# CLINICAL STUDY PROTOCOL

**Title: A Randomized, Double-Blind, Placebo-Controlled, Multi-Center Study to Assess the Antipsychotic Efficacy of ITI-007 in Patients With Schizophrenia.**

**Protocol Number:** **ITI-007-301**

**Protocol Version:** **1.1**

**Protocol Date:** **17 September 2014**

**Product:** **ITI-007**

**Sponsor**
Intra-Cellular Therapies, Inc. (ITI)
[redacted]

**Sponsor Study Contacts**
[redacted]

**Prepared by:**
[redacted]

# SIGNATURE PAGE

**Declaration of Sponsor**

**Protocol Title: A Randomized, Double-Blind, Placebo-Controlled, Multi-Center study to Assess the Antipsychotic Efficacy of ITI-007 in Patients with Schizophrenia**

This clinical study protocol was subjected to critical review. The information it contains is consistent with current knowledge of the risks and benefits of the investigational medicinal product (IMP), as well as with the moral, ethical and scientific principles governing clinical research as set out in the Declaration of Helsinki (Version 2008), and the guidelines on Good Clinical Practices (GCP) applicable to this clinical study.

**Sponsor Signatory**

# SIGNATURE PAGE

**Declaration of the Principal Investigator**

**Protocol Title: A Randomized, Double-Blind, Placebo-Controlled, Multi-Center study to Assess the Antipsychotic Efficacy of ITI-007 in Patients with Schizophrenia**

This clinical study protocol was subjected to critical review and has been released by the Sponsor. The information it contains is consistent with current risk and benefit evaluation of the investigational medicinal product (IMP), as well as with the moral, ethical and scientific principles governing clinical research as set out in the Declaration of Helsinki (Version 2008), and the guidelines on Good Clinical Practices (GCP) applicable to this clinical study.

**Principal Investigator**

| | |
|---|---|
| Signature | Date |
| Printed Name | Institution Name |

# TABLE OF CONTENTS

# PROTOCOL SYNOPSIS ITI-007-301 (PHASE III)

## TITLE: A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTI-CENTER STUDY TO ASSESS THE ANTIPSYCHOTIC EFFICACY OF ITI-007 IN PATIENTS WITH SCHIZOPHRENIA

### Phase

III

### Indication

Schizophrenia

### Objectives

#### *Primary Objective*

To determine whether ITI-007 administered to patients with acutely exacerbated schizophrenia demonstrates antipsychotic efficacy compared to placebo, as measured by change from baseline to Day 28 on the Positive and Negative Syndrome Scale (PANSS) total score.

#### *Secondary Objectives*

To determine whether, compared to placebo, ITI-007 administered to patients with acutely exacerbated schizophrenia:

- demonstrates enhanced social function as measured by change from baseline on a PANSS-derived Social Factor;
- demonstrates enhanced social function as measured by change from baseline on the Personal and Social Performance Scale (PSP);
- demonstrates efficacy as assessed by change from baseline on any of the individual PANSS subscales: Positive, Negative, and General Psychopathology;
- demonstrates efficacy as assessed by change from baseline on the negative symptoms subscale of the PANSS in a subgroup of patients with prominent negative symptoms at baseline (a score of 4 or higher on at least 3 negative

subscale symptom items) and as assessed by change from baseline on a PANSS-derived Negative Symptom Factor in a subgroup of patients with prominent negative symptoms at baseline (a score of 4 or higher on at least 3 Negative Symptom Factor items);

- demonstrates an improvement in symptoms of depression as measured by the Calgary Depression Scale for Schizophrenia (CDSS) in a subgroup of patients with co-morbid depressive symptoms at baseline (CDSS total score > 6);
- demonstrates an improvement in symptoms of psychosis as measured by the total PANSS and PANSS subscales in a subgroup of patients with co-morbid depressive symptoms at baseline (CDSS total score > 6);
- demonstrates therapeutic efficacy as assessed by change from baseline on the Clinical Global Impression Scale-Severity (CGI-S).

### *Safety Objectives*

To determine whether ITI-007 is safe and well tolerated.

### *Exploratory Objectives*

The exploratory objectives of the study include, but are not limited to, the evaluation of the association between response to treatment and any genetic trait or change from baseline in protein biomarkers, such as p11.

## Study Design

The study will be conducted as a randomized, double-blind, parallel-group, placebo-controlled, multi-center study in patients diagnosed with schizophrenia having an acute exacerbation of psychosis.

Approximately 440 subjects will be randomized and receive one of three study treatments in a ratio of 1:1:1; ITI-007 40 mg QD, ITI-007 60 mg QD, placebo QD, for 28 days. Based on its mechanism of action with mesolimbic/mesocortical selectivity and data from animal models, it was hypothesized that ITI-007 would demonstrate antipsychotic efficacy at doses

with striatal $D_2$ receptor occupancy of 50% and lower, in contrast to most antipsychotic drugs that require >60% striatal $D_2$ receptor occupancy. The 60 mg ITI-007 dose is projected to have about 50% striatal $D_2$ receptor occupancy and demonstrated statistically significant improvement in psychosis as measured by the PANSS total score after 28 days of treatment in a Phase 2 clinical trial (ITI-007-005). A higher dose, 120 mg ITI-007 that was projected to have approximately 70% striatal D2 receptor occupancy did not separate from placebo in that Phase 2 trial. Therefore, the 60 mg dose and a lower dose of 40 mg ITI-007 will be evaluated in order to explore the dose-response relationship of ITI-007. A dose of 40 mg ITI-007 demonstrated peak striatal $D_2$ receptor occupancy of approximately 40% and an average striatal D2 receptor occupancy of 29% in healthy volunteers (ITI-007-003).

Study participation will last approximately 7 weeks and include an inpatient *Screening Period* (2 to 7 days prior to Day 1), an inpatient *Study Treatment Period* with 4 weeks of daily administration of study treatment, up to a 5-day inpatient *Stabilization Period* during which subjects will be stabilized on standard antipsychotic medication, and an outpatient *Final Safety Follow-up* approximately 2 weeks after the end of the Treatment Period. Subjects with symptoms that are considered by the site investigator to be appropriately controlled may receive a day-pass to leave the clinic during the inpatient portion of the study in order to attend to personal business accompanied and supervised by clinic staff. Such day-passes must not occur on scheduled Study Visit days and no more than two day passes should be allowed per subject.

Subjects will be evaluated at baseline and periodically thereafter using the PANSS, CGI-S, PSP, CDSS, Simpson and Angus Scale (SAS), Barnes Akathisia Rating Scale (BARS), Abnormal Involuntary Movement Scale (AIMS), Columbia Suicide Severity Rating Scale (C-SSRS), 12-lead electrocardiogram (ECG), 3-positional vital signs, clinical laboratory values, body weight/waist circumference, and reported/observed adverse events. In addition, blood samples will be collected for exploratory biomarker analysis.

### Population

Subjects will be male or female, 18-60 years of age, with a current clinical diagnosis of schizophrenia according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5); confirmed by the Structured Clinical Interview for DSM Disorders-Clinical Trial Version (SCID-CT, modified for use in this study), who are experiencing an acute exacerbation of psychosis. Acute exacerbation is defined as a total score on the Brief

Psychiatric Rating Scale (BPRS; using an item range of 1-7) of at least a 40 with a score of 4 or higher on at least two of the positive symptom items of suspiciousness, conceptual disorganization, hallucinatory behavior, unusual thought content; onset of the current acute episode must have occurred within 4 weeks of screening. In addition, patients will have a score of at least 4 on the CGI-S. Appropriate severity of illness will be confirmed by a PANSS score of 70 or higher and a CGI-S score of 4 or higher at baseline as assessed by a centralized rater blinded to study visit.

To ensure the inclusion of appropriate subjects, the information collected during the Screening Period that pertains to diagnosis of schizophrenia, evaluation of the level of disease severity and exacerbation, and the distinction between symptoms of exacerbated schizophrenia and co-morbid conditions will be recorded (e.g., using audio and/or video recording device). The information and recorded interviews will be submitted for review by an experienced independent psychiatrist (MD) or clinical psychologist (PhD) to confirm that the subject is appropriate for inclusion in the study.

### Sample size

Approximately 440 subjects will be randomized in the study. See Statistical Procedures below for sample size calculation.

### Treatments

ITI-007 and placebo will be administered as visually-matched Size 0 capsules. A dose of 60 mg ITI-007 will be administered as a single 60 mg capsule and one placebo capsule. A dose of 40 mg ITI-007 will be administered as two 20 mg capsules. Patients randomized to placebo will receive two placebo capsules per dose.

Study treatments will be administered in the morning during breakfast. Study treatments will be administered once daily for 28 days.

### Pharmacokinetic data

Blood samples will be collected pre-dose and 3-4 h post-dose on Study Days 1, 8, 28, again at 6-7 h post-dose on Day 28, and once on Day 33 for determination of ITI-007 and metabolite concentrations. Samples will be collected from all subjects in order to maintain the blind. All samples will be analyzed from ITI-007 recipients; selected samples from

placebo recipients will be analyzed for presence/absence of active drug to ensure appropriate randomization. Descriptive statistics for trough and peak ITI-007/metabolite levels will be calculated.

### Pharmacodynamic data

The PANSS will be determined by remote central independent raters by video conference. The remote central raters will be blinded to study treatment and to study visit. CGI-S will be determined by a remote central independent rater and a separate CGI-S will be determined by a qualified site-based rater; all other assessments will be site-based assessments by qualified individuals.

Change from baseline scores will be calculated for:

- Total PANSS to assess psychosis, as well as
  - PANSS Social Factor
  - PANSS Positive subscale
  - PANSS Negative subscale
  - PANSS General Psychopathology subscale
- PSP to assess social function
- CGI-S to assess global disease severity
- CDSS to assess symptoms of depression

### Safety data

Tolerability will be assessed by reported and observed adverse events. Frequency and severity of adverse events will be summarized.

Routine safety assessments will be performed throughout the study, with a final assessment performed when a subject completes the study on Day 42 or if a subject discontinues. Safety assessments will include: modified physical examinations including height, weight, waist circumference, 12-lead electrocardiograms, vital signs (3-positional blood pressure and pulse rate, respiration rate, and oral body temperature), clinical laboratory tests (hematology, serum chemistry, and urinalysis), and pregnancy test (where applicable). Motor tolerability and

safety will be assessed by the SAS, BARS and AIMS and suicidality will be evaluated by the C-SSRS.

### Exploratory Data

For the evaluation of the association between response to treatment and any genetic trait or change from baseline in protein biomarkers, blood samples will be collected. On Day 1 (pre-dose) and Day 28 blood samples will be collected to measure levels of protein biomarkers, including but not limited to the protein p11. An additional blood sample will be collected once during the study from those subjects who provide separate, optional consent to have their samples analyzed for genetic mutations and single-nucleotide polymorphisms.

### Statistical procedures

A formal and detailed statistical data analysis plan will be prepared before clinical conduct begins. Briefly, descriptive statistics for safety, pharmacodynamic and pharmacokinetic variables will be summarized according to treatment group. Shift tables for safety parameters will be prepared for appropriate variables. Change from baseline and effect sizes will be calculated for pharmacodynamic measures. Inferential statistical comparisons of ITI-007 treatment groups to placebo will be made for efficacy endpoints. The primary endpoint will be the change from baseline to Day 28 on the total PANSS score for 60 mg ITI-007 compared to placebo. Only if 60 mg ITI-007 significantly separates from placebo will 40 mg ITI-007 be compared to placebo.

This step-wise approach will preserve the pre-specified level of significance for the primary efficacy analysis.

Approximately 440 subjects will be randomized in a 1:1:1 ratio to one of the three treatment arms, which will provide approximately 396 evaluable subjects, assuming a 10% early discontinuation rate before a post-dose assessment is made. With a sample size of approximately 132 in each of the treatment arms: ITI-007 60 mg QD, ITI-007 40 mg QD, or placebo, the study is designed to have 90% power to demonstrate an effect size of 0.4 corresponding, for example, to a 6 point difference from placebo, with a common standard deviation of 15, in change from baseline of total PANSS score between a treatment arm and placebo, two-sided significance level of 0.05. The Mixed-effect Model for Repeated

Measures (MMRM) will be used for the primary efficacy analysis on the intent-to-treat (ITT) population.

### Study duration

The duration of participation for each individual subject in this study will be approximately 7 weeks including the up-to-7-day Screening Period, 28-day Inpatient Study Treatment Period, up to a 5-day inpatient Stabilization Period, and the Final Safety Follow-up approximately 2 weeks after the completion of the Treatment Period.

The duration of the clinical conduct portion of the trial is estimated to be 12 months. Approximately 12 sites are planned to randomize 440 patients over a 10 month period. An additional two months allow for the screening period, treatment period and follow up period to complete clinical conduct. The actual schedule will vary depending on clinic and subject availability.

# LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

| | |
|---|---|
| IC200056 | Active moiety of the drug product, ITI-007 |
| ITI-007 | Tosylate salt of the active drug substance IC200056 |
| 5-HT2A | Serotonin 2A receptor subtype |
| AE | Adverse event |
| AIMS | Abnormal Involuntary Movement Scale |
| ALT | Alanine transaminase (synonymous with serum glutamic-pyruvic transaminase – SGPT) |
| AST | Aspartate transaminase (synonymous with serum glutamic-oxaloacetic transaminase SGOT) |
| BARS | Barnes Akathisia Rating Scale |
| BPRS | Brief Psychiatric Rating Scale |
| BMI | Body mass index |
| BUN | Blood urea nitrogen |
| CDSS | Calgary Depression Scale for Schizophrenia |
| CGI-S | Clinical Global Impression-Severity |
| CNS | Central nervous system |
| CPK | Creatine phosphokinase |
| CRA | Clinical Research Associate |
| CRF | Case report form |
| CRO | Contract research organization |
| C-SSRS | Columbia Suicide Severity Rating Scale |

| | |
|---|---|
| CYP | Cytochrome P450 isozyme |
| D1 | Dopamine 1 receptor subtype |
| D2 | Dopamine 2 receptor subtype |
| DPPM | Dopamine receptor phosphoprotein modulator |
| DSM-5 | Diagnostic and Statistical Manual of Mental Disorders, Version 5 |
| ECG | Electrocardiograms |
| GCP | Good Clinical Practices |
| GGT | Gamma-glutamyl transferase |
| H | Hour |
| HepB | Hepatitis B surface antigen |
| HepC | Hepatitis C antibody |
| HIV | Human immunodeficiency virus |
| HR | Heart rate |
| ICH | International Conference on Harmonization |
| IEC | Independent Ethics Committee |
| IRB | Institutional Review Board |
| Kg | Kilogram |
| LC/MS/MS | Liquid chromatographic tandem mass spectroscopy |
| LDH | Lactate dehydrogenase |
| LLQ | Lower limit of quantification |
| MCH | Mean corpuscular hemoglobin |

| | |
|---|---|
| MCHC | Mean corpuscular hemoglobin concentration |
| MCV | Mean corpuscular volume |
| MedDRA | Medical Dictionary for Regulatory Activities |
| Mg | Milligram |
| Min | minute(s) |
| ML | Milliliter |
| MMRM | Mixed-effects Model for Repeated Measures |
| MRI | Magnetic Resonance Imaging |
| PANSS | Positive And Negative Syndrome Scale |
| PET | Positron emission tomography |
| PK | Pharmacokinetic |
| PR | PR interval of ECG |
| PSP | Personal and Social Performance Scale |
| QRS | QRS interval of ECG |
| QT | QT interval of ECG |
| QTc | QT interval corrected for heart rate of ECG |
| QTcB | QT interval corrected for heart rate of ECG using Bazett's formula |
| QTcF | QT interval corrected for heart rate of ECG using Fridericia's formula |
| RBC | Red blood cell |
| ROI | Region(s) of interest |
| SAS | Simpson Angus Scale |

| | |
|---|---|
| SAE | Serious adverse event |
| SAP | Statistical analysis plan |
| SCID-CT | Structured Clinical Interview for DSM Disorders - Clinical Trials Version |
| SGOT | Serum glutamic-oxaloacetic transaminase (synonymous with AST) |
| SGPT | Serum glutamic-pyruvic transaminase (synonymous with ALT) |
| SOP | Standard operating procedures |
| SRTM | Simplified reference tissue modeling |
| TEAE | Treatment-emergent adverse event |
| WBC | White blood cell |

# 1.0 INTRODUCTION

Modulation of serotonergic, dopaminergic, and glutamatergic neurotransmission in the central nervous system has proven to be an effective means of treating a variety of abnormal behaviors in humans. Among the many behaviors affected by altered neurotransmission in these systems are a variety of psychiatric conditions, including schizophrenia.

Schizophrenia is a chronic psychiatric disorder with the prevalence of about 1-2% in general population. Schizophrenia disorder is characterized by positive symptoms such as hallucination, delusion, though disorder and negative symptoms i.e. social withdrawal, lack of motivation. Schizophrenia is one of the most serious mental illnesses that has significant impact to patients' families as well as social and economic implications**.**

The blockade of dopamine receptors is thought principally to mediate the ability of antipsychotic drugs to control psychosis, hallucinations and delusions in patients with schizophrenia. However, excessive blockade of dopamine receptors can lead to tardive dyskinesia, parkinsonism and other and extrapyramidal side-effects (EPS), neuroleptic malignant syndromes and hyperprolactinemia. Most antipsychotics block between 60 to 75% of striatal dopamine receptors in brain to achieve antipsychotic efficacy. Unfortunately, many of the above-mentioned side-effects emerge when slightly more of the dopamine receptors are occupied (> 80%). A delicate balance between too little and too much dopamine blockade must be achieved for safe antipsychotic drug therapy.

The possibility that antagonism of 5-$HT_{2A}$ receptors may act synergistically with antagonism of dopamine receptors to control psychotic behavior emerged with the introduction of clozapine. The uniqueness of atypical antipsychotic drugs like clozapine appears to be driven by high affinity interactions of 5-$HT_{2A}$ receptors that add to or counteract some of the influences of strong $D_2$ receptor blockade, improving efficacy while lowering the propensity to induce EPS and hyperprolactinemia. At therapeutic doses in patients with schizophrenia, clozapine exhibits mean striatal $D_2$-receptor occupancies of approximately 45% whereas the 5-$HT_2$-receptor occupancy is approximately 90%. Other atypical antipsychotic drugs, like risperidone, have less selectivity for 5-$HT_{2A}$ receptors relative to dopamine receptors and do

not fully occupy cortical 5-$HT_{2A}$ receptors, when occupancy of $D_2$ receptors is optimal and between 60 to 80%. Atypical antipsychotics other than clozapine and quetiapine do elevate prolactin levels and can induce EPS at therapeutic doses.

Thus, based on positron emission tomography (PET) receptor occupancy studies, human striatal $D_2$ receptor occupancy in the basal ganglia regions should not exceed 60% and should be targeted to around 40% for maximal therapeutic efficacy and to avoid EPS liability in patients with schizophrenia. However, higher levels of $D_2$ occupancy in the neocortex may be advantageous. In contrast, 5-$HT_{2A}$ receptors (which have been extensively studied in clinical trials with specific 5-$HT_{2A}$ receptor antagonists where full saturation was shown to be safe) can be and probably should be saturated at therapeutically useful doses in these patients in order to achieve maximum therapeutic benefit and/or reduce neuromotor side effect burden.

ITI-007 represents a novel small molecule therapeutic agent under development for the treatment of schizophrenia, sleep disorders, mood disorders and other neuropsychiatric and neurological disorders. ITI-007 is a potent serotonin 5-$HT_{2A}$ receptor antagonist, a dopamine receptor phosphoprotein modulator (DPPM) and serotonin transporter (SERT) inhibitor. As a mesolimbic/mesocortical selective DPPM, ITI-007 exhibits pre-synaptic partial agonism and post-synaptic antagonism at dopamine $D_2$ receptors *in vivo*. ITI-007 also increases phosphorylation of glutamatergic N-methyl-D-aspartate (NMDA) NR2B or GluN2B receptors downstream of dopamine $D_1$ receptor activation in mesolimbic brain regions. Together, this unique pharmacological profile predicts low dose enhancement of sleep, antipsychotic and antidepressant efficacy at higher doses, and a favorable side effect profile. Consistent with its pharmacological profile, preclinical models demonstrate that ITI-007 increases sleep at low doses and exhibits antipsychotic-like efficacy and antidepressant-like efficacy at higher doses.

Clinically, ITI-007 is safe and well-tolerated across a broad range of doses (1 to 140 mg) evaluated in healthy volunteer and patient populations. A positron emission tomography (PET) study (ITI-007-003) was conducted as a single center, open-label study in sequential groups of healthy volunteers. PET was used to determine dopamine $D_2$ receptor, serotonin transporter, and/or serotonin 5-$HT_{2A}$ receptor occupancy in the brain at various time points after ITI-007 administration. Single oral doses ranging from 10 mg to 40 mg were evaluated. This study demonstrated that striatal $D_2$ receptor occupancy is about 12% after 10 mg ITI-007, about 19%

after 20 mg ITI-007, about 27% after 30 mg ITI-007, and about 29% after 40 mg ITI-007 (with a peak occupancy of about 39% after a 40 mg dose of ITI-007 in one of two subjects tested). ITI-007 rapidly penetrates brain as demonstrated by peak striatal $D_2$ receptor occupancy achieved within 30 min after oral administration. ITI-007 shows long-lasting brain residency time as determined by multiple PET scans after a single dose administration of ITI-007. Consistent with a centrally-acting drug, brain residency time for ITI-007 outlasts its plasma half-life. Cortical 5-$HT_{2A}$ receptors are essentially fully occupied at 10 mg ITI-007 (88% occupancy). ITI-007 occupies striatal serotonin transporters (SERT) to a similar extent as it occupies striatal $D_2$ receptors; a dose of 40 mg ITI-007 occupies about 22% of striatal SERTs (with a peak of about 31% after 40 mg ITI-007 in one of two subjects tested). Lower occupancy levels at a given dose tended to correspond to lower ITI-007 plasma levels. There appears to be a good relationship between plasma levels of ITI-007 and brain occupancy.

ITI-007 has been demonstrated to be safe and well tolerated in healthy volunteers at single oral doses up to and including 40 mg (highest single dose tested in healthy volunteers; ITI-007-001 and ITI-007-003) and at multiple oral doses up to and including 20 mg (highest multiple dose tested in healthy volunteers; ITI-007-002). Multiple oral doses up to and including 140 mg ITI-007 were generally well tolerated in patients with stable schizophrenia (ITI-007-002). ITI-007 has been safely administered for 28 days at doses of 60 mg and 120 mg in patients with acute schizophrenia (ITI-007-005). Moreover, a 60 mg dose of ITI-007 demonstrated statistically significant improvement in psychosis as demonstrated by a reduction in Positive and Negative Syndrome Scale (PANSS) total score compared to placebo after 28 days of oral administration in patients with acute schizophrenia (ITI-007-005). Based on the PET data collected in healthy volunteers at lower doses (10 – 40 mg; ITI-007-003) and the plasma levels achieved across the range of doses administered to patients with schizophrenia (up to and including 140 mg; ITI-007-002), we projected approximately 50% striatal $D_2$ receptor occupancy at the 60 mg ITI-007 efficacious dose for treating symptoms of schizophrenia. Given the mesolimbic/mesocortical selectivity of ITI-007 measured in preclinical models, we hypothesize higher cortical (extrastriatal) $D_2$ receptor occupancy than striatal $D_2$ receptor occupancy of any given dose of ITI-007. Such a neurobiological profile, with relatively higher extrastriatal $D_2$ receptor occupancy than striatal $D_2$ receptor occupancy, predicts mesolimbic/mesocortically mediated antipsychotic efficacy with a relatively lower risk of striatally mediated motor side effects. This hypothesis is consistent with the clinical profile observed with 60 mg ITI-007 to date. Given similar *in vitro* affinities of ITI-007 for $D_2$

receptors, $D_1$ receptors and serotonin transporters, we hypothesize similar occupancy of these brain targets by ITI-007 in patients with schizophrenia. Activity through these additional targets, such as inhibition of serotonin reuptake and partial agonism of $D_1$ receptors (with subsequent downstream phosphorylation of NR2B or GluN2B receptor subunits), predict antidepressant efficacy and cognitive improvement, respectively. Such activity would be of clinical benefit to patients suffering from schizophrenia and residual symptoms of depression and cognitive impairment.

The higher dose of 120 mg ITI-007 evaluated in the ITI-007-005 Phase 2 trial was projected to occupy approximately 70% striatal $D_2$ receptors, but did not separate from placebo. It has been speculated that the serotonergic/dopaminergic/glutamatergic balance is optimal at 60 mg ITI-007 and becomes suboptimal at the higher dose. It could also be that the frequent mild to moderate daytime sedation/somnolence occurring in about a third of the patients at the 120 mg ITI-007 dose in the ITI-007-005 trial, interfered with the ability to measure antipsychotic efficacy using remote central raters who conducted the PANSS interviews over video conference. Likely it was a combination of factors, complex pharmacology combined with study specific design features. However, 60 mg ITI -007 was clearly effective in reducing psychosis in patients with schizophrenia given once daily in the morning and safe in the ITI-007-005 trial. In order to more fully explore the dose response for antipsychotic efficacy in the Phase 3 program, a dose lower than 60 mg ITI-007 will be evaluated. A dose of 40 mg ITI-007 was demonstrated to occupy an average of 29% striatal $D_2$ receptors with 39% peak occupancy in healthy volunteers (ITI-007-003) and remains in the target range of dopamine receptor occupancy for antipsychotic efficacy. Lower doses, though having measureable levels of striatal $D_2$ receptor occupancy, might not have sufficient $D_2$ receptor occupancy for antipsychotic efficacy. Selective 5-$HT_{2A}$ receptor antagonists have not shown antipsychotic efficacy in schizophrenia without separate additional $D_2$ receptor occupancy. If 40 mg ITI-007 is effective in treating schizophrenia, the exploration of lower doses may be warranted in future studies.

The purpose of the present study is to determine the efficacy of ITI-007 for the treatment of schizophrenia. Acutely exacerbated patients with schizophrenia will be randomized to receive one of two doses of ITI-007 (40 mg or 60 mg) or placebo. Because 60 mg ITI-007 has demonstrated statistically significant antipsychotic efficacy in a well-controlled randomized

trial (ITI-007-005), a separate positive control will not be included in this study. Rather, the 60 mg ITI-007 dose will serve to confirm assay sensitivity if the previous results are replicated. The primary objective will be to assess the effects of ITI-007 on total psychopathology related to schizophrenia. Secondary objectives will assess the effects of ITI-007 on separate symptom domains related to schizophrenia and explore potential differentiating features of ITI-007 to reduce symptoms of depression and improve negative symptoms. The safety and tolerability in the target patient population will also be assessed.

Details on the profile of ITI-007 can be found in the ITI-007 Investigator's Drug Brochure issued September 2014.

# 2.0 STUDY OBJECTIVES

### *Primary Objective*

To determine whether ITI-007 administered to patients with acutely exacerbated schizophrenia demonstrates antipsychotic efficacy compared to placebo, as measured by change from baseline to Day 28 on the Positive and Negative Syndrome Scale (PANSS) total score.

### *Secondary Objectives*

To determine whether, compared to placebo, ITI-007 administered to patients with acutely exacerbated schizophrenia:

- demonstrates enhanced social function as measured by change from baseline on a PANSS-derived Social Factor;
- demonstrates enhanced social function as demonstrated by change from baseline on the Personal and Social Performance Scale (PSP);
- demonstrates efficacy as assessed by change from baseline on any of the individual PANSS subscales: Positive, Negative, and General Psychopathology;
- demonstrates efficacy as assessed by change from baseline on the negative symptoms subscale of the PANSS in a subgroup of patients with prominent negative symptoms at baseline (a score of 4 or higher on at least 3 negative subscale symptom items) & as assessed by change from baseline on a PANSS-derived Negative Symptom Factor in a subgroup of patients with prominent negative symptoms at baseline (a score of 4 or higher on at least 3 Negative Symptom Factor items);
- demonstrates an improvement in symptoms of depression as measured by the Calgary Depression Scale for Schizophrenia (CDSS) in a subgroup of patients with co-morbid depressive symptoms at baseline (CDSS total score > 6);
- demonstrates an improvement in symptoms of psychosis as measured by the total PANSS and PANSS subscales in a subgroup of patients with co-morbid depressive symptoms at baseline (CDSS total score > 6);

- demonstrates therapeutic efficacy as assessed by change from baseline on the Clinical Global Impression Scale-Severity (CGI-S).

### *Safety Objectives*

To determine whether ITI-007 is safe and well tolerated.

### *Exploratory Objectives*

The exploratory objectives of the study include, but are not limited to, the evaluation of the association between response to treatment and any genetic trait or change from baseline in protein biomarkers, such as p11.

# 3.0 INVESTIGATIONAL PLAN

## 3.1 Study Design and Rationale for Study Design

### 3.1.1 Study Design

The study will be conducted as a randomized, double-blind, placebo-controlled, multi-center study in patients diagnosed with schizophrenia having an acute exacerbation of psychosis.

Approximately 440 subjects will be randomized and receive one of three study treatments in a ratio of 1:1:1; ITI-007 40 mg QD, ITI-007 60 mg QD, placebo QD. Based on its mechanism of action with mesolimbic/mesocortical selectivity and data from animal models, it was hypothesized that ITI-007 would demonstrate antipsychotic efficacy at doses with striatal $D_2$ receptor occupancy of 50% and lower, in contrast to most antipsychotic drugs that require >60% striatal $D_2$ receptor occupancy. The 60 mg ITI-007 dose is projected to have about 50% striatal $D_2$ receptor occupancy and demonstrated statistically significant improvement in psychosis as measured by the PANSS total score after 28 days of treatment in a Phase 2 clinical trial (ITI-007-005). A higher dose, 120 mg ITI-007 that was projected to have approximately 70% striatal D2 receptor occupancy did not separate from placebo in that Phase 2 trial. Therefore, the 60 mg dose and a lower dose of 40 mg ITI-007 will be evaluated in order to explore the dose-response relationship of ITI-007. A dose of 40 mg ITI-007 demonstrated peak striatal $D_2$ receptor occupancy of approximately 40% and an average striatal D2 receptor occupancy of 29% in healthy volunteers (ITI-007-003).

Study participation will last approximately 7 weeks and include an inpatient *Screening Period* (2 to 7 days prior to Day 1), an inpatient *Study Treatment Period* with 4 weeks of daily administration of study treatment, up to a 5-day inpatient *Stabilization Period* during which patients will be stabilized on standard antipsychotic medication, and an outpatient *Final Safety Follow-up* approximately 2 weeks after the end of the Treatment Period. Subjects with symptoms that are considered by the site investigator to be appropriately controlled may receive a day-pass to leave the clinic during the inpatient portion of the study in order to attend to personal business accompanied and supervised by clinic staff. Such day-passes must not occur on scheduled Study Visit days and no more than two day passes should be allowed per subject.

Subjects will be evaluated at baseline and periodically thereafter using the PANSS, CGI-S, PSP, CDSS, Simpson and Angus Scale (SAS), Barnes Akathisia Rating Scale (BARS),

Abnormal Involuntary Movement Scale (AIMS), Columbia Suicide Severity Rating Scale (C-SSRS), 12-lead electrocardiogram (ECG), 3-positional vital signs, clinical laboratory values, body weight/waist circumference, and reported/observed adverse events. In addition, blood samples will be collected for exploratory biomarker analysis.

The study design is shown schematically in **Figure 1** and the Schedule of Assessments for each scheduled visit is presented in *Section 6.1*.

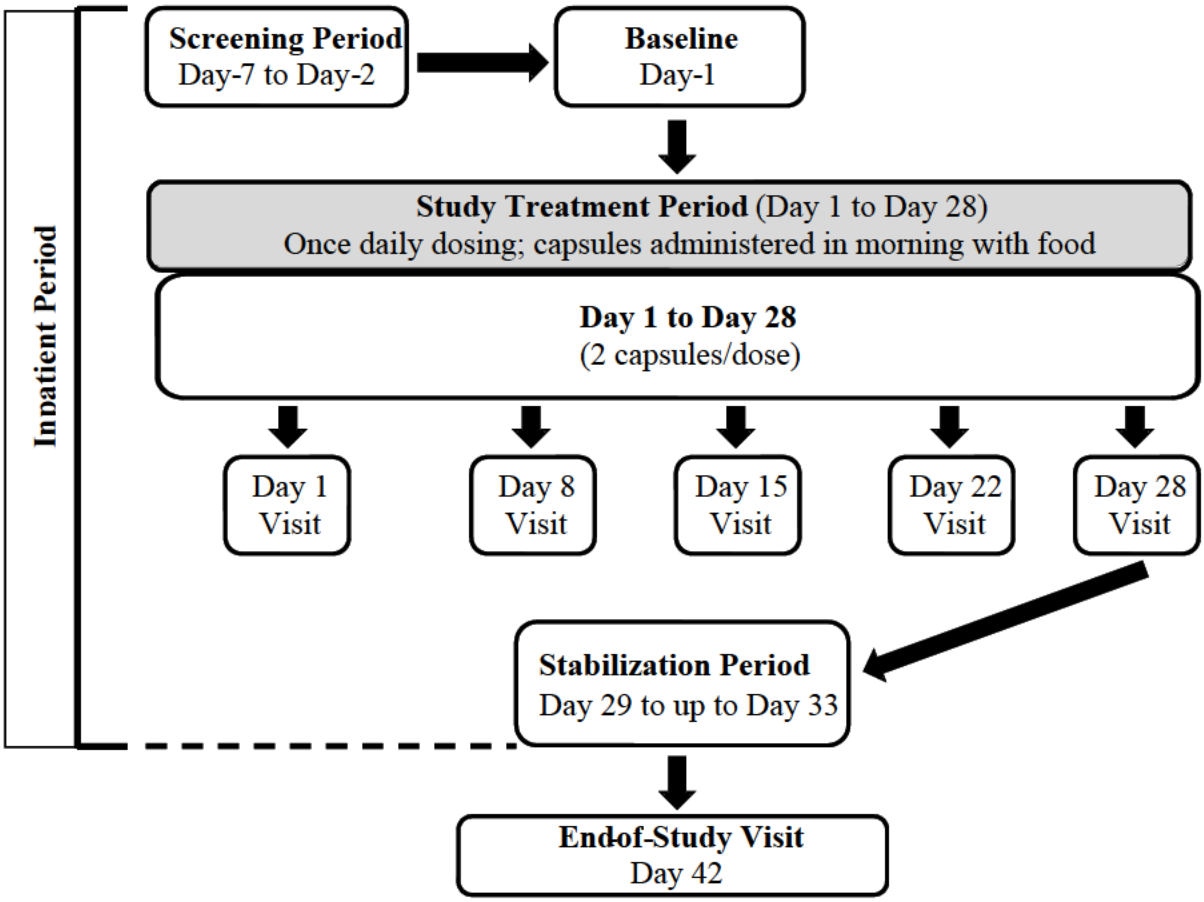



**Figure 1: Schematic of Study Design**

### 3.1.2 Rationale

The blockade of dopamine receptors is thought principally to mediate the ability of antipsychotic drugs to control psychosis, hallucinations, and delusions in patients with schizophrenia (Snyder 1976). However, excessive blockade of dopamine receptors can lead to tardive dyskinesia, parkinsonism and other extrapyramidal symptoms (EPS), neuroleptic

malignant syndromes and hyperprolactinemia. Most antipsychotics occupy between 60% and 75% of striatal dopamine receptors at doses required to achieve antipsychotic efficacy. Unfortunately, many of the above-mentioned side effects emerge when slightly more of the dopamine receptors are occupied (>80%). A delicate balance between too little and too much dopamine blockade must be achieved for safe and effective antipsychotic drug therapy (Farde et al 1995; Bushe et al 2008; Rummel-Kluge et al 2012).

The possibility that antagonism of 5-$HT_{2A}$ receptors may act synergistically with antagonism of dopamine receptors to control psychotic behavior emerged with the introduction of clozapine. The uniqueness of the newer atypical antipsychotic drugs like clozapine appears to be driven by high affinity interactions of 5-$HT_{2A}$ receptors that add to or counteract some of the influences of strong $D_2$ receptor blockade, improving efficacy while lowering the propensity to induce EPS and hyperprolactinemia (Meltzer 1989). At therapeutic doses in patients with schizophrenia, clozapine exhibits mean striatal $D_2$ receptor occupancies of approximately 45%, whereas the 5-$HT_2$ receptor occupancy is approximately 90% (Nordstom et al 1995). Other atypical antipsychotic drugs, like risperidone, have less selectivity for 5-$HT_{2A}$ receptors relative to dopamine receptors and do not fully occupy cortical 5-$HT_{2A}$ receptors, when occupancy of $D_2$ receptors is between 60% and 80% (required occupancy for the efficacy of most atypical antipsychotic drugs (Farde et al 1995). These latter atypical antipsychotic drugs still elevate prolactin levels and induce EPS at therapeutic doses.

ITI-007 represents an approach more similar to that of clozapine with full saturation of cortical 5-$HT_{2A}$ receptors when striatal $D_2$ receptor occupancy is relatively low. Moreover, the preclinical in vivo findings that ITI-007 exhibits mesolimbic/mesocortical selective pre-synaptic partial agonism and post-synaptic antagonism of dopamine receptors may help to achieve antipsychotic efficacy with limited motor side effects. Uniquely, at doses that modulate dopamine receptors, ITI-007 also increases phosphorylation of GluN2B receptor subunits of NMDA receptors preclinically in vivo, thought to indirectly enhance glutamatergic signaling by increasing cell surface expression of NMDA receptors, as well as inhibiting serotonin transporters. Modulation of serotonergic, dopaminergic, and glutamatergic neurotransmission in the central nervous system has proven to be an effective means of treating a variety of abnormal behaviors in humans, but driving the doses too high might upset the balance among these ubiquitous neurotransmitter systems.

In an animal model predictive of antipsychotic efficacy, ITI-007 significantly inhibited conditioned avoidance response (CAR) in rats. The $ED_{50}$ for inhibition of CAR was 1.5 mg/kg, by mouth (PO), at the time of peak effect (2 hours after oral administration). Based on scaling by BSA, this translates to a projected human dose of about 17 mg.

Importantly, no haloperidol-like frank catalepsy is observed with ITI-007 even at doses up to and including 30 mg/kg, PO (projected 340 mg human dose scaled by BSA) with the maximal increase in step-down latency with ITI-007 occurring at a level that was half the maximal haloperidol response. Mesocortical selectivity of ITI-007 was demonstrated using in vivo microdialysis. A dose of 3 mg/kg ITI-007 PO significantly increased extracellular concentrations of dopamine in the prefrontal cortex with lesser effects in the striatum. Interestingly, increasing the dose to 10 mg/kg ITI-007 PO resulted in a less robust effect on dopamine release than 3 mg/kg, suggesting that increasing the dose does not provide an increased benefit on modulating dopamine. Based on scaling by BSA, these data indicate an efficacious human dose around 30 – 40 mg ITI-007 and that increasing to a human dose over 100 mg would not provide additional modulation of dopamine in the cortex.

In the ITI-007-003 PET study in healthy volunteers, a dose of 20 mg ITI-007 (the lowest predicted antipsychotic dose based on CAR) demonstrated a peak of about 20% striatal $D_2$-receptor occupancy, with essentially full (~88%) cortical 5-$HT_{2A}$ receptor occupancy already achieved at a dose of 10 mg ITI-007. This would put the striatal $D_2$ receptor occupancy for ITI-007 lower than that of the vast majority of approved antipsychotic drugs. Rather than test this hypothesis in Phase 2, doses of 60 mg ITI-007 (projected to have ~ 50% striatal $D_2$ receptor occupancy) and 120 mg ITI-007 (projected to have ~70% striatal $D_2$ receptor occupancy) were included. These dopamine receptor occupancies would be in the range of other antipsychotic drugs. Indeed, 60 mg ITI-007, but not 120 mg ITI-007, demonstrated antipsychotic efficacy in the ITI-007-005 Phase 2 clinical trial, consistent with the initial hypothesis that ITI-007 should show antipsychotic efficacy at relatively low levels of striatal $D_2$ receptor occupancy and at doses projected to be mesocortical selective.

***Dose Selection for the Present Trial***

A dose of 60 mg ITI-007 exhibited antipsychotic efficacy in a randomized, double-blind, placebo- and active-controlled clinical trial in patients with acute schizophrenia (ITI-007-005). This dose was also found to be safe and well tolerated in patients with schizophrenia after 28 days of once daily oral administration in the morning. Based on earlier PET data using lower doses of ITI-007 in healthy volunteers, the dose of 60 mg ITI-007 is projected to occupy approximately 50% of the striatal $D_2$ receptors. The 60 mg ITI-007 dose and a lower dose of 40 mg ITI-007 will be evaluated in the present trial.

## 3.2 Selection of Study Population

The subjects will be healthy male and female subjects between 18-60 years (inclusive).

### 3.2.1 Inclusion Criteria

Subjects meeting all of the following criteria will be considered for admission to the study

1. male or female subjects of any race, ages 18-60 inclusive, with a clinical diagnosis of schizophrenia according to DSM-5 criteria as confirmed by the modified Structured Clinical Interview for DSM Disorders – Clinical Trials Version (SCID-CT);
2. experiencing an acute exacerbation of psychosis where the following criteria are met:
   a. Screening score on the BPRS is 40 or greater;
   b. Screening score on the BPRS includes a minimum score of 4 or higher on at least two of the following four positive symptom items: suspiciousness, conceptual disorganization, hallucinatory behavior, unusual thought content;
   c. Screening score on the CGI-S of at least 4;
   d. Current exacerbated episode started within 4 weeks of screening;
   e. Sufficient history and/or independent reporter (such as a family member, caregiver, outside practitioner) must verify that current state is an exacerbated state for the individual subject;
   f. Illness of the subject is confirmed at baseline with a PANSS score of 70 or higher and baseline CGI-S score of 4 or higher before randomization;
3. has a history of at least three months exposure to one or more antipsychotic therapy(ies) and a prior response to antipsychotic therapy within the previous five years where response is defined as a clinically significant decrease in delusions and/or hallucinations during an exacerbated episode;
4. body mass index (BMI) of 19 – 40 $kg/m^2$, inclusive, at screening;
5. female subjects must be of non-childbearing potential (defined as either permanently sterilized or post-menopausal; the latter is defined as at least 1 year with no menses without an alternative medical explanation) or must use highly effective methods of birth control (defined as those, alone or in combination, that result in a low failure rate [i.e., less than 1 percent per year] when used consistently and correctly) from Screening through to the End-of-Study visit, if subject has sexual intercourse during the trial;

6. subject must be able to provide informed consent, where the subject are fluent and literate in a language spoken by the Investigator and staff;
7. subject must be willing to be hospitalized for the duration of the inpatient period of the study, and willing to comply with all Investigator and staff instructions.

### 3.2.2 Exclusion Criteria

Subjects presenting with any of the following will not be included in the study:

1. any subject unable to provide informed consent;
2. any female subject who is pregnant or breast-feeding; female subjects of childbearing potential must have a negative urine pregnancy test at screening and on Day 1 prior to study treatment administration;
3. any subject who is identified as a duplicate subject on the DupCheck.org website and on other methods for identifying duplicate subjects;
4. any subject presenting with concurrent dementia or suspicion thereof, delirium, mental retardation, epilepsy, drug-induced psychosis, history of significant brain trauma, history of prolonged loss of consciousness or history of seizure disorder other than single seizure episode;
5. any subject presenting with major depressive disorder with psychotic features or bipolar disorder with psychotic features, schizoaffective disorder, schizophreniform disorder and brief psychotic disorder, or delusional disorder [Note: comorbid symptoms of depression are allowed and should be well-characterized and documented according to the SCID-CT];
6. any subject presenting with symptoms that meet DSM-5 criteria for any moderate or severe substance use disorder within the 6 months prior to screening (including for cannabis);
7. a positive qualitative urine drug or alcohol test at screening, or evidence of either withdrawal from, or acute intoxication with cocaine, opiates, (meth)amphetamines, alcohol, barbiturates, or hallucinogens or similar compounds;

   Note: Previous occasional use of alcohol or cannabis is allowed as long as the level of use does not meet DSM-5 criteria for any moderate or severe substance use disorder within the 6 months prior to Screening and the use of alcohol or cannabis is not considered to be the precipitating factor of the current psychotic episode in the opinion of the Investigator and confirmed by the Eligibility Adjudication process; subjects are required to abstain from alcohol and cannabis use during the trial. Any subject that tests positive for any other drug (besides alcohol or cannabis) at Screening will be excluded from participating in the study. One repeat of urine and

alcohol screen is allowed on a case-by-case basis after discussion and agreement with medical monitor.

8. any subject considered to be an imminent danger to themselves or others;
9. any subject reporting suicidal ideation of type 4 or 5 on the C-SSRS within 6 months prior to screening or any suicidal behavior in the last 2 years prior to screening, as indicated by any 'yes' answers on suicidal behavior section of C-SSRS;
10. any subject who has not had a stable living environment for at least 3 months before the current exacerbating episode;
11. use of depot antipsychotic within 1.5 treatment cycle prior to Study Day –1 (baseline);
12. use of any antipsychotic medication within the Screening Period through to Study Day –1 (tapering antipsychotic medication during screening period may be considered on a case-by-case basis, but must be approved in writing by the study medical monitor and be discontinued by Study Day -3);
13. use of the any of the following agents: mianserin, mirtazapine, nefazodone, cyproheptadine, pimavanserin, or fluvoxamine within 30 days prior to Study Day –1;
14. use of any strong or moderate CYP450 3A4 inhibitor or inducer (see Appendix D) within 7 days prior to the Day -1;
15. abnormal laboratory values or clinical findings on screening that are judged clinically significant, including, but not limited to
    a. QTcF > 450 ms and QTcB >450 ms,
    b. HR $\leq$ 55 beats per minute,
    c. Evidence of bundle branch blocks on ECG,
    d. ALT, AST, or CPK values > 3X upper limit of normal,

    Note: medical conditions that are stable with medication (e.g. hypertension, high cholesterol, and hyperthyroidism) are allowed as long as the condition has been stable for at least 3 months prior to screening, the medications are documented and kept stable during the study, and the condition is not thought to affect safe participation in the trial in the opinion of the Investigator and confirmed by the Eligibility Adjudication process
16. uncontrolled angina, or history of a myocardial infarction within three months prior to screening, or history of a clinically significant cardiac arrhythmia including antipsychotic drug-induced QTc prolongation;
17. presence, or history within 6 months prior to screening, of significant and/or uncontrolled hematological, renal, hepatic, endocrinological (including poorly controlled diabetes defined as HbA1c > 53 mmol/mol [7.0%] on screening and one repeat of screening serum

chemistry analysis & after discussion and agreement with the medical monitor), neurological, or cardiovascular disease;

18. prior history of neuroleptic malignant syndrome induced by any antipsychotic medication;
19. subjects with a history of HIV infection or demonstration of HIV antibodies;
20. subjects with a history of Hepatitis B or C infection and evidence of active disease defined as elevated ALT, AST or bilirubin levels >2x ULN;
21. subjects with demonstration of Hepatitis B or C antibodies at screening and evidence of active disease defined as elevated ALT, AST or bilirubin levels >2x ULN;
22. any subject who has a known or likely allergy to ITI-007 or any psychoactive drug based on history;
23. any subject who has participated in a prior clinical trial of ITI-007 with exposure to ITI-007;
24. any subject who has had exposure to any investigational product within 3 months of Day –1;
25. any subject who has received clozapine (verified by the clozapine patient registry);
26. any subject who is unable to be safely discontinued from current antipsychotic therapy, mood stabilizers, lithium, anticholinergics, and antidepressant medications;
27. any subject judged by the Investigator to be inappropriate for the study.

Appropriateness of subjects for the trial will be reviewed and approved by an independent psychiatrist (or panel for eligibility review). Once a final decision is made, no waivers for exceptions will be provided.

### 3.2.3 Removal, Replacement, or Early Withdrawals of Subjects from Therapy or Assessment

Every effort will be made to determine why any subject withdraws from the study prematurely, consistent with FDA recommendations. All subjects who withdraw from the study with an ongoing AE must be followed until the event is resolved or deemed stable. If a subject withdraws prematurely after dosing, all examinations scheduled for the End of Study evaluations must be performed on all subjects who receive study medication but do not

complete the study according to protocol (see *Section 6.2.12, End of Study procedures*). Subjects will be discharged from inpatient unit per standard-of-care treatment.

Subject participation may be terminated before completing the study and the reason recorded as follows:

- Adverse event
- Lack of efficacy
- Protocol violation
- Subject withdrew consent at own request – if consent is withdrawn, the subject must be questioned by the Investigator or site staff whether the withdrawal is due to an adverse event, lack of efficacy, personal or family reasons, or whether the subject withdrew consent and refused all End of Study procedures including refusing to give a reason; these reasons must be documented in the case report form, consistent with FDA recommendations
- In the Investigator's opinion, continuation in the study would be detrimental to the subject's well-being
- Other – reason must be clearly documented in the CRF in order to determine whether the withdrawal was associated with an adverse event or lack of efficacy

In all cases, the reason for withdrawal must be recorded in the case report form (CRF) and in the subject's medical records. The subject must be followed to establish whether the reason was an adverse event, and, if so, this must be reported in accordance with the procedures in *Section 8. Adverse Events.*

The Investigator must make every effort to contact subjects lost to follow-up.

The Sponsor may discontinue subject enrollment at any time, but will allow any ongoing subjects to complete the study as long as it is considered safe, in the opinion of the Investigator, to do so.

# 4.0 TREATMENTS

## 4.1 Identity of Investigational Product(s)

ITI-007 will be supplied as 20 mg and 60 mg capsules.

The composition of the size 0, Swedish orange capsules is listed in the **Table 1** below.

**Table 1 Composition of the Capsules**

| **Composition** | **Active (20 mg ITI-007) W/W%** | **Active (60 mg ITI-007) W/W%** | **Placebo** |
|---|---|---|---|
| ITI-007 | 6.67% | 20% | |
| [redacted] | [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] | [redacted] |
| [redacted] | [redacted] | [redacted] | |
| [redacted] | [redacted] | [redacted] | [redacted] |
| [redacted] | | | [redacted] |

Capsules are being manufactured by:

[redacted]

Capsules are being packaged by:

[redacted]

A certificate of analysis, GMP compliance statement and Product Release Certificate will be provided to the clinical site pharmacy.

## 4.2 Shipping, Storage and Handling

The Sponsor will supply study drug (ITI-007 and placebo) to all study sites. Study drug will be shipped under ambient conditions and should be stored at room temperature 15-30 °C (59-86 °F) in a secure and locked location. At randomization the clinic pharmacist or designated clinical staff member authorized to dispense medications will assign a study medication kit to each patient according to the randomization code and instructions provided by the Sponsor. Each day study medication will be dispensed from the appropriate kit by the clinic pharmacist or designated staff member.

## 4.3 Treatment Assignment

Randomization will occur on Day 1 immediately prior to dosing, after confirmation of continued eligibility. Patients will be randomized centrally in a 1:1:1 ratio to one of the three study treatments.

Assignment of a randomization number will be via an automated phone or web-based system. The system will assign a unique randomization number for each patient and randomize the patient to 1 of the 3 treatment arms. This patient identification number will be recorded on all electronic case report forms (eCRFs) and correspondence regarding the patient. Further information and instructions regarding the randomization process will be provided in the study reference manual.

## 4.4 Dose, Dose Schedule and Route of Administration

ITI-007 will be administered orally as a powder-filled capsule, once daily for 28 consecutive days. Treatment will be administered between 8:00 am and 10:30 am during breakfast; furthermore, study treatment will be administered at the same time each day ± 1 hour. Study treatment will include dosing from Study Day 1 through Day 28. The dosing is shown in Table 2.

### Table 2: Dosing Schedule

| Study Drug | Day 1-28: |
|---|---|
| **60 mg ITI-007** | 1 capsule, 60 mg ITI-007 and<br>1 capsule, placebo = 2 capsules<br>**Daily Dose:** 60 mg ITI-007 |
| **40 mg ITI-007** | 2 capsules, 20 mg ITI-007<br>**Daily Dose:** 40 mg ITI-007 |
| **Placebo** | 2 capsules, placebo<br>**Daily Dose:** 0 mg ITI-007 |

ITI-007 will be provided to the clinical site by [redacted] on behalf of the Sponsor. Each ITI-007 dosing container will be labeled according to the local law.

After dose administration, empty containers should be returned to the pharmacy. Empty containers should be stored in the pharmacy until accountability is documented by the clinical monitor. After accountability is documented, empty containers will be shipped to [redacted]. Documentation regarding accountability and disposition of the clinical supplies and empty containers should be placed in the Trial Master File. Any unused clinical supplies should be returned to [redacted] at the end of the study.

## 4.5 Blinding, Packaging, Labeling and Retention of Supplies

The Investigators, study staff at the clinical site, patients, primary study team at the contract research organization (CRO), and the Sponsor will be blinded to study treatment. At the Sponsor, clinical site, and CRO, selected individuals may be unblinded to the study treatments on a need to know basis as described in study documentation and/or standard operating procedures (SOPs) of the site and CRO. Such individuals may include personnel generating the randomization scheme, the clinical supply packaging and labeling team, and the bioanalytical personnel handling and analyzing the pharmacokinetic samples. The study will be unblinded following database lock according to the CRO's SOPs.

If unblinding on an individual basis becomes necessary due to safety reasons, the Investigator will attempt to contact the Sponsor or representative at the CRO before unblinding. See *Section 9.2 Emergency Identification of Study Medication* for details on the unblinding procedure.

Each patient kit will include four dose cards, one for each week of treatment. The dose card for each study treatment will contain 1x8 strips of capsules as described below:

**Treatment Cards (4 Cards per patient kit):**

| | |
|---|---|
| Placebo: | Two 1x8 strips of placebo capsules = 16 capsules |
| 60 mg ITI-007: | One 1x8 strip of 60 mg ITI-007 capsules = 8 capsules |
| | One 1x8 strip of placebo capsules = 8 capsules |
| | Total = 16 capsules |
| 40 mg ITI-007: | Two 1x8 strips of 20 mg ITI-007 capsules = 16 capsules |

Study treatment will be administered once daily for 28 days.

On Study Days 1 – 28, the clinic pharmacist or designated clinical staff member authorized to dispense medications will dispense two capsules each morning from the dose treatment cards. Labeling will indicate that each dose is to be one capsule from each strip.

Each dose treatment card will contain 16 capsules. Two capsules will be taken each day. There is one dose treatment card for each week of study treatment, so each card contains 7 days per week x 2 capsules/dose, plus an extra dose of 2 capsules in case a capsule is dropped or lost.

The Investigator or study Pharmacist or qualified designee is responsible for taking an inventory of each shipment of investigational supplies received and comparing it with the accompanying drug accountability form. The Investigator or Pharmacist or qualified designee will verify the accuracy of the information on the form, sign and date it, and return the form to the clinical study monitor for retention in the trial master file. Sites must retain accurate, original site records of study drug inventory as well as copies of all invoices of

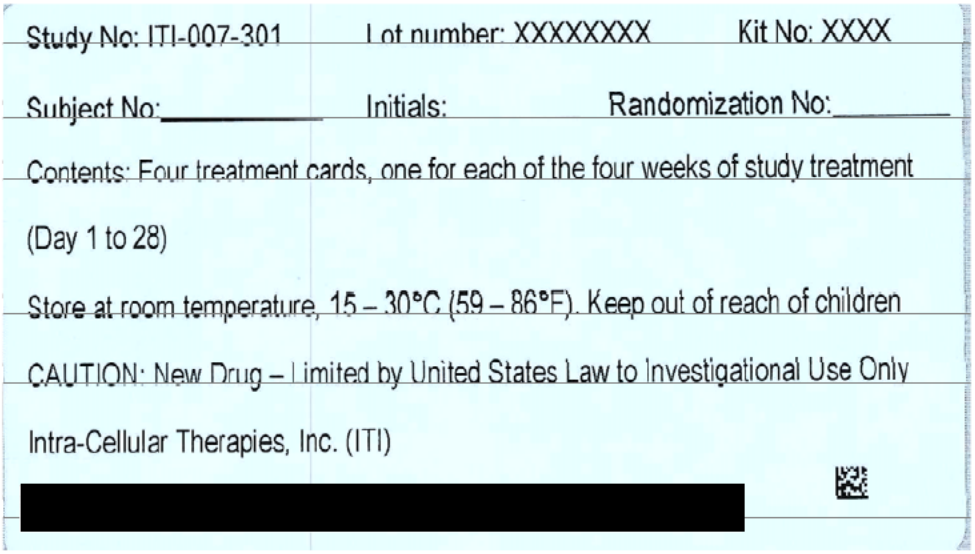
study drug shipments and records of study drug disposition. All investigational drug supplied is for use only in this study and should not be used for any other purpose.

The investigational drug product must be kept in a securely locked area with restricted access. Neither the Investigator, Pharmacist, qualified designee nor any of his or her designees may provide investigational drug to any person not participating in the study. The investigational drug must be stored and handled in accordance with the manufacturer's instructions.

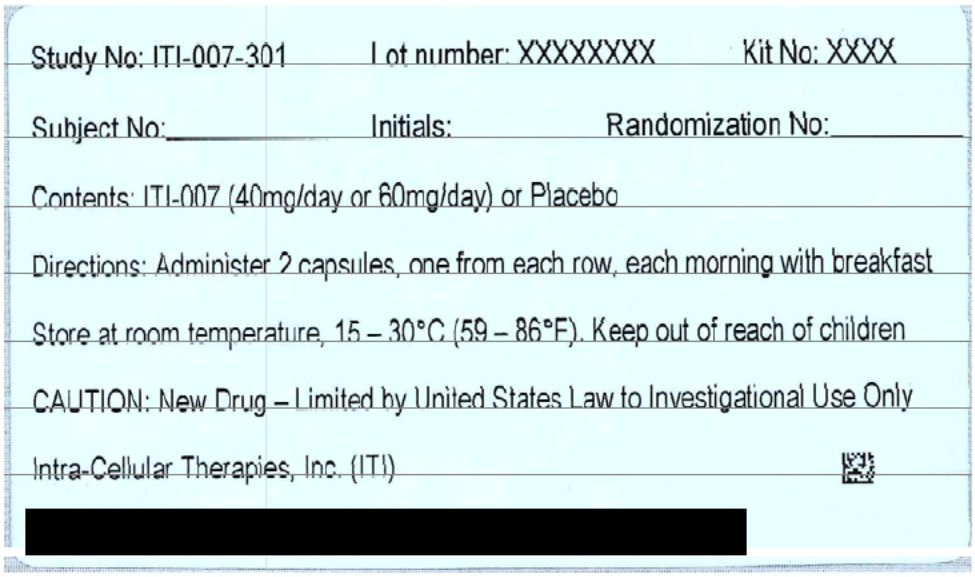
The Investigator or study Pharmacist or qualified designee will keep accurate records of the quantities of the investigational drug product dispensed, used, and returned by each subject. At the conclusion of the study, final drug reconciliation will be conducted at the clinic. Final drug accountability documentation will be maintained at both the clinic and in the trial master file. Each site must keep all used and unused study drug until the Study Monitor either arranges return to the Sponsor or their designee or gives instruction for their disposal.

All treatment cards and kits will be labeled in a manner consistent with local and Federal laws.

Clinical Supply Kit – Carton Label Text:



Clinical Supply Treatment Card Label Text (4 Cards to Kit):




## 4.6 Treatment Compliance

The treatments will be administered by qualified study personnel under the direction of the Investigator. A witness will verify that the treatment has been given correctly and that the patient has indeed swallowed the capsule(s). Details of the exact date and time of administration (day/month/year, h: min) will be documented in the eCRF. In the event the Investigator or study staff suspects a subject's non-compliance with study medication, the appropriate study personnel could perform hand and mouth checks.

Given the importance of treatment compliance to the success of the clinical study, compliance will be evaluated by reviewing pharmacokinetic data in an ongoing, blinded manner. If a subject assigned to receive ITI-007 is shown to have blood levels that are below the limit of quantitation, indicating likely non-compliance with study treatment, then the Sponsor may request an investigation by an unblinded third party to determine whether the subject should be discontinued from the study as per *Section 3.2.3*.

## 4.7 Prior and Concomitant Treatments

Starting on the first day of Screening Period and as applicable, after written informed consent for the study has been provided, all antipsychotic or other psychotropic medications will be

stopped with clinically appropriate titration down as per Investigator clinical judgment. Thereafter, subjects will be instructed not to take any medications during the conduct of the study unless approved by the Investigator. Any patient who is unable to be safely discontinued from current antipsychotic therapy or other psychotropic medications will not be eligible for the study.

Patients will not be eligible for the study if they have used any depot antipsychotic within one treatment cycle prior to Day –1 (baseline), have used mianserin, mirtazepine, nefazadone, cyproheptadine, or fluvoxamine within 30 days prior to the Day -1, or have used any strong or moderate CYP3A4 inhibitor or inducer (see Appendix D) within 7 days prior to the Day -1.

Any treatment given, other than the study treatment, is to be considered a concomitant medication and must be documented in the case report form. Any subject who requires emergency treatment with prescription or nonprescription medications during the Screening Period through to the End-of Study visit will be approved by the Investigator. If the Investigator determines that the subject requires medication during the study, the project medical officer (or designee) or clinical monitor must be contacted within 48 hours.

During the study lorazepam may be administered for agitation or anxiety or to aid sleep; other medications to manage insomnia are not allowed, including but not restricted to zaleplon and zolpidem. Lorazepam, benztropine and propanolol are allowed according to the instructions in Table 3 below. Within the daily limits shown in Table 3, lorazepam, benztropine, and propanolol are not allowed on a PRN basis; each and every administration must be considered and approved by the Investigator prior to administration.

The adverse event associated with the need for lorazepam, benztropine, or propanolol should be recorded on the Adverse Events CRF. The amount and timing or each lorazepam, benztropine, or propanolol administration should be recorded for each patient.

Table 3: Restrictions on the Use of Lorazepam, Benztropine, and Propanolol through Study Day 28

| Medication | Study Period | Dosage Allowed | Timing Restrictions |
|---|---|---|---|
| **Lorazepam** | Screening Period through to Day 7, inclusive | Maximum of 6 mg/day | Not within 8 hours prior to PANSS, CGI, PSP or CDSS |
| | Day 8 through to Day 14, inclusive | Maximum of 4 mg/day | |
| | Day 15-Day 28, inclusive | Maximum of 2 mg/day on no more than 4 days/week | |
| **Benztropine** | Screening Period through to Day 28, inclusive | Maximum of 4 mg/day | Not within 8 hours of SAS, BARS, or AIMS |
| **Propanolol** | Screening Period through to Day 28, inclusive | Maximum of 40 mg/day | |

During the End-of-Stabilization period, the use of Lorazepam, Benztropine, and Propanolol is not restricted.

# 5.0 PHARMACOKINETIC, PHARMACODYNAMIC AND SAFETY ANALYSIS VARIABLES

## 5.1 Pharmacokinetic

Pharmacokinetic (PK) plasma samples will be analyzed for concentrations of ITI-007 and its metabolites, as applicable based on randomization. Blood samples will be collected prior to dosing, and then 3-4 h post-dose on Study Days 1, 8, and 28 and again at 6-7 h post-dose on Study Day 28. A blood sample will also be collected once on Day 33 (or last day of inpatient stabilization period, if earlier than Day 33).

Select samples from patients randomized to placebo will be analyzed for presence/absence of active drug to ensure appropriate randomization.

In addition, ongoing blinded monitoring of pharmacokinetic data will inform evaluation of study treatment compliance. If patients assigned to receive ITI-007 are shown to be non-compliant with study treatment, as indicated by levels of study drug below the limit of quantitation, then patients may be discontinued from the study.

## 5.2 Pharmacodynamic

The PANSS will be determined by remote central independent raters; CGI-S will be determined by both a remote central independent rater and a qualified site-based rater; all other assessments will be site-based assessments by qualified individuals.

Change from baseline scores will be calculated for:

- Total PANSS to assess psychosis, as well as
  - PANSS Social Factor
  - PANSS Positive subscale
  - PANSS Negative subscale
  - PANSS General Psychopathology subscale
- PSP to assess social function
- CGI-S to assess global disease severity
- CDSS to assess symptoms of depression

## 5.3 Safety

Tolerability will be assessed by reported and observed adverse events. Frequency and severity of adverse events will be summarized.

Routine safety assessments will be performed throughout the study, with a final assessment performed when a subject completes the study on Day 42 or if a subject discontinues. Safety assessments will include: modified physical examinations including height, weight, waist circumference, 12-lead electrocardiograms, vital signs (3-positional blood pressure and pulse rate, respiration rate, and oral body temperature), clinical laboratory tests (hematology, serum chemistry, and urinalysis), and pregnancy test (where applicable). Motor tolerability and safety will be assessed by the SAS, BARS and AIMS and suicidality will be evaluated by the C-SSRS.

## 5.4 Exploratory

The exploratory objectives of the study include, but are not limited to, the evaluation of the association between response to treatment and any genetic trait or change from baseline in protein biomarkers, such as p11.

# 6.0 STUDY PROCEDURES AND SCHEDULE

## 6.1 Study Events Schedule

The study consists of a Screening visit (Day –7 to Day -2), Baseline (Day -1), Study Treatment period (Day 1 to up to Day 28), End of Stabilization Period (Day 33) and End-of-Study Visit (Study Day 42 ± 2 days). Therefore, the maximal study duration for an individual subject will be approximately up to 7 weeks.

Please see Table 4 for the schedule of assessments.

## TABLE 4 SCHEDULE OF ASSESSMENTS

| Event | Screening Period | Baseline | Study Treatment Period | | | | End-of-Stabilization Period | End-of-Study Visit |
|---|---|---|---|---|---|---|---|---|
| | -7 to -2 [a] | -1 | 1 | 8 | 15 | 28 | 33 [b] | Day 42 ± 2 |
| Visit Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Informed consent | Before Any Study-Specific Procedures are Conducted | | | | | | | |
| Admission to Inpatient Unit | X | | | | | | | |
| Medical History | X | | | | | | | |
| Inclusion/Exclusion Criteria Review | X | X | | | | | | |
| SCID-CT [c] | X | | | | | | | |
| BPRS [c] | X | | | | | | | |
| Screening Adjudication Form(s) and Recordings [c] | X | | | | | | | |
| Modified Physical Exam | X | | | | | X | | |
| Hepatitis/HIV Testing | X | | | | | | | |
| Urine Drug and Alcohol Screen [d] | X | | X | | | | | |
| Urine pregnancy Test | X | | X | | | X | | |
| Clinical Laboratories [e] | X | | X | X | | X | X | |
| 12-Lead ECG [f] | X | X | X | X | | X | X | |
| Vital Signs [g] | X | X | X | X | X | X | X | |

| Event | Screening Period | Baseline | Study Treatment Period | | | | End-of-Stabilization Period | End-of-Study Visit |
|---|---|---|---|---|---|---|---|---|
| | -7 to -2 [a] | -1 | 1 | 8 | 15 | 28 | 33 [b] | Day 42 ± 2 |
| Visit Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Randomization | | | X | | | | | |
| Study Drug Dosing | | | X | X | X | X | | |
| PK Sample Collection [h] | | | X | X | | X | X | |
| Samples for protein biomarkers [i] | | | X | | | X | | |
| Sample for genetic testing [j] | | | X | | | X[i] | | |
| PANSS [k] | | X | | X | X | X | | |
| Central CGI-S [k,l] | | X | | X | X | X | | |
| Site-based CGI-S [l] | X | | | X | X | X | X | X |
| PSP | | X | | | | X | | |
| CDSS | | X | | | | X | | |
| SAS [m] | | X | | X | | X | | |
| BARS [m] | | X | | X | | X | | |
| AIMS [m] | | X | | X | | X | | |
| C-SSRS | X | X | | X | X | X | X | X |
| Concomitant medication | X | X | X | X | X | X | X | X |
| Adverse events [n] | X | X | X | X | X | X | X | X |
| Restart Standard Antipsychotics [o] | | | | | | X | | |

a) Upon signing of the ICF, patients are to be admitted to the inpatient unit, if they are not inpatient already.

b) The End-of-Stabilization period is up to 5 days
c) These assessments are to be recorded and submitted to an independent psychiatrist or clinical psychologist (or panel of psychiatrists and/or clinical psychologists) for review.
d) The Urine Drug and Alcohol screen will be also performed upon patient's return to the inpatient unit from the day pass.
e) Clinical Laboratory samples are to be taken after an overnight fast of at least 10 hours.
f) ECGs are to be triplicate 10 second epochs with 5 min between recordings. One triplicate ECG will be performed during Screening and one triplicate ECG will be performed on Day 33. On Days 1, 8, and 28, triplicate ECGs will be taken pre-dose (trough, before breakfast) and at 3-4 h post-dose (peak, before lunch). An additional triplicate ECG will be performed on Day 28 at 6-7 h post-dose. Triplicate ECGs on Day -1 are to be performed at the same time of day planned for Day 28: once before breakfast (approximately 7:00 am corresponding to the same time of day for the trough measure), again before lunch (approximately 11:00 am – noon, corresponding to the same time of day for the 3-4 h post-dose peak measure), and again approximately 2:00-3:00 pm corresponding to the same time of day of the 6-7 h post-dose measure. In all cases ECGs are conducted before other assessments scheduled in the same time window; for example, when ECG, vital signs and blood sample collection for PK measures are scheduled for the same time window, ECG measures should be conducted first, followed by vital signs and then blood sample collection.
g) Respiratory rate, oral temperature, and 3-positional blood pressure and pulse (after at least 10 minutes lying down, after approximately 1 minute sitting, immediately upon standing, and after approximately 3 minutes standing) will be taken at least once at all scheduled visits. On Days 1, 8, 15, 22, and 28 vitals will be taken pre-dose (trough) and 3-4 h post-dose (peak). Additionally, vital signs will be assessed on Days -1 and 28 at 6-7 h post-dose. Vital signs are always taken after conducting the ECGs, as applicable, and prior to any other assessments scheduled in the same time window.
h) Blood samples for PK collection are to be collected prior to dosing (trough) and then 3-4 h after dosing (peak) on Days 1, 8, and 28, again at 6-7 h post-dose only on Day 28, and then once on Day 33. Samples are collected after ECGs and vital signs.
i) Samples for protein biomarkers on Day 1 has be collected pre-dose.
j) Blood sample is to be collected once for genetic testing, if the optional ICF addendum is signed. If the addendum is signed prior to Day 1 then the sample will be taken on Day 1. If it signed after Day 1, then the sample will be taken on Day 28. It is not to be taken twice.
k) PANSS and CGI-S are to be conducted by a remote central rater. Efforts should be made to schedule the remote interviews at approximately the same time of day for each visit. Due to potential limitations around scheduling of the remote rater, the PANSS may be conducted within a 1-day window provided for this assessment (see *Section 6.2*). PANSS and CGI scores at Baseline Visit are reviewed for inclusion of subject (subject will be included with the baseline PANSS score of 70 or higher and baseline CGI-S score of 4 or higher).
l) CGI-S is conducted by a qualified site rater at screening (in addition to the CGI-S conducted by the remote central rater) at baseline through Day 28. Site-based CGI-S is conducted by qualified site rater; during the treatment period, the site rater will interview the subject and review baseline and current PANSS and CGI-S scores from the remote centralized rater prior to rating the site-based CGI-S for the visit.
m) On Days 8, 15, and 28, the SAS, BARS, and AIMS assessments are to be conducted 3-6 hours after the administration of study treatment
n) The recording of adverse events is to start immediately after the ICF is signed and continue through until End-of-Study, Visit 9.
o) On Day 28, after all study assessments have been completed, or on Day 29 patients are to be started on standard antipsychotic medication as prescribed by the Investigator.

## 6.2 Procedures by Visit

### 6.2.1 Visit 1: Screening

After obtaining written informed consent, the following assessments are to be performed 2-7 days prior to Day 1; assessments can be conducted on different days within the screening period.

Eligibility of potential subjects will be confirmed through a formal adjudication process in which screening data obtained to evaluate psychiatric status (including CGI-S, BPRS, SCID-CT, DSM-5 diagnosis) is reviewed by a member of an independent Eligibility Adjudication Committee (EAC; Clintara). In conjunction with completing the assessments the investigator or designated staff member will complete a summary of the information called the Clinical Validation Inventory for Study Admission (C-VISA©) and submit it to the EAC. The scales and C-VISA© form will be completed using the Clintara audio-digital pen which will digitally record the completion of the forms and also create audio recordings of the interviews conducted with the subject in order to make the assessments. Both the forms and the audio recordings will be transferred to Clintara EAC.

The EAC review process will involve review of all information by a Tier 1 independent reviewer who will confirm eligibility of the subject as assessed by the Investigator. If the Tier 1 reviewer disagrees with the Investigator's assessment of a subject as eligible for participation in the study, a Tier 2 expert reviewer will review the data and in communication with the Investigator will make a final decision regarding eligibility of a subject. Once the final EAC decision is made, no waivers for exceptions will be provided. Note: The Sponsor has no involvement in the adjudication process.

Subjects will be checked for duplicate enrollment by site staff using the [redacted] website and on other methods for identifying duplicate subjects.

The following procedures are to be conducted during the Screening Period:

- Informed Consent
- DSM-5 diagnosis of schizophrenia
- confirmation of diagnosis by the modified SCID-CT

- BPRS
- CGI-S, conducted by the Investigator or qualified site-based rater
- medical history
- C-VISA© completion
- C-SSRS
- modified physical examination, including standard neurological examination, weight, waist circumference, and height measurement; excluding genital/rectal exam
- review of past and current medications; for past medications collect information on antipsychotics up to 5 years prior to Day -1 and all other medications up to 6 months prior to Day -1
- 12-lead electrocardiogram (ECG), prior to vital signs and blood/urine sample collection
- vital signs, after ECG and prior to blood/urine sample collection
- clinical laboratory sample collection, fasted
- HepB, HepC, HIV testing
- urine drug and alcohol testing
- urine pregnancy test (as applicable)
- review of inclusion and exclusion criteria by Investigator
- submission of applicable data to EAC for confirmation of eligibility

### 6.2.2 Visit 2: Study Day -1

At Study Day -1, the following assessments will be performed:

- 12-lead ECG, prior to any other assessments; ECG to be performed at about the same time of day as it would be during the study treatment period on Study Day 28 [pre-dose, at 3-4 h post-dose (before lunch) and again at 6-7 h post-dose]
- 3-positional vital signs, after ECG and prior to remaining procedures; vital signs to be performed at about the same time as it would be during the study treatment period on Study Day 28 at pre-dose, 3-4 h post-dose (before lunch) and again at 6-7 h post-dose
- review of concomitant medications
- adverse event assessment
- PANSS, conducted by a remote centralized rater; subjects with PANSS score is 70 or above will be allowed to continue study participation

- CGI-S, conducted by a remote centralized rater; subjects with CGI-S score of 4 or higher will be allowed to continue study participation
- SAS
- BARS
- AIMS
- CDSS
- C-SSRS
- PSP

Whenever possible, the assessments should be performed at the same time of the day and in the same order of the assessments for Study Visit 2 through Study Visit 7. However, in each instance, the ECG will be performed prior to 3-positional vital signs, and phlebotomy to be performed after ECG and vital signs.

### 6.2.3 Visit 3: Study Day 1

On Study Day 1 the following assessments will be performed:

- 12-lead ECG, pre-dose and 3-4 hour post-dose before lunch; prior to any other assessments scheduled within same time window
- 3-positional vital signs, pre-dose and 3-4 h post-dose; after ECG and prior to remaining procedures scheduled within same time window
- pharmacokinetic sample collection, pre-dose (fasted), 3-4 h post-dose
- clinical laboratory sample collection, fasted
- urine drug and alcohol screen
- blood samples for biomarker proteins (eg, p11) and pharmacogenomics
- urine pregnancy test (as applicable)
- randomization
- study drug administration, during breakfast eaten between 8:00 and 10:30 am
- review of concomitant medications
- adverse event assessment

Whenever possible the assessments should be performed at the same time of the day and in the same order of the assessments for Study Visit 2 through Study Visit 7. However, in each instance, the ECG will be performed prior to 3-positional vital signs, and phlebotomy to be performed after ECG and after vital signs.

### 6.2.4 Study Days 2 through Day 7

On Days 2, 3, 4, 5, 6, and 7 each patient should receive the daily administration of study treatment. No other assessments are scheduled and therefore no formal visits are shown in the Schedule of Assessments. However, adverse events observed or reported and concomitant medications will be recorded. If any assessments are required in order to evaluate an adverse event an unscheduled visit will be conducted.

### 6.2.5 Visit 4: Study Day 8

On Study Day 8 the following assessments will be performed on the same day as study drug administration:

- 12-lead ECG, pre-dose and 3-4 h post-dose before lunch; prior to any other assessments scheduled within same time window
- 3-positional vital signs, pre-dose and 3-4 h post-dose; after ECG and prior to remaining procedures scheduled within same time window
- pharmacokinetic sample collection, pre-dose (fasted) and 3-4 h post-dose
- clinical laboratory sample collection, fasted
- study drug administration, during breakfast eaten between 8:00 and 10:30 am
- SAS, 3-6 h post-dose
- BARS, 3-6 h post-dose
- AIMS, 3-6 h post-dose
- review of concomitant medications
- adverse event assessment

The following Study Day 8 assessments can be performed on Day 8 ± 1 day and yet should all be performed on the same day.

- PANSS, CGI-S conducted by a remote centralized rater
- CGI-S conducted by the Investigator or qualified site-based rater
- C-SSRS

Whenever possible the assessments should be performed at the same time of the day and in the same order of the assessments for Study Visit 2 through Study Visit 7. However, in each instance, the ECG will be performed prior to 3-positional vital signs, and phlebotomy to be performed after ECG and vital signs.

### 6.2.6 Study Days 9 through 14

On Days 9, 10, 11, 12, 13, and 14 each patient should receive the daily administration of study treatment. No other assessments are scheduled and therefore no formal visits are shown in the Schedule of Assessments. However, any adverse events observed or reported and concomitant medications will be recorded. If any assessments are required in order to evaluate an adverse event an unscheduled visit will be conducted.

### 6.2.7 Visit 5: Study Day 15

On Study Day 15 the following assessments will be performed:

- 3-positional vital signs, pre-dose and 3-4 h post-dose; prior to all other procedures
- study drug administration, during breakfast eaten between 8:00 and 10:30 am
- review of concomitant medications
- adverse event assessment

The following Study Day 15 assessments can be performed on Day 15 ± 1 day and yet should all be performed on the same day.

- PANSS, CGI-S conducted by a remote centralized rater
- CGI-S conducted by the Investigator or qualified site-based rater
- C-SSRS

Whenever possible the assessments should be performed at the same time of the day and in the same order of the assessments for Study Visit 2 through Study Visit 7. However, in each instance, the ECG will be performed prior to 3-positional vital signs, and phlebotomy to be performed after ECG and after vital signs.

### 6.2.8 Study Days 16 through 27

On Days 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 and 27 each patient should receive the daily administration of study treatment. No other assessments are scheduled and therefore no formal visits are shown in the Schedule of Assessments. However, any adverse events observed or

reported and concomitant medications will be recorded. If any assessments are required in order to evaluate an adverse event an unscheduled visit will be conducted.

### 6.2.9 Visit 6: Study Day 28

On Study Day 28 and, if possible, when any patient discontinues prior to this visit, the following assessments will be performed:

- 12-lead ECG, pre-dose and 3-4 hour post-dose, before lunch, and again at 6-7 hours post-dose; prior to any other assessments scheduled within same time window
- 3-positional vital signs, pre-dose and 3-4 h post-dose and again at 6-7 h post-dose; after ECG and prior to remaining procedures scheduled within same time window
- Clinical laboratory sample collection, fasted
- blood samples for biomarker proteins (eg, p11)
- pharmacokinetic sample collection, pre-dose (fasted) and 3-4 h post-dose and again at 6-7 h post-dose
- final study drug administration, during breakfast and between 8:00 and 10:30 am
- SAS, 3-6 h post-dose
- BARS, 3-6 h post-dose
- AIMS, 3-6 h post-dose
- modified physical examination, including weight and waist circumference, excluding genital/rectal exam and including standard neurological examination
- urine pregnancy test (as applicable)
- review of concomitant medications
- adverse event assessment

The following Study Day 28 assessments can be performed on Day 27 or Day 28 and yet must be conducted on the same day.

- PANSS, CGI-S conducted by a remote centralized rater
- CGI-S conducted by the Investigator or qualified site based rater
- PSP

- CDSS
- C-SSRS

Whenever possible the assessments should be performed at the same time of the day and in the same order of the assessments for Study Visit 2 through Study Visit 7. However, in each instance, the ECG will be performed prior to 3-positional vital signs, and phlebotomy to be performed after ECG and after vital signs.

If optional ICF addendum is signed after Day 1, the blood sample for genetic testing is collected on Day 28. The following can be performed on Day 28 *after all study related procedures have been completed* or on Day 29, at the discretion of the Investigator.

Restart of standard-of-care antipsychotics.

### 6.2.10 Study Days 29 through 32

Patients are to remain inpatient and be re-stabilized on standard-of-care antipsychotics which may be started on Day 29 or on Day 28 (see above). No other assessments are scheduled and therefore no formal visits are shown in the Schedule of Assessments. However, any adverse events observed or reported and concomitant medications will be recorded. If any assessments are required in order to evaluate an adverse event an unscheduled visit will be conducted.

### 6.2.11 Visit 7: Study Day 33

On Study Day 33 (or the last day of stabilization if it occurs before Day 33) the following assessments will be conducted:

- 12-lead ECGs, prior to any other assessments
- 3-positional vital signs (following 12-lead ECGs)
- clinical laboratory sample collection, fasted
- PK sample, fasted
- urine pregnancy test (females of child-bearing potential only)
- CGI-S conducted by the Investigator or qualified site-based rater
- C-SSRS
- review of concomitant medications
- adverse event assessment

Based on known pharmacokinetic properties of ITI-007 it is expected that within 5 days of the end of the Study Treatment Period, circulating levels of ITI-007 and its metabolites will have

diminished to below detection. Therefore, as of Day 33 direct exposure to ITI-007 and its metabolites will be minimal and patients will be considered discharged from the inpatient period of the study.

### 6.2.12 Visit 8: End-of-Study

On Day 42 ± 2 days, the following procedures will be conducted:

- CGI-S conducted by the Investigator or qualified site-based rater
- C-SSRS
- review of concomitant medications
- adverse event assessment

## 6.3 Study Procedures

### 6.3.1 Informed Consent

Prior to any study-related activities the patient must sign and date and Institutional Review Board (IRB)-approved Informed Consent Form (ICF). The format and content of ICF must be agreed upon by the Investigator, the appropriate IRB, and the Sponsor.

A separate ICF will be provided for the collection of samples to be used in the determination of genetic biomarkers. Patients may withhold consent to provide such samples and still participate in the study without prejudice.

The original signed ICFs (together with any subsequent IRB-approved amended versions) must be retained by the Investigator in the patient's study file. A copy of the original and date ICF must be given to the patient.

### 6.3.2 Medical History and Other Information

Medical history information will be collected at the Screening Visit and should includ (but not be limited to) demographic information, current and past medical conditions, current and past medications. The medical history must be documented in the patients study chart prior to study treatment administration and also recorded in the appropriate eCRF.

In addition to conventional medical history, information pertaining the patients' average alcohol and caffeine consumption and average tobacco usage should be recorded on the appropriate CRF.

### 6.3.3 Modified Physical Examination

A modified physical examination, *including neurological and excluding genital/rectal exams*, will be performed on the patient during the Screening Visit and on Day 28. The examination should include:-

- Height (Screening only)
- Body weight
- Waist circumference
- Appearance
- Skin
- Head-neck
- Eyes-ears-nose-throat
- Lungs-chest
- Heart
- Abdomen
- Extremities

All physical exam findings must be documents in the patients study chart and also recorded on the appropriate CRF.

Note: A directed physical exam may be performed on patients at additional time points if indicated by safety monitoring.

### 6.3.4 Electrocardiogram

Electrocardiogram (ECG) assessments will be made at screening, baseline visit (time-matched to Day 28) and then pre-dose (trough) and 3-4 hours after dosing on Days 1, 8, and 28 with an additional measure on Day 28 at 6-7 hours post-dose. The 3-4 hour window for the ECG assessment represents the approximate Tmax for plasma concentrations of ITI-007 parent (IC200056) and its active metabolites, IC200131 and IC200161. Based on the ITI-007-006 study, ITI-007 administered orally as a formulated capsule with food resulted in a median Tmax

of the ITI-007 parent molecule (IC200056) of 3.5 h, and a median Tmax of its metabolites IC200131 and IC200161 each at 4.0 h. Additionally, ECGs will be captured at baseline and on Day 28 at the 6-7 h post-dose time point at the approximate Tmax of the IC200565 metabolite. This schedule of assessments will ensure that ECG safety will be monitored at the approximate time of peak plasma concentrations and trough plasma concentrations throughout the study. A single assessment will be made on Day 33 and End-of Study Visit Day 42 ± 2 days. Each ECG assessment will comprise triplicate ten second epochs from 12-lead ECGs recorded 5 minutes apart. ECG parameters to be measured include HR, QRS, PR, QT, QTcB, QTcF and RR intervals.

The ECG recordings will be made on ECG machines supplied by a central ECG laboratory. ECG data will be transferred to the central ECG laboratory on the same day as collected and interpretation will be provided to the site within approximately 48 hours.

If any 12-lead ECG recording shows an arrhythmia other than a sinus arrhythmia, sinus tachycardia or sinus bradycardia, an additional 12-lead ECG will be recorded to confirm the original tracing. Any other clinically significant treatment-emergent cardiac conduction abnormalities will be followed until no longer deemed necessary by the Investigator.

Central interpretations of ECG recordings obtained at Screening will be the basis for determination that a patient is eligible for inclusion in the study. Similarly central interpretations of ECG recordings at baseline and all other visits will be included in the final study data. However, given that interpretations of recordings will not be available for up to 48 hours, Investigators are to use machine generated parameters and clinical judgment to assess cardiac function for the purposes of immediate safety concerns.

### 6.3.5 Vital Signs

Vital signs will be recorded at every scheduled visit and will include 3-positional blood pressure and pulse rate, respiratory rate, and oral temperature. The 3-positional blood pressure and pulse rate will be measured after 10 minute supine, 1 minute sitting, immediately after standing, and 3 minutes after standing. Pulse will be measured by counting pulse over 60 seconds.

Blood pressure and pulse should be measured after the ECG recordings have been made and prior to any other scheduled assessments.

Vitals signs should also be collected, if feasible, at the time of an adverse event such as vertigo, dizziness, fall, or any sign that might indicate a fall in blood pressure.

### 6.3.6 Hepatitis B and C Screen

Blood samples will be collected at the Screen Visit from all patients in order to perform Hepatitis B (HbsAg) and Hepatitis C (IgG) testing. Test results will send to each site and must be reviewed prior to the Day -1 Visit. Any patient that tests positive for Hepatitis B or C will be further evaluated and excluded from participating in the study if active disease is appreciated defined as ALT, AST, or bilirubin levels > 2 times the upper limit of normal. Further details regarding sample collection, processing and shipping can be found in the Laboratory Manual.

### 6.3.7 HIV Screen

Patients are required to provide blood samples for HIV virus types 1 and 2 testing. Test results will send to each site and must be reviewed prior to the Day -1 Visit. Any patient that tests positive for HIV will be excluded from participating in the study. Patients will be informed of positive HIV results and referred for follow up testing and counseling and health authorities will be notified consistent with federal, state and local laws. Further details regarding sample collection, processing and shipping can be found in the Laboratory Manual.

### 6.3.8 Drug and Alcohol Screen

Qualitative urine drug (urine drug or alcohol test at screening, or evidence of either withdrawal from, or acute intoxication with cannabis, cocaine, opiates, (meth) amphetamines, alcohol, barbiturates, or hallucinogens or similar compounds) and alcohol tests will be performed at screening and on Day 1. Any subject that tests positive for any drug (other than cannabis) or alcohol at Screening will be excluded from participating in the study. For any subject testing positive for cannabis, the cannabis use must be confirmed as having been subsequent to the psychosis exacerbation and not the cause of psychosis exacerbation. Furthermore, the subject must not meet DSM-5 criteria for moderate or severe substance use disorder for cannabis. Test

results obtained from Day 1 samples will be for information only. Further information regarding sample collection, processing and shipping can be found in the Laboratory Manual.

### 6.3.9 Urine Pregnancy Test

Female subjects, who are of childbearing potential, will undergo a urine pregnancy test at the Screen Visit, on Day 1 prior to administration of study treatment, and Day 28. The pregnancy test will be done at the study clinic using a urine dipstick.

If the pregnancy test at Screening or Day 1 is positive, the subject will not be eligible to participate in the study. If a urine pregnancy test performed after initiation of study treatment is positive it should be confirmed by a serum pregnancy test. Any pregnancies that occur in women who have received study treatment must be promptly reported to the Sponsor (see *Section 8.3.3.*).

Further details regarding sample collection and processing can be found in the Laboratory Manual.

### 6.3.10 Clinical Laboratory Tests

Blood and urine samples will be collected from all patients at the Screening, Day 1, Day 8, Day 28, and Day 33 Visits and forwarded to the Central Laboratory for analysis. Further details regarding sample collections, processing and specific testing can be found in the Laboratory Manual.

All samples for clinical laboratory analysis will be collected after an overnight fast (≥ 10 hours), *after any scheduled ECG or vital signs have been recorded*, and prior to dosing with study treatment.

The following laboratory variables will be determined as outlined below:

**Routine laboratory tests** (hematology, clinical chemistry, and urinalysis) will include:

- **Hematology**: The following hematology parameters will be assessed: hematocrit, hemoglobin, HbA1c, red blood cell count with indices (mean corpuscular volume [MCV], mean corpuscular hemoglobin [MCH], and mean corpuscular hemoglobin concentration [MCHC]), reticulocytes, white blood cell count and differential (neutrophils, lymphocytes, monocytes, eosinophils, and basophils) reported as absolute values, and platelets (platelet count, PT and PTT);

- **Clinical chemistry**: The following clinical chemistry parameters will be assessed: albumin, alkaline phosphatase, blood urea nitrogen (BUN), gamma-glutamyl transferase (GGTP), calcium, creatinine, glucose, insulin, cholesterol, triglycerides, phosphate, potassium, prolactin (blinded during the study conduct), AST (serum glutamic-oxaloacetic transaminase [SGOT]), ALT (serum glutamic-pyruvic transaminase [SGPT]), LDH, sodium, chloride, bilirubin (total, direct), total protein, uric acid CPK and thyroid panel (Screening only);
- **Urinalysis**: macroscopic (pH, specific gravity, glucose, protein, ketones, nitrates, blood) and microscopic – report only if present (RBCs/hpf, WBCs/hpf, casts, epithelial cells, crystals, granulation).

**Serology**: anti-HIV antibodies, HBsAg and Hepatitis C antibody. Serology will be performed at Screening.

Clinical laboratory test results should be within normal limits or for certain clinical laboratory tests, within an "acceptable" range (refer to Laboratory Manual).

The results will be forwarded to the Investigator for review. Any post-study treatment abnormalities deemed clinically significant by the Investigator, will be recorded as AEs in the patient's eCRF. Any clinically significant abnormal laboratory values that persist should be followed by the Investigator in consultation with the Sponsor (or designee). Such events should be followed until that have resolved or the Investigator assesses them to be chronic or stable.

The Investigator should file all copies of the reports in the patients study chart.

Note: Clinical laboratory testing may be performed at additional time points if indicated by AE monitoring.

Blood will be collected for clinical laboratory testing as outlined below:

- Hematology: Blood (4 mL) will be collected into an EDTA tube.
- Chemistry: Blood (8.5 mL) will be collected into a non-additive tube.
- Serology: Blood (8.5 mL) will be collected into a non-additive tube.

The total blood volume collected for clinical labs over the duration of the study will be approximately 71 mL. See Section 6.3.25 for the total blood sample collections (up to 169 mL), for clinical labs and pharmacokinetic samples and biomarker samples.

### 6.3.11 Pharmacokinetic Sample Collection

Blood samples will be collected for determination of plasma concentrations of ITI-007 and its metabolites. Blood samples will be collected on Days 1, 8, and 28 pre-dose and 3-4 hours after dosing with an additional collection 6-7 hours after dosing on Day 28. Furthermore, a sample will be collected on Day 33 to document that levels of ITI-007 are below the limit of quantitation. For the amount of blood and frequency of sampling, see Section 6.3.25 below.

Further details regarding sample collections, processing and specific testing can be found in the Laboratory Manual.

### 6.3.12 Sample Collection for Biomarker Measurements

On Day 1 (pre-dose for baseline) and Day 28 blood samples will be collected for preparation of PBMC which will then be used to evaluate p11, a member of the S100 family of proteins, which might be associated with an antidepressant therapeutic response to ITI-007 (Svenningsson and Greengard, 2007).

Additional blood samples may be collected from those patients who provide separate optional consent to have their samples analyzed for targeted genotyping. The Sponsor has identified approximately 12 genes in which single nucleotide polymorphisms have been associated with treatment effects in schizophrenia, but the whole genome may be sequenced for each subject. Blood samples will be collected and DNA extracted. The DNA isolated will be aliquotted into at least three separate archives; the archived samples will be stored until the end of the study. If, at the end of the study, the data indicate that ITI-007 has at least a tendency towards efficacy in the patient population, then analysis of targeted genotypes will be conducted on one set of archived samples. The other sample archives will be retained for future targeted genotyping based on additional information that may become available regarding other genes that may be associated with treatment effects in this patient population.

Further details regarding sample collections, processing and specific testing can be found in the Laboratory Manual.

### 6.3.13 SCID-CT

The Structured Clinical Interview for DSM-5 Disorders-Clinical Trial Version (SCID-CT) is a semi-structured interview for making the major DSM-5 Axis I diagnoses (First et al., 2014). It will be used in this trial at the Screening Visit only to confirm the diagnoses of schizophrenia in subjects evaluated for inclusion in the study. It will be completed by the Investigator or a trained clinical rater using the Clintara audio-digital pen recorder and be part of the C-VISA© evaluation (see *Section 6.2.1*).

### 6.3.14 BPRS

The 18-point Brief Psychiatric Rating Scale was developed by Overall and Gorham (1988) to measure changes in inpatients during clinical pharmacology trials. The BPRS is a clinician-based rating scale that emphasizes significant psychiatric symptoms including somatic concerns, anxiety, emotional withdrawal, conceptual disorganization, guilt feelings, tension, mannerisms and posturing, grandiosity, depressive mood, hostility, suspiciousness, hallucinatory behavior, motor retardation, uncooperativeness, unusual thought content, blunted affect, excitement, disorientation, suicidality, elated mood, bizarre behavior, distractibility, and motor hyperactivity. Ratings are made based on a 7-point Likert scale, from "Not Present" to "Extremely Severe." The scale is administered in context of a clinical interview and includes observation of subject's behavior.

The scale will be administered at the Screening visit only by the Investigator or a trained clinical rater and a patient will be eligible for participation in this study only if the total BPRS score at Screening is at least 40, and is comprised of a score of 4 or higher on at least two positive symptom core items (suspiciousness, conceptual disorganization, hallucinatory behavior, unusual thought content). The scale will be completed using the Clintara audio-digital pen recorder and be part of the C-VISA© evaluation (see *Section 6.2.1*)

### 6.3.15 PANSS

The PANSS is a 30-item scale used to measure symptoms of schizophrenia (Kay et al., 1987). The scale has seven positive symptom items, seven negative symptom items, and sixteen general psychopathology symptom items. Each item is scored on a seven-point scale by the clinical rater based on a clinical interview with the patient.

In this study patients will be rated on this scale by a remote independent central rater who will, through audiovisual communication with the patient, assess the patient using the Structured Clinical Interview for PANSS (SCI-PANSS) which was developed to increase objectivity and therefore decrease inter-rater and intra-rater variability (Kay et al., 1991). The central assessment of the PANSS requires site-based informant data that will be provided in a standardized format in advance of the interview. However, the remote independent central rater will remain blinded to study treatment and study visit. Additionally, a staff member from the site should stay in the same room as the subject during the entire remote interview and be available for a phone call with the central rater following the interview to address any discrepancies between the informant data and the interview, if needed.

The PANSS will be administered on Day -1 (baseline), Day 8, Day 15 and Day 28.

### 6.3.16 CGI

The Clinical Global Impressions (CGI) Scale (Guy, 1976) is a standardized assessment tool that the clinician can use to rate the severity of illness, change over time, and efficacy of medication, taking into account the patient's clinical condition and the severity of side effects. The CGI Scale consists of three global subscales, only two of which will be used in the present trial. The first subscale, Severity of Illness, assesses the clinician's impression of the patient's current illness state; it is often used both before and after treatment. Scores on the Severity of Illness subscale range from 1 = not ill at all to 7 = among the most extremely ill. The second subscale, the Global Improvement subscale, also goes from 1 = very much improved to 7 = very much worse. Global Improvement, assesses the patient's improvement or worsening from baseline, which is usually the beginning of a clinical trial. Sometimes a global improvement rating from the patient and the clinician is recorded.

The CGI only takes a clinician a minute or two to score following a clinical interview. A CGI-S assessment will be completed at screening by the Investigator or another Sponsor-approved trained clinical rater, using the Clintara digital-audio pen recorder and will be incorporated into the C-VISA evaluation. A site-based CGI-S will also be conducted on Study Days 8, 15, and 28 to determine treatment effects and again on Day 33 and at the final End of Study Visit for clinical follow-up. On Study Days -1, 8, 15 and 28 CGI-S will be conducted by a remote central rater (MedAvante) who will, through audiovisual communication with the patient, assess the

patient. The rater of the site-based CGI-S must conduct a psychiatric interview with the patient and this information must be combined with any other information available for the time period under study from such sources as chart notes, unit nurses, the screening BPRS and screening CGI-S as well as the all PANSS and central CGI-S ratings that are available at the time the site-based CGI-S scores are rated.

### 6.3.17 CDSS

The Calgary Depression in Schizophrenia scale (CDSS) was developed to assess the level of depression in schizophrenia (Addington et al., 1990), separate from positive, negative, and extrapyramidal symptoms associated with schizophrenia.

Nine items are rated (0 = absent, 1 = mild, 2 = moderate, 3 = severe) by the Investigator or a trained clinical rater based on a semi-structured clinical interview assessing a patient's depression, hopelessness, self deprecation, guilty ideas of reference, pathological guilt, morning depression, early wakening, suicidality, and observed depression over the previous two weeks.

The CDSS depression score is obtained by adding each of the item scores. A score above 6 has 82% specificity and 85% sensitivity for predicting the presence of a major depressive episode. The scale will be conducted at Days -1 (baseline) and Day 28.

### 6.3.18 PSP

The Personal and Social Performance scale (PSP) is a clinician-rated measure of the severity of personal and social dysfunction (Morosini et al., 2000) that has been validated in patients with acute schizophrenia (Patrick et al., 2009). The scale is a 100-point rating scale that is calculated based on ratings in four areas: 1) socially useful activities, 2) personal and social relationships, 3) self-care, and 4) disturbing and aggressive behavior.

### 6.3.19 C-SSRS

The C-SSRS is a questionnaire developed and validated by Kelly Posner and colleagues (2011) for the assessment of suicidal ideation and behavior. Several versions have been developed including the "Baseline" and "Screening" versions and a combined "Baseline/Screening" version of the scale which assesses suicidal ideation and behavior in a patient's lifetime and during a

predefined time period. This version can assess a patient's lifetime suicidality as well as eligibility based on inclusion/exclusion criteria. A separate "Since Last Visit" version of the scale has been developed which is used to assess suicidality since the patient's last visit. This version is meant to assess patients who have completed at least one initial C-SSRS assessment, and should be used in every subsequent visit. The 'Since Last Visit' version of the C-SSRS addresses any suicidal thoughts or behaviors the patient/participant may have had since the last time the C-SSRS was administered.

The C-SSRS will be administered by the Investigator or a trained clinical rater at Screening, Day -1 (baseline), and Days 8, 15, 28, 33 and 42. At Screening a potential patient will not be eligible if they report suicidal ideation of type 4 or 5 on the C-SSRS within 30 days prior to screening or any suicidal behavior in the last 2 years prior to screening, as indicated by any 'yes' answers on suicidal behavior section of C-SSRS.

### 6.3.20 BARS

The Barnes Akathisia Rating Scale (BARS) is a rating scale for drug-induced akathisia developed by Barnes (1989). It includes the rating of observable restless movements, the subjective awareness of restlessness, and the distress associated with the akathisia. There is also a global rating for severity. The scale is completed by the Investigator or a trained clinical rater after a standard examination. Objective Akathisia, Subjective Awareness and Subjective Distress are rated on a 4-point scale from 0 – 3, yielding a total score from 0 – 9. The Global Clinical Assessment of Akathisia is rated separately, on a 5-point scale from 0 – 4.

The BARS is to be conducted on Days -1 (baseline), and then on Days 8 and 28 between 3 and 6 hours after administration of study medication.

### 6.3.21 SAS

The Simpson-Angus Scale (SAS) is a measure of extrapyramidal side effects (Simpson and Angus, 1970). Ten items rating gait, arm dropping, shoulder shaking, elbow rigidity, wrist rigidity, leg pendulousness, head dropping, glabella tap, tremor, and salivation are rated on a scale from 0 (normal) to 4 (extreme in severity).

The SAS should be conducted by the Investigator or a trained clinical rater in a room where the subject can walk a sufficient distance to allow a natural pace (e.g., 15 paces). Each side of the body should be examined.

The SAS is to be conducted on Day -1 (baseline) and then on Days 8 and 28 between 3 and 6 hours after administration of study medication.

### 6.3.22 AIMS

The Abnormal Involuntary Movement Scale (AIMS; Guy, 1976) measures facial and oral movements, extremity movements, and trunk movements. Seven items are rated on a scale from none (0) to severe (4). A score of 'mild' (2) in two or more categories or a score of 'moderate' or 'severe' in any one category results in a positive AIMS score (i.e., the scores are not averaged). Additionally, overall severity is scored on the basis of severity of abnormal movements and incapacitation due to abnormal movements. The patient's awareness of and distress caused by the abnormal movements are also noted.

The AIMS is to be conducted by the Investigator or a trained clinical rater on Day -1 (baseline) and then on Days 8 and 28 between 3 and 6 hours after administration of study medication.

### 6.3.23 C-VISA© and Eligibility Adjudication

In conjunction with completing the BPRS, CGI-S, and SCID-CT, the Investigator or trained clinical rater will complete the Clinical Validation Inventory for Study Admission (C-VISA©) form. The form and the scales will be completed using the Clintara digital-audio pen recorder which will digitally record the completion of the forms and also create audio recordings of the interviews conducted with the patient in order to make the assessments. Both the forms and the audio recordings will be submitted to an Eligibility Adjudication Committee (EAC) at Clintara.

A two-tier review process will be employed by the EAC to confirm eligibility of screened patients. A Tier 1 independent reviewer will review all submitted information including audio recordings of interviews in order to independently confirm suitability of the patient. If the Tier 1 reviewer agrees with the Investigator's assessment of a potential patient then the reviewer will submit written confirmation of eligibility to the Investigator. If the Tier 1 reviewer disagrees with the Investigator's assessment of a patient and concludes that the patient is not suitable for the study a Tier 2 expert reviewer will be asked to review the data. In close communication with

the Investigator, the Tier 2 reviewer will make a final decision regarding eligibility of a patient. Once the final EAC decision is made, no waivers for exceptions will be provided.

Note: The Sponsor has no involvement in the adjudication process.

### 6.3.24 General and Dietary Restrictions

ITI-007 will be administered with a meal or light snack.

Strenuous activity is prohibited from signing ICF until discharge. After discharge, mild physical activity can be resumed, but strenuous physical activity is prohibited until End-of-Study.

All study procedures and meals will be performed at approximately the same time during the day for a given subject.

Cigarette smokers will be allowed to smoke in designated areas provided by the clinic and the number of cigarettes per day should be recorded. However, smoking is not permitted approximately 30 minutes prior to the clinical assessments, vital signs, ECGs.

Patients requiring special diets, such as vegetarian diets or diets without pork, will be allowed in the study as long as the patient agrees to adhere to the restrictions stated above, the clinic can accommodate such diet preparations, and that this is the normal day-to-day diet for that patient.

### 6.3.25 Blood Samples

Blood samples will be collected for determination of plasma concentrations of ITI-007 and three metabolites and for clinical laboratories. The sampling times relative to the time of dose administration for all blood sample collections are given in the *Study Events Schedule,* section 6.1. The estimates of blood volume and sample numbers are as follows:

| Day of study | Number of ITI-007 samples (7 mL per sample) | Clinical laboratory | Biomarker samples |
|---|---|---|---|
| Screening | | 21 | |
| Day -1 | 21 | | |
| Day 1 | 14 | 12.5 | 14 (protein & optional genetic) |
| Day 8 | 14 | 12.5 | |
| Day 28 | 21 | 12.5 | 7 (protein) |
| End-of-Study | 7 | 12.5 | |
| Total volume collected and analyzed per subject | 77 | 71 | 21 |
| Total number of samples | 11 | 5 | Up to 3 |

**Total volume collected per subject up to mL 169**

Blood samples for ITI-007 concentrations will be collected or by venipuncture from an appropriate vein into 7 mL tubes containing sodium heparin. The exact date and time of the pharmacokinetic blood collection must be recorded in the subject's case report form. Blood samples for clinical laboratory measurements will be collected into appropriate vacuum blood collection tubes.

Each label will state the study number, subject number, analyte (plasma), study day of sample, and scheduled time of sample. The actual date and time of blood collection will be recorded in the subject's case report form.

All blood samples for PK will be processed immediately and stored at -70° Celsius freezer. Additional information about the collection, handling, and storage of plasma for PK analysis will be provided separately prior to start of the study.

Blood samples for clinical laboratories should be collected and processed according to the standard operating procedures (SOPs) of the clinical site.

### 6.3.26 Bioanalytical Methods and Sample Shipment

Analysis of plasma samples for ITI-007 and its metabolites will be performed by [redacted]. The bioanalytical method will be a validated high performance liquid chromatography/tandem mass

spectrometric (LC/MS/MS) assay. The bioanalytical facility shall validate the method's reproducibility, lower and upper limit of quantitation, linearity, specificity, and recovery. Also, day-to-day performance of the methods shall be assessed by monitoring quality control samples and standard curve summaries.

All plasma samples will be shipped to [redacted] according to the instructions provided by [redacted].

# 7.0 STATISTICS

## 7.1 Data Analysis Plans

### 7.1.1 General Considerations

A formal and detailed statistical analysis plan (SAP) will be finalized before database lock.

Efficacy analyses will be conducted on the intent-to-treat (ITT) population. The ITT population will include all randomized subjects who received at least one dose of any study medication and who have a baseline measurement as well as at least one valid post baseline efficacy measurement. Subject data will be analyzed according to the treatment assigned at randomization, regardless of the treatment received during the course of the trial. Some of the sensitivity analyses of the efficacy variables will be performed on the per-protocol (PP) population, which will include all ITT subjects without any major protocol deviations.

Safety analyses will be conducted on the safety population which will include all randomized subjects who received at least one dose of study medication. Subjects will be analyzed according to the actual treatment received.

PK analyses will be performed on a subset of the ITT population, which will include all randomized subjects who received at least one dose of the study medication, have a baseline measurement as well as at least one valid post baseline efficacy measurement, and for whom a valid assay result (according to laboratory guidelines) has been obtained.

All treatment comparisons will be evaluated at a two-sided significance level of 0.05, unless otherwise specified. No adjustments for multiple comparisons will be made, unless otherwise stated.

Unless otherwise specified, baseline is defined as the last non-missing measurement at Visit 2 on Study Day -1 (or time-matched assessment when multiple assessments are scheduled on Study Day -1). Measurements on the same day of first administration of study medication (Visit 3) for which the time (hours and/or minutes) is missing, and are, according to the study schedule, supposed to be collected prior to treatment, will be considered baseline. If there are multiple measurements of the same variable on the same day, prior to the first administration of study

medication, the last value will be considered the baseline unless a time-matched baseline is appropriate, as defined in the SAP.

In case procedures and assessments are conducted during an unscheduled visit, due to early withdrawal, the collected data will be carried forward to the next scheduled visit when used in the repeated measures analyses. As a result, the time points in these analyses will match the scheduled pre-specified time points.

Safety and Efficacy data will be summarized and listed for all treated subjects by treatment group, site identifier (ID), patient ID, and time point (visit and relative study day), unless stated otherwise in the SAP. Continuous variables will be presented using the total number of subjects (N), number and percentage of subjects in the specified categories [n (%)], mean and standard deviation, median, minimum, and maximum. Categorical variables will be summarized using frequencies and percentages.

Changes made to the data analysis methods as described in the protocol will be documented in the SAP and will not necessitate a protocol amendment. All departures from the statistical analyses described in the approved protocol, whether made before or after unblinding, will be documented and justified in the final clinical study report. Additional exploratory analyses of the data will be conducted as deemed appropriate.

### 7.1.2 Pharmacokinetics

Blood samples will be collected pre-dose and 3-4 h post-dose on Study Days 1, 8, 28, at 6-7 h post-dose on Day 28, and once on Day 33 for determination of ITI-007 and metabolite concentrations. Samples will be collected from all subjects in order to maintain the blind. All samples will be analyzed from ITI-007 recipients; selected samples from placebo recipients will be analyzed for presence/absence of active drug to ensure appropriate randomization. Descriptive statistics for trough and peak ITI-007/metabolite levels will be calculated.

Subjects with partial data will be evaluated on a case-by-case basis to determine if sufficient data are available for meaningful analysis.

Values excluded from the pharmacokinetic analysis will be flagged with an asterisk and concentration values reported as below the level of quantification (BLQ) will be listed with the lower limit of quantification in parentheses. Descriptive statistics (N, arithmetic mean, SD, arithmetic CV%, median, minimum and maximum) will be used to summarize ITI-007 and metabolite concentration data at each planned sampling time point for each study part and treatment. BLQ values will be set to zero prior to calculation of descriptive statistics for the plasma concentration-time profile.

Plasma concentration data will be summarized by analyte, dose level, planned sampling time and day.

### 7.1.3 Patient Disposition

Summaries of patient disposition by treatment group, when applicable, and overall will be provided and will include the number of patients screened, the number and percentage of randomized subjects as well as the number and percentage of screening failures by reason; the number and percentage of subjects completed the 28-day study treatment, completed the study, or discontinued. A subject is defined to have completed the study if the subject has completed the End-of-Study procedures on Day 42. Subjects who discontinued from study treatment or from the study will be listed and summarized by treatment group, reason for discontinuation, and time.

Time to discontinuation (measured in days) due to AEs, lack of efficacy or due to any other reason of special interest, will be evaluated by each reason separately using Kaplan-Meier method. For each of these analyses, subjects who complete the Treatment Period, or who discontinue for a reason other than the one being evaluated will be considered as censored. In comparing time to discontinuation for any reason, subjects who complete the double-blind treatment phase will be considered as censored observations. The Log-rank Test will be used to compare the time to discontinuation between each ITI-007 treatment group and the placebo group.

### 7.1.4 Patient Characteristics

Demographic and baseline disease characteristics will be listed and summarized by treatment group. No formal statistical testing comparing treatment groups will be performed. The

following baseline will be summarized by treatment group: Continuous demographic characteristics: age, body weight, waist circumference, and body mass index; Categorical demographic characteristics: gender, race, and ethnicity group. Consumption habits (tobacco, alcohol and caffeine). Schizophrenia diagnosis type, time since first diagnosis of schizophrenia, presence of depression, and assessment of acute exacerbation: BPRS total score.

Baseline efficacy and safety variables will be listed and summarized by treatment group: PANSS total score, PANSS pro-social factor score, PANSS positive score, PANSS negative score, PANSS general psychopathology score, CGI-S score, PSP score, CDSS depression score, BARS (total score and Global Assessment of Akathisia), SAS, AIMS and C-SSRS.

### 7.1.5 Prior and Concomitant Therapy

Prior medications (defined as medications that start and stop prior to the date of first dose of study medication), and prior concomitant medication (defined as medications that start prior to and stop on or after the date of first dose of study medication), as well as concomitant medication (defined as medications that start on or after the date of first dose of study medication) will be listed summarized by treatment group.

The number and percent of subjects receiving lorazepam, benztropine, or propranolol during the treatment period and the total average dose of each of these drugs will be summarized for each treatment group.

### 7.1.6 Efficacy Analyses

The primary efficacy endpoint is the change from baseline to Day 28 on the PANSS total score. Secondary efficacy endpoints will include change from baseline to Days 8, 15, and 22 on the PANSS total score as well as change from baseline to Days 8, 15, 22, and 28 on the CGI-S, the PANSS Positive, Negative, and General Pathology sub-scales, the pro-social factor for the PANSS, PSP (Day 28 only), and CDSS (Day 28 only). Comparisons between each of the treatment groups (including both ITI-007 doses) versus placebo will be conducted for secondary measures, without adjustments made for multiple comparisons.

The treatment effect on the primary efficacy endpoint will be analyzed using a Mixed-effects Model for Repeated Measure (MMRM). The model may include treatment, site, Study Day, treatment-by-Study Day interaction, baseline PANSS score, and baseline PANSS score-by-Study Day interaction as fixed effects; the model will be detailed in the SAP. Within-patient errors will be modeled using an unstructured covariance matrix. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom, or similar procedure detailed in the SAP. If the model fails to converge, other covariance structures will be evaluated until model convergence is met.

A similar analysis will be conducted on the PP population.

In addition to the MMRM model, sensitivity analysis on the ITT and PP will be conducted and detailed in the SAP. Briefly, sensitivity analyses will be performed using the pattern mixture approach with multiple imputations (based on reasons for discontinuation, such as discontinuation for lack of efficacy or discontinuation due to adverse event) under the assumption of missing not at random.

Similar to the analysis of the primary efficacy endpoint, the MMRM will be used to evaluate the secondary endpoints.

Additionally, the PANSS total score will be used for a responder analysis will be conducted, calculating the percent of patients in each treatment group showing at least 20%, at least 30%, and at least 40% reduction from baseline in PANSS total score. The Fisher's Exact Test will be used to compare responder rate between each ITI-007 treatment group and the placebo group at each post baseline time point. The LOCF method will be applied to impute missing data.

Time to the first response (measured in days from the date of first administration of study drug to the date that the earliest 20%, 30%, and 40% or greater reduction is observed) will be analyzed using the Kaplan-Meier method. Subjects who do not show response will be considered as censored observation. The Log-rank Test will be used to compare the time to the first response between each ITI-007 treatment group and the placebo group.

The number and percent of subjects receiving lorazepam, benztropine, or propranolol and the total average dose of each of these drugs will be summarized for each treatment group.

The primary endpoint will be statistically powered for a two-tailed (alpha level 0.05) comparison of 60 mg ITI-007 to placebo. Only if 60 mg ITI-007 is statistically significantly different from placebo will the 40 mg ITI-007 treatment group be compared to placebo on the primary endpoint. This hierarchical testing will preserve the alpha level at 0.05 for the two-tailed comparison of 40 mg ITI-007 to placebo. No other adjustments for multiple comparisons are planned.

### 7.1.6 Safety Analyses

All available safety data from subjects receiving at least one dose of study medication will be included in the safety analyses. The frequency of adverse events will be tabulated. Baseline, end-of-study, and change from baseline laboratory, 3-positional vital signs, and ECG parameters will be summarized: whenever appropriate, out-of-range values will be flagged on data listings and tabulated. Shift tables will be prepared for laboratory parameters. Parameters collected in duplicate or triplicate will be analyzed as an average of the measures for the relevant time point, including baseline. A formal statistical comparison of each dose of ITI-007 to placebo will be made for the relative risk (95% confidence interval and p-value) of 'any' treatment-emergent adverse event and for each individual treatment-emergent adverse event recorded during the study.

## 7.2 Determination of Sample Size

Approximately 440 subjects will be randomized in a 1:1:1 ratio to 1 of 3 treatment arms: ITI-007 60 mg QD, ITI-007 40 mg QD, or placebo QD, which will provide 396 evaluable subjects, assuming a 10% early discontinuation rate before a post-dose assessment is made. With a sample size of approximately 132 in each of the treatment arms, the study is designed to have 90% power to demonstrate an effect size of 0.4, corresponding, for example, to a 6 point difference in change from baseline of total PANSS score between a treatment arm and placebo, at a one-sided significance level of 0.025 (two-sided significance level of 0.05). The sample size was calculated using a two-sample t-test, assuming the standard deviation of the change from baseline in the PANSS total score is about 15.

# 8.0 ADVERSE EVENTS

## 8.1 Adverse Event Definitions

### 8.1.1 Adverse event

An AE is any untoward medical occurrence in a study subject administered an IMP, whether or not considered drug related. This can be any unfavorable and unintended sign (such as an abnormal laboratory finding), symptom or disease temporally associated with the use of an IMP, without any judgment of causality.

The AE may be:

- A new illness
- A worsening sign or symptom the condition under treatment, or of a concomitant illness;
- An effect of the study medication, including comparator; or
- A combination of 2 or more of these factors.

No causal relationship with the study medication or with the clinical study itself is implied by the use of the term AE.

Surgical procedures themselves are not AEs; they are therapeutic measures for conditions that require surgery. The condition for which the surgery is required may be an AE. Planned surgical measures, if permitted by the clinical study protocol and the conditions leading to those measures are not AEs.

All AEs fall into the categories of "nonserious" or "serious. (see *Section 8.1.2 Serious adverse event*).

### 8.1.2 Serious adverse event

A serious adverse event (SAE) is defined as any untoward medical occurrence that at any dose:

- Results in death;

- Is life-threatening; this means that the subject was at risk of death at the time of the event; it does not mean that the event hypothetically might have caused death if it were more severe;
- Requires hospitalization or prolongation in existing hospitalization;
- Results in persistent or significant disability or incapacity;
- Is a congenital anomaly or birth defect;

The term "life-threatening" refers to an event in which the subject is at risk of death at the time of the event; it does not refer to an event, which hypothetically might have caused death if it were more severe.

Medical and scientific judgment should be exercised in deciding whether an AE is serious. Some important medical events may not result in death, be life -threatening or require hospitalization may be considered SAEs when they may jeopardize the subject such that medical or surgical intervention is needed to prevent one of the outcomes listed above.

Examples of such medical events include intensive emergency treatment for an allergic reaction, bloody dyscrasias or convulsions that do not results in in-patient hospitalization, and the development of drug dependency or drug abuse.

The study itself requires an inpatient stay in the clinic. Therefore, prolongation of the inpatient portion of the study may not be considered serious if the prolongation is due to events that are otherwise not considered serious or would not otherwise be sufficient to warrant hospitalization.

If either the Sponsor or PI believes that any event is serious, the event must be considered and evaluated by the Sponsor for possible expedited reporting.

Clarification of the difference in meaning between "severe" and "serious":

The term "severe" is often used to describe the intensity (severity) of a specific event (as in mild, moderate, or severe myocardial infarction); the event itself, however, may be of relatively minor medical significance (such as severe headache). This is not the same as "serious", which is based on patient/event outcome or action criteria usually associated with events that pose a threat to a patient's life or functioning. Seriousness (not severity) serves as a guide for defining regulatory reporting obligations.

### 8.1.3 Nonserious adverse event

A **nonserious adverse event** is any adverse event not meeting the serious adverse event criteria.

### 8.1.4 Definition of relationship to study medication

By definition, any AE that starts before the first dose of IMP is administered is considered to be "unrelated". For example, in the present study schizophrenia, an exacerbation of schizophrenia or worsening of schizophrenia is considered to start before the first dose of IMP by inclusion criteria and is considered to be "unrelated" to study medication.

The PI will assess the causality/relationship between the IMP and the AE. One of the following categories should be selected based on medical judgment, considering the definitions below and all contributing factors.

| | |
|---|---|
| Definitely related | A clinical event, including laboratory test abnormality, occurs in a plausible time relationship to treatment administration, and which concurrent disease or other drugs or chemicals cannot explain. The response to withdrawal of the treatment (dechallenge*) should be clinically plausible. The event must be definitive pharmacologically or phenomenologically, using a satisfactory rechallenge† procedure if necessary. |
| Probably related | A clinical event, including laboratory test abnormality, with a reasonable time sequence to administration of the treatment, unlikely to be attributed to concurrent disease or other drugs or chemicals, and which follows a clinically reasonable response on withdrawal (dechallenge). Rechallenge information is not required to fulfill this definition. |
| Possibly related | A clinical event, including laboratory test abnormality, with a reasonable time sequence to administration of the treatment, but which could also be explained by concurrent disease or other drugs or chemicals. Information on treatment withdrawal may be lacking or unclear. |
| Unlikely to be related | A clinical event, including laboratory test abnormality, with a temporal relationship to treatment administration which makes a causal relationship improbable, and in which other drugs, chemicals or underlying disease provide plausible explanations. |
| Unrelated | A clinical event, including laboratory test abnormality, with little or no temporal relationship with treatment administration. May have negative dechallenge and rechallenge information. Typically explained by extraneous factors (e.g., concomitant disease, environmental factors or other drugs or chemicals) |

* Dechallenge is when a drug suspected of causing an AE is discontinued. If the symptoms of the AE disappear partially or completely, within a reasonable time from drug discontinuation, this is termed a positive dechallenge. If the symptoms continue despite withdrawal of the drug, this is termed a negative dechallenge. Note that there are exceptions when an AE does not disappear upon discontinuation of the drug, yet drug-relatedness clearly exists (for example, as in bone marrow suppression, fixed drug eruptions, or tardive dyskinesia).

† Rechallenge is when a drug suspected of causing an AE in a specific subject in the past is readministered to that subject. If the AE recurs upon exposure, this is termed a positive rechallenge. If the AE does not recur, this is termed a negative rechallenge.

### 8.1.5 Definition of intensity

The PI will assess all AEs for severity in accordance with the following standard ratings.

| | |
|---|---|
| Mild | Ordinarily transient symptoms, does not influence performance of subject's daily activities. Treatment is not ordinarily indicated. |
| Moderate | Marked symptoms, sufficient to make the subject uncomfortable. Moderate influence on performance of subject's daily activities. Treatment may be necessary. |
| Severe | Symptoms cause considerable discomfort. Substantial influence on subject's daily activities. May be unable to continue in the study and treatment may be necessary. |
| Life-threatening | Extreme limitation in activity, significant assistance required; significant medical intervention/therapy required, hospitalization or hospice care probable. |

When changes in the intensity of an AE occur more frequently than once a day, the maximum intensity for the event should be noted for that day. Any change in severity of signs and symptoms over a number of days will be captured by recording a new AE, with the amended severity grade and the date (and time, if known) of the change.

## 8.2 Period of observation

For the purposes of this study, the period of observation extends from the time the subject gives his study specific informed consent until the End of Study procedures are completed.

If the Investigator detects a serious adverse event in a study subject after the end of the period of observation, and considers the event possibly related to prior study treatment, he or she should contact the Sponsor to determine how the adverse event should be documented and reported.

## 8.3 Documentation and Reporting of Adverse Events by Investigator

Adverse events should be collected and recorded for each subject from the date the informed consent form (ICF) was signed until the end of their participation in the study, i.e. the subject has discontinued or completed the study.

Adverse events may be volunteered spontaneously by the study subject, or discovered by the study staff during physical examinations or by asking an open, non-leading question such as 'How have you been feeling since you were last asked?' All AEs and any required remedial action will be recorded. The nature of AE, date (and time, if known) of AE onset, date (and time, if known) of AE outcome to date, severity, and action taken for the AE will be documented together with the

PI's assessment of the seriousness of the AE and causal relationship to IMP and/or study procedure.

All AEs should be recorded individually in the study subject's own words (verbatim) unless, in the opinion of the PI, the AEs constitute components of a recognized condition, disease or syndrome. In the latter case, the condition, disease or syndrome should be named rather than each individual symptom. The AEs will subsequently be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

## 8.4 Notification about Serious or Unexpected Adverse Events

The PI will review each SAE and evaluate the intensity and the causal relationship of the event to IMP. All SAEs will be recorded from signing of informed consent until Follow-up. Serious AEs occurring after the Follow-up visit and coming to the attention of the PI must be reported only if there is (in the opinion of the PI) reasonable causal relationship with the IMP.

The PI is responsible for providing notification to the Sponsor of any SAE, whether deemed IMP-related or not, that a subject experiences during their participation in study within 24 hours of becoming aware of the event.

As a minimum requirement, the initial notification should provide the following information:

- Study number
- Subject/Patient number
- Gender
- Date of birth
- Name of PI and full clinical site address
- Details of SAE
- Criterion for classification as 'serious'
- IMP name, or code if unblinded, and treatment start date
- Date of SAE onset
- Causality assessment (if sufficient information is available to make this classification)

The Sponsor will request clarification of omitted or discrepant information from the initial notification. The PI or an authorized delegate is responsible for faxing the requested information to the Sponsor within 24 hours of the Sponsor's request.

Initial reports of SAEs must be followed later with detailed descriptions, including clear photocopies of other documents as necessary (e.g. hospital reports, consultant reports, autopsy reports etc.), with the study subject's personal identifiers removed. All relevant information obtained by the PI through review of these documents will be recorded and faxed to the Sponsor within 24 hours of receipt of the information. If a new SAE Report Form is faxed, then the PI must sign and date the form. The Sponsor may also request additional information on the SAE, which the PI or an authorized delegate must fax to the Sponsor within 24 hours of the request.

**SERIOUS ADVERSE EVENT REPORTING INSTRUCTIONS**

Safety Contact Information:

Intra-Cellular Therapies, Inc.

# 9.0 EMERGENCY PROCEDURES

During this inpatient study, patients will not be allowed to leave the inpatient unit at anytime. The clinic staff is present on the site at all times. In addition, the Investigator or designated physician is reachable by telephone 24 hours a day, 7 days week.

Any control or follow-up examinations can be initiated at any time by the Investigator or the emergency physician. In case of any clinically significant abnormalities, medical assistance is assured until complete resolution or stabilization. The Investigator may discontinue the treatment at any time based on clinical judgment.

In case of emergency, the patient will be treated as described in the clinical site's standard operating procedures. The emergency care is under the responsibility of the emergency physician. First aid and resuscitation can be provided on site. If acute dystonic reactions are seen following administration of ITI-007, standard clinical intervention should be used until this reaction subsides. This intervention may include the administration of anticholinergic therapy.

## 9.1 Emergency Sponsor Contact

In emergency situations, the Investigator should contact the Sponsor by telephone at the number listed on the title page of the protocol.

## 9.2 Emergency Identification of Study Medication

Unblinding of treatment assignment during the study is discouraged and should occur only if it is absolutely necessary to know what the patient received. If the Sponsor or Investigator deems identification of the study drug as necessary for the purpose of providing urgent patient care, and knowledge of the patient's treatment assignment (ITI-007 or placebo) will alter subsequent care, the randomization envelope for that single patient will be opened. The date and reason for the unblinding must be recorded on the patient's eCRF. When possible, the Sponsor Study Manager should be notified prior to unblinding; otherwise, the study manager and site monitor must be notified within 24 hours after unblinding.

## 9.3 Emergency Treatment

During and following a subject's participation in the trial, the Investigator should ensure that adequate medical care is provided to a subject for any adverse events, including clinically significant laboratory values, related to the trial. The Investigator should inform a subject when medical care is needed for intercurrent illness(es) of which the Investigator become aware.

# 10.0 ETHICS

## 10.1 Institutional Review Board or Independent Ethics Committee

Before initiation of the study, the Sponsor (see below) will submit the study protocol, sample Informed Consent Form, and any other documents that pertain to subject information to the Institutional Review Board (IRB)/Independent Ethics Committee (IEC). The Sponsor must also submit any other information that may be requested to the IRB/IEC for review and approval. The Sponsor will request that the IRB/IEC provide written approval of the study and will keep on file records of approval of all documents pertaining to this study. A letter confirming the approval must be forwarded to the clinical monitor prior to initiation of this study. This letter will be forwarded to Intra-Cellular Therapies, Inc. before the initiation of the study.

The Sponsor must submit and, where necessary, obtain approval from the IRB/IEC for all subsequent protocol amendments and changes to the informed consent document.

The Sponsor will be responsible for obtaining annual IRB/IEC approval or renewal throughout the duration of the study. Copies of the Investigator's reports and the IRB/IEC continuance of approval must be sent to the Sponsor.

## 10.2 Ethical Conduct of the Study

The procedures set out in this study protocol, pertaining to the conduct, evaluation, and documentation of this study, are designed to ensure that the Sponsor and Investigator abide by the principles of the good clinical practice (GCP) guidelines and the ethical principles laid down in the current revision of the Declaration of Helsinki (see Appendix A). The study will also be carried out in keeping with local legal and regulatory requirements in the United States.

The Investigator should ensure that all persons assisting with the trial are adequately informed about the protocol, any amendments to the protocol, the study treatments, and their trial-related duties and functions.

## 10.3 Subject Information and Informed Consent

Informed consent will be obtained from each subject in writing before the start of the study. The nature of the informed consent will comply with the current version of the Declaration of Helsinki, the current requirