetes mellitus, Type 2 diabetes mellitus, Urine ketone body present, Diabetes mellitus insulin-dependent, Diabetes mellitus non-insulin-dependent, Insulin-requiring type II diabetes mellitus.

6.1.3. Potentially Prolactin-Related Adverse Events

Treatment-emergent adverse events that may be associated with changes in serum prolactin levels will be summarized. MedDRA preferred terms considered as being potentially related to serum prolactin levels are defined as below:

Amenorrhoea, Amenorrhoea-galactorrhoea syndrome, Galactorrhoea, Gynaecomastia, Hyperprolactinaemia, Oligomenorrhoea, Blood prolactin increased, Anorgasmia, Ejaculation delayed, Ejaculation disorder, Erectile dysfunction, Female sexual dysfunction, Libido decreased, Libido disorder, Loss of libido, Male sexual dysfunction, Orgasm abnormal, Orgasmic sensation decreased, Sexual dysfunction, Breast discharge, Breast enlargement, Breast pain, Prolactin-producing pituitary tumour, Prolactinoma, Blood prolactin, Blood prolactin abnormal, Breast tenderness, Hypertrophy breast, Menstruation irregular.

These adverse events will also be tabulated separately by sex.

6.1.4. Other Adverse Events of Special Interest

Incidence of other treatment-emergent adverse events of clinical interest will be presented. Search terms relevant to the adverse events of clinical importance are listed in Attachment 1. These terms are classified into the following groups:

Suicidality, Aggression and Agitation, Somnolence and Sedation, Seizures and Convulsions, Neuroleptic Malignant Syndrome, Cardiac Arrhythmias, Orthostatic Hypotension, Adverse Events Suggestive of Proarrhythmic Potential, Ischemia-related, Potential Rhabdomyolysis-related, Overdose-related, Weight Gain-related, Tachycardia-related, and Injection-site Related.

6.2. Clinical Laboratory

Descriptive statistics (N, mean, median, minimum, maximum and range) for values and changes from baseline will be provided for clinical laboratory tests (hematology, chemistry and urinalysis) for all the phases.

Clinical laboratories will be summarized at the following visits: Part I Baseline, Part I End Point, Part II Baseline, Part II End Point, Part III Baseline, Part III End Point, EDP Baseline, and EDP End Point. Baseline for the Part II period will be defined as the last Part I data point on or before the randomization date, similar to other efficacy and safety scales. If there are no Part I post-baseline laboratory assessments the baseline for the Part II period will consist of data from the screening period. Part II baseline will serve as the baseline score for the EDP Phase.

Clinical laboratory tests that meet the criteria for markedly abnormal will be listed by subject for each treatment phase. The incidence of treatment-emergent markedly abnormal laboratory values that occurred at any time during each treatment phase will be presented. Clinical laboratory test values will be considered "TEMA using the criteria defined by the Sponsor (Janssen Research & Development, LLC)" listed in Attachment 2. The identification of TEMA laboratory values is based on the postbaseline value being out of range while the baseline value (defined above) is either missing or within the range given in Attachment 2. If post-baseline laboratory results are above the upper limit and the baseline value is below the lower limit, then the post-baseline abnormality will also be considered TEMA. The same applies to the postbaseline value being below the lower limit with the baseline value being above the upper limit.

For prolactin laboratory results, values over time and treatment-emergent abnormal results based on the laboratory reference range will be presented by sex.

6.3. Vital Signs, Weight, and BMI

Vital signs will be assessed at each visit, and will consist of body temperature, sitting blood pressure, and pulse. Body weight will also be assessed at each visit. Height will be recorded at screening only. Height will be recorded at screening only. BMI will be calculated from measurements of height and weight, $(kg)/(height\ (m))^2$.

Actual and change from baseline in vital signs will be summarized at each assessment time point including LOCF end points within each phase. Data summaries will be listed by treatment group.

For body weight, the incidence of increases/decreases from baseline to endpoint by ≥7% will be summarized for each study period using the corresponding baseline score. For each of the vital signs parameters, the following categories for abnormality will be tabulated and presented with percentages by treatment group at each assessment time point and at end point for all the study

phases. The tabulation of weight abnormality classes will be repeated by baseline BMI category (OL Baseline) [Normal=<25; Overwieght= 25 - <30; Obese= ≥30].

Treatment-emergent abnormality categories for vital signs are defined as follows:

	Period Post-baseline value outside of normal limit if:	
	Abnormally low	Abnormally high
Pulse (bpm)	A decrease from baseline of ≥ 15 to a value ≤ 50	An increase from baseline of ≥ 15 to a value ≥ 100
Systolic BP (mmHg)	A decrease from baseline of ≥ 20 to a value ≤ 90	An increase from baseline of ≥ 20 to a value ≥ 180
Diastolic BP (mmHg)	A decrease from baseline of ≥ 15 to a value ≤ 50	An increase from baseline of ≥ 15 to a value ≥ 105
Body Weight (kg)	A decrease from baseline of ≥7%	An increase from baseline of ≥7%

BP = blood pressure

Physical examinations will be performed at screening and the end of study, including examination of general appearance, skin, neck (including thyroid), eyes, ears, nose, throat, heart, lungs, abdomen, lymph nodes, extremities, and nervous system. The number and percent of subjects with abnormal values will be summarized for each system at baseline and study endpoint.

6.4. ESRS-A - Extrapyramidal Symptoms Rating Scale – Abbreviated

The ESRS-A is an abbreviated manualized version of the ESRS, a semistructured interview that rates parkinsonian symptoms, dystonia, dyskinesias, and akathisia over the previous 7 days. The ratings include a motor examination for rigidity, tremor, reduced facial expression or speech, impaired gait/posture, postural instability, and bradykinesia/hypokinesia. Twenty-four individual items are rated on a 6-point scale: 0=Absent, 1=Minimal, 2=Mild, 3=Moderate, 4=Severe, or 5=Extreme. Frequency is included as an index of severity. Symptoms are divided into the 4 corresponding subscales and each subscale is summarized in a Clinical Global Impression of Movement Severity (CGI-MS) score.

A qualified clinician is required to administer the ESRS-A. A trained clinician rater for these scales must meet one the following criteria: a physician (M.D. or D.O.), advanced practice nurse (NP), physician assistant (PA), or other allied health professional who is trained and holds a valid license to perform physical examinations.

The ESRS-A will be administered at the time points indicated in the Time and Events Schedule. If possible, the same person should administer these scales at all visits.

Part I

CGI-MS scores will be presented for parkinsonism, dystonia, dyskinesia, and akathisia. The actual values and change from screening in numeric CGI-MS scores will be summarized for both observed data as well as LOCF data.

Part II

The actual values and change from baseline scores of CGI-MS domains will be summarized for both observed data as well as LOCF data by treatment group in each phase separately. Frequency counts, percentages, and cumulative percentages of subjects reporting each CGI-MS level will be summarized for both observed data as well as LOCF data by treatment group. At each post-baseline assessment and LOCF, a shift from baseline table will summarize the worsening, no change, and improvement relative to baseline for the CGI-MS. Worsening in the CGI-MS change from baseline score is defined as a decrease by one grade or more, and an improvement in the CGI-MS change from baseline score is defined as an increase by one grade or more. No testing will be done for this shift summary. Individual ESRS-A items for each domain will be summarized using frequency counts, percentages, and cumulative percentages of subjects reporting each level of the score within each phase as well.

EDP Phase

ESRS-A items will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP and OAPOAP. The objective is to examine the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment.

Part III

ESRS-A items will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP, OAPPP, and OAPOAP. All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

6.5. ISST-Plus Short Form - InterSePT Scale for Suicidal Thinking – Plus

The ISST-Plus is a clinician-rated 4-part instrument for collecting data on suicidal thinking and behavior. The ISST-Plus Short Form is an abbreviated form of the ISST-Plus used to screen subjects between ISST-Plus assessments. The ISST-Plus meets the requirements recently announced by FDA for assessment of suicidality.

The ISST-Plus includes 4 parts:

- **Part I** is designed to collect information on the severity of suicidal ideation during the ***7 days*** (or other pre-defined time frame) prior to the subject's visit. It is comprised of 13 items, with three levels of severity: 0 (none), 1 (weak), or 2 (moderate or strong).

- **Part II** is designed to collect information on suicidal behaviors that have occurred ***since the last visit or the last assessment of suicidal behavior***.
- **Part III** is a global rating of suicide ideation and behavior or status ***at the time of the subject interview*** as judged by a clinically experienced rater. It includes and integrates both Part I and Part II of the scale. It should only be completed after Parts I and II have been completed and should take all available information into consideration.
- **Part IV** is collected at the end of the study to record whether or not the subject died by suicide during the study.

Rating the ISST-Plus requires completion of a semi-structured interview. Subjects who exhibit a high suicide risk as evidenced by a total score of ≥7 on the ISST-Plus Part I or a score of 2 on items 7, 10, or 11, or an ISST-Plus Part III score of ≥2 must be evaluated by either a clinical psychologist or psychiatrist, and this must be documented in the source documents. If a subject answers in the affirmative (ie, "YES") to any part of the suicidal behaviors section (Part II) of the ISST-Plus, completion of a detailed potential suicide attempt narrative is required.

The ISST-Plus Short Form will be administered at the study visits indicated in the Time and Events Schedule. If suicidality is identified (ie, 'yes' is answered to questions 1.0, 2.2, 2.3, or 2.4 of the ISST-Plus Short Form) the full ISST-Plus must be administered in its entirety.

Part I

ISST-Plus item scores will be presented for both observed data as well as LOCF data. Frequency counts and percentages of subjects reporting each item level will be summarized. Additionally, Total Score of Suicidal Thinking Part will be computed. The actual values and change from baseline in total score will be summarized. Within group difference for change from baseline will be evaluated using a paired t-test.

Part II

Frequency counts and percentages of subjects reporting each item level will be summarized for both observed data as well as LOCF data by treatment group.

Total Score of Suicidal Thinking Part will be computed. The actual values and change from baseline in total score will be summarized. Within treatment group difference for change from baseline will be evaluated using a paired t-test. The actual values and change from baseline scores of total score will be summarized. The treatment group differences will be analyzed using an ANCOVA model. The model will include treatment and country as fixed effect design factors, and the baseline score as a covariate.

EDP Phase

ISST-Plus items will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP and OAPOAP. The objective is to examine the cumulative effect on extended disease progression at the end of Part III following early start for PP compared to continuing OAP after a total of 18 months of treatment.

Part III

ISST-Plus items will be analyzed using the same method as the Part II analysis. Treatment groups will be listed as PPPP, OAPPP, and OAPOAP. All data from the Part III phase of the study will be categorized and presented using three columns: those patients who are not re-randomized (PPPP) were on PP in Part II, those oral patients in Part II who are re-randomized to paliperidone palmitate (OAPPP), and those patients who are re-randomized to oral antipsychotics (OAPOAP).

REFERENCES

1. WHO drug dictionary. WHO Collaborating Centre for International Drug Monitoring. Uppsala, Sweden, (Box 26, S-751 03), 1993. http://www.who-umc.org/
2. Demitrienko A, Tamhane A, Bretz F. Multiple testing problems in pharmaceutical statistics. 2010, Chapman&Hall/CRC Biostatistics Series.
3. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association*, 1958: 53; 457-81*12.
4. Kalbfleisch, J.D. and Prentice, R.L. The Statistical Analysis of Failure Time Data. New York: John Wiley & Sons, Inc.; 1980.
5. Cox DR. Regression models and life tables. J Royal Stat Soc 1972; 34: 87-220.
6. Taylor M., Murray S., Hsu C-H. Survival estimation and testing via multiple imputation. Statistics and Probability letters. 2002; 58:221-232.
7. Lipkovich I., Ratitch, B., Kelly, M. Sensitivity to censored-at-random assumption in the analysis of time-to-event endpoints. Pharmaceutical Statistics. 2016; 15:216-229.
8. Schoenfeld D. Chi-squared goodness of fit test for the proportional hazards regression model. Biometrika 1981, 67:147-53.
9. Nelson, W. B. (2003), Recurrent Events Analysis for Product Repairs, Disease Recurrences, and Other Applications, The ASA-SIAM Series on Statistics and Applied Probability.
10. Siddiqui O, James Hung HM, O'Neill R. MMRM vs. LOCF: Comprehensive Comparison Based on Simultion Study and 25 NDA Datasets. Journal of Biopharmaceutical Statistics. 2009, 19:2, 227-246.
11. Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. Springer-Verlag, 2000.
12. Diggle P, Heagerty P, Lian KY, and Zeger S. Analysis of Longitudinal Data, Second Edition. Oxford University Press, USA. August 2002
13. Neter J, Wasserman W, Kutner M. Applied linear statistical models. Richard D. Irwin, Inc., Homewood, Illinois, 1990: 1037-1049
14. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 1959: 22; 719-48.
15. Stokes, Maura E., Davis, Charles S., Koch, Gary G., Categorical Data Analysis Using the SAS® System, Cary, NC: SAS Institute Inc., 1997 (3rd Printing).
16. Medical Dictionary for Regulatory Activities (MedDRA), Version 18.1. IFPMA, International Committee on Harmonization. Geneva, Switzerland.

LIST OF ATTACHMENTS

ATTACHMENT 1: Special Interest Adverse Events

MAGCAT	AEDECOD
AGITATION	Agitation
AGITATION	Psychomotor agitation
AGITATION	Psychomotor hyperactivity
AGGRESSION	Aggression
AGGRESSION	Homicidal ideation
AGGRESSION	Hostility
AGGRESSION	Homicide

MCARCAT	AEDECOD
CARDIOVASCULAR	Torsade de pointes
CARDIOVASCULAR	Sudden death
CARDIOVASCULAR	Ventricular tachycardia
CARDIOVASCULAR	Ventricular fibrillation
CARDIOVASCULAR	Ventricular flutter
CARDIOVASCULAR	Syncope
CARDIOVASCULAR	Presyncope
CARDIOVASCULAR	Syncope vasovagal

MISCCAT	AEDECOD
ISCHAEMIA	Acute coronary syndrome
ISCHAEMIA	Acute myocardial infarction
ISCHAEMIA	Angina pectoris
ISCHAEMIA	Angina unstable
ISCHAEMIA	Subendocardial ischaemia
ISCHAEMIA	Myocardial ischaemia
ISCHAEMIA	Cardiac ischaemia
ISCHAEMIA	Coronary artery disease
ISCHAEMIA	Coronary artery insufficiency
ISCHAEMIA	Myocardial infarction
ISCHAEMIA	Myocardial ischaemia
ISCHAEMIA	Papillary muscle infarction
ISCHAEMIA	Postinfarction angina
ISCHAEMIA	Prinzmetal angina
ISCHAEMIA	Silent myocardial infarction
ISCHAEMIA	Subendocardial ischaemia
ISCHAEMIA	Amaurosis fugax
ISCHAEMIA	Brain stem infarction
ISCHAEMIA	Brain stem ischaemia
ISCHAEMIA	Cerebellar infarction
ISCHAEMIA	Cerebral infarction
ISCHAEMIA	Cerebral ischaemia
ISCHAEMIA	Cerebrovascular accident
ISCHAEMIA	Cerebrovascular disorder

ISCHAEMIA	Cerebrovascular insufficiency
ISCHAEMIA	Embolic cerebral infarction
ISCHAEMIA	Embolic stroke
ISCHAEMIA	Haemorrhagic cerebral infarction
ISCHAEMIA	Haemorrhagic stroke
ISCHAEMIA	Ischaemic cerebral infarction
ISCHAEMIA	Ischaemic stroke
ISCHAEMIA	Subclavian steal syndrome
ISCHAEMIA	Thromboembolic stroke
ISCHAEMIA	Thrombotic stroke
ISCHAEMIA	Lacunar infarction
ISCHAEMIA	Reversible ischaemic neurological deficit
ISCHAEMIA	Transient ischaemic attack
ISCHAEMIA	Vascular encephalopathy
ISCHAEMIA	Vertebrobasilar insufficiency
ISCHAEMIA	Ischaemia
ISCHAEMIA	Ischaemic cardiomyopathy
ISCHAEMIA	Thrombotic cerebral infarction
ISCHAEMIA	Cerebral microangiopathy
ISCHAEMIA	Cerebellar ischaemia

MORTHCAT	**AEDECOD**
Orthostatic Hypotension	Blood pressure orthostatic abnormal
Orthostatic Hypotension	Blood pressure orthostatic decreased
Orthostatic Hypotension	Dizziness postural
Orthostatic Hypotension	Orthostatic hypotension
Orthostatic Hypotension	Orthostatic intolerance
Orthostatic Hypotension	Orthostatic heart rate response increased

MOVCAT	**AEDECOD**
Overdose	Accidental overdose
Overdose	Intentional overdose
Overdose	Multiple drug overdose
Overdose	Multiple drug overdose accidental
Overdose	Multiple drug overdose intentional
Overdose	Overdose

MQTCAT	**AEDECOD**
TORSADE DE POINTES	Torsade de pointes
SUDDEN DEATH	Cardiac arrest
SUDDEN DEATH	Cardiac death
SUDDEN DEATH	Cardio-respiratory arrest
SUDDEN DEATH	Sudden death
SUDDEN DEATH	Sudden cardiac death
SUDDEN DEATH	Ventricular asystole
VENTRICULAR TACHYCARDIA	Accelerated idioventricular rhythm

VENTRICULAR TACHYCARDIA	Ventricular tachycardia
VENTRICULAR TACHYCARDIA	Ventricular tachyarrhythmia
VENTRICULAR FIBRILLATION AND FLUTTER	Cardiac fibrillation
VENTRICULAR FIBRILLATION AND FLUTTER	Ventricular arrhythmia
VENTRICULAR FIBRILLATION AND FLUTTER	Ventricular fibrillation
VENTRICULAR FIBRILLATION AND FLUTTER	Ventricular flutter
SYNCOPE	Circulatory collapse
SYNCOPE	Loss of consciousness
SYNCOPE	Syncope
SYNCOPE	Presyncope
SYNCOPE	Syncope vasovagal
SEIZURES	Clonic convulsion
SEIZURES	Convulsion
SEIZURES	Epilepsy
SEIZURES	Grand mal convulsion
SEIZURES	Tonic clonic movements
SEIZURES	Tonic convulsion
SEIZURES	Status epilepticus
SEIZURES	Atonic seizures
SEIZURES	Complex partial seizures
SEIZURES	Generalised non-convulsive epilepsy

MRHACAT	**AEDECOD**
RHABDOMYOLYSIS	Rhabdomyolysis
RHABDOMYOLYSIS	Blood creatine phosphokinase increased
RHABDOMYOLYSIS	Myoglobinuria
RHABDOMYOLYSIS	Myoglobin urine

MSEIZCAT	**AEDECOD**
SEIZURES	Acquired epileptic aphasia
SEIZURES	Alcoholic seizure
SEIZURES	Atonic seizures
SEIZURES	Atypical benign partial epilepsy
SEIZURES	Automatism epileptic
SEIZURES	Baltic myoclonic epilepsy
SEIZURES	Clonic convulsion
SEIZURES	Complex partial seizures
SEIZURES	Convulsion
SEIZURES	Convulsion in childhood
SEIZURES	Convulsion neonatal
SEIZURES	Convulsions local
SEIZURES	Convulsive threshold lowered
SEIZURES	Deja vu
SEIZURES	Dreamy state
SEIZURES	Drug withdrawal convulsions
SEIZURES	Eclampsia
SEIZURES	Epilepsy
SEIZURES	Epilepsy congenital

SEIZURES	Epileptic aura
SEIZURES	Epileptic psychosis
SEIZURES	Febrile convulsion
SEIZURES	Frontal lobe epilepsy
SEIZURES	Generalised non-convulsive epilepsy
SEIZURES	Grand mal convulsion
SEIZURES	Hypoglycaemic seizure
SEIZURES	Infantile spasms
SEIZURES	Lafora's myoclonic epilepsy
SEIZURES	Lennox-Gastaut syndrome
SEIZURES	Myoclonic epilepsy
SEIZURES	Myoclonic epilepsy and ragged-red fibres
SEIZURES	Myoclonic epilepsy and ragged-red fibers
SEIZURES	Partial seizures
SEIZURES	Partial seizures with secondary generalisation
SEIZURES	Petit mal epilepsy
SEIZURES	Post-traumatic epilepsy
SEIZURES	Psychomotor seizures
SEIZURES	Seizure anoxic
SEIZURES	Simple partial seizures
SEIZURES	Status epilepticus
SEIZURES	Sudden unexplained death in epilepsy
SEIZURES	Temporal lobe epilepsy
SEIZURES	Tonic clonic movements
SEIZURES	Tonic convulsion
SEIZURES	Uncinate fits

MSOMCAT	**AEDECOD**
SOMNOLENCE	Somnolence
SOMNOLENCE	Sedation
SOMNOLENCE	Lethargy
SOMNOLENCE	Hypersomnia

MSUICAT	**AEDECOD**
SUICIDALITY	Depression suicidal
SUICIDALITY	Intentional overdose
SUICIDALITY	Intentional self-injury
SUICIDALITY	Multiple drug overdose intentional
SUICIDALITY	Poisoning deliberate
SUICIDALITY	Self injurious behavior
SUICIDALITY	Self-injurious ideation
SUICIDALITY	Completed suicide
SUICIDALITY	Suicidal ideation
SUICIDALITY	Suicide attempt
SUICIDALITY	Suicidal behavior

MNMSCAT	AEDECOD
NMS	Hyperthermia malignant
NMS	Neuroleptic malignant syndrome
NMS	Serotonin syndrome
NMS	Body temperature increased
NMS	Hyperpyrexia
NMS	Pyrexia
NMS	Catatonia
NMS	Dyskinesia
NMS	Dystonia
NMS	Freezing phenomenon
NMS	Hyperkinesia
NMS	Hypertonia
NMS	Muscle necrosis
NMS	Muscle rigidity
NMS	Oculogyric crisis
NMS	Oculogyration
NMS	Opisthotonus
NMS	Rhabdomyolysis
NMS	Altered state of consciousness
NMS	Autonomic nervous system imbalance
NMS	Blood creatine phosphokinase abnormal
NMS	Blood creatine phosphokinase increased
NMS	Blood creatine phosphokinase MM increased
NMS	Blood pressure abnormal
NMS	Blood pressure decreased
NMS	Blood pressure fluctuation
NMS	Blood pressure increased
NMS	Cardiovascular insufficiency
NMS	Coma
NMS	Confusional state
NMS	Consciousness fluctuating
NMS	Delirium
NMS	Depressed level of consciousness
NMS	Disorientation
NMS	Extrapyramidal disorder
NMS	Heart rate abnormal
NMS	Heart rate increased
NMS	Hyperhidrosis
NMS	Hypertension
NMS	Hypotension
NMS	Labile blood pressure

NMS	Labile hypertension
NMS	Leukocytosis
NMS	Loss of consciousness
NMS	Muscle enzyme increased
NMS	Myoclonus
NMS	Myoglobin blood increased
NMS	Myoglobin blood present
NMS	Myoglobin urine present
NMS	Myoglobinaemia
NMS	Myoglobinuria
NMS	Parkinsonian crisis
NMS	Parkinsonian rest tremor
NMS	Parkinsonism
NMS	Parkinson's disease
NMS	Stupor
NMS	Tachycardia
NMS	Tremor
NMS	Unresponsive to stimuli
NMS	White blood cell count abnormal
NMS	White blood cell count increased

MTACCAT	AEDECOD
Tachycardia	Heart rate increased
Tachycardia	Sinus tachycardia
Tachycardia	Tachycardia
Tachycardia	Tachycardia paroxysmal

MWEICAT	AEDECOD
WEIGHT GAIN	Increased appetite
WEIGHT GAIN	Hyperphagia
WEIGHT GAIN	Obesity
WEIGHT GAIN	Overweight
WEIGHT GAIN	Abnormal weight gain
WEIGHT GAIN	Waist circumference increased
WEIGHT GAIN	Weight above normal
WEIGHT GAIN	Weight increased
WEIGHT GAIN	Waist circumference increased

MINJCAT	AEDECOD
INJECTION SITE	Injection site abscess
INJECTION SITE	Injection site abscess sterile
INJECTION SITE	Injection site anaesthesia
INJECTION SITE	Injection site atrophy
INJECTION SITE	Injection site bruising
INJECTION SITE	Injection site calcification
INJECTION SITE	Injection site cellulitis
INJECTION SITE	Injection site coldness
INJECTION SITE	Injection site cyst
INJECTION SITE	Injection site dermatitis
INJECTION SITE	Injection site desquamation
INJECTION SITE	Injection site discharge
INJECTION SITE	Injection site discolouration
INJECTION SITE	Injection site discomfort
INJECTION SITE	Injection site eczema
INJECTION SITE	Injection site erosion
INJECTION SITE	Injection site erythema
INJECTION SITE	Injection site extravasation
INJECTION SITE	Injection site fibrosis
INJECTION SITE	Injection site haematoma
INJECTION SITE	Injection site haemorrhage
INJECTION SITE	Injection site hypersensitivity
INJECTION SITE	Injection site hypertrophy
INJECTION SITE	Injection site induration
INJECTION SITE	Injection site infection
INJECTION SITE	Injection site inflammation

INJECTION SITE	Injection site injury
INJECTION SITE	Injection site irritation
INJECTION SITE	Injection site ischaemia
INJECTION SITE	Injection site lymphadenopathy
INJECTION SITE	Injection site mass
INJECTION SITE	Injection site movement impairment
INJECTION SITE	Injection site necrosis
INJECTION SITE	Injection site nerve damage
INJECTION SITE	Injection site nodule
INJECTION SITE	Injection site oedema
INJECTION SITE	Injection site pain
INJECTION SITE	Injection site pallor
INJECTION SITE	Injection site papule
INJECTION SITE	Injection site paraesthesia
INJECTION SITE	Injection site phlebitis
INJECTION SITE	Injection site photosensitivity reaction
INJECTION SITE	Injection site pruritus
INJECTION SITE	Injection site pustule
INJECTION SITE	Injection site rash
INJECTION SITE	Injection site reaction
INJECTION SITE	Injection site scab
INJECTION SITE	Injection site scar
INJECTION SITE	Injection site swelling
INJECTION SITE	Injection site thrombosis
INJECTION SITE	Injection site ulcer
INJECTION SITE	Injection site urticaria
INJECTION SITE	Injection site vesicles
INJECTION SITE	Injection site warmth
INJECTION SITE	Buttock pain
INJECTION SITE	Pain in extremity
INJECTION SITE	Puncture site pain
INJECTION SITE	Administration site pain
INJECTION SITE	Application site pain
INJECTION SITE	Injection site dryness
INJECTION SITE	Injection site dysaesthesia
INJECTION SITE	Injection site exfoliation
INJECTION SITE	Injection site granuloma
INJECTION SITE	Injection site hyperaesthesia
INJECTION SITE	Injection site laceration
INJECTION SITE	Injection site macule
INJECTION SITE	Injection site plaque
INJECTION SITE	Injection site streaking
INJECTION SITE	Injection site vasculitis

ATTACHMENT 2: Criteria of Markedly abnormal laboratory values

Laboratory Parameter[unit]	Markedly Abnormal Limits	
	Low	High
Albumin [g/L]	24	60
Alkaline phosphatase [U/L]	N/A	250
Alanine Aminotransaminase (SGPT) [U/L]	N/A	200
Aspartate Aminotransaminase (SGOT) [U/L]	N/A	250
Bicarbonate [mmol/L]	15.1	34.9
Blood urea nitrogen [mmol/L]	N/A	17.9
Calcium [mmol/L]	1.5	3
Chloride [mmol/L]	94	112
Cholesterol [mmol/L]	N/A	7.8
Creatinine [µmol/L]	N/A	265.2
Gamma glutamyl transferase [U/L]	N/A	300
Glucose [mmol/L]	2.2	16.7
HDL (mmol/L)	0.9	N/A
LDL (mmol/L)	2.3	4.1
Lactate Dehydrogenase[U/L]	N/A	500
Phosphate [mmol/L]	0.7	2.6
Potassium [mmol/L]	3.0	5.8
Sodium [mmol/L]	125	155
Bilirubin, total [µmol/L]	N/A	51.3
Protein, total (g/L)	50	N/A
Triglycerides [mmol/L]	N/A	5.7
Urate [µmol/L]	89.2	594.8
Hematocrit (fraction) -- female	0.28	0.5
-- male	0.24	0.55
Hemoglobin [g/L]	80	190
Neutrophils, Segmented [%]	30	90
Monocytes [%]	N/A	20
Eosinophils [%]	N/A	10
Basophils [%]	N/A	6
Lymphocytes [%]	10	60
Platelet count [$x10^9$/L]	100	600
Erythrocytes [$x10^{12}$/L] -- female	3.0	5.5
-- male	3.0	6.4
Leukocytes [$x10^9$/L]	2.5	15.0

Note: The same limits apply to both males and females unless gender is indicated; N/A = Not applicable.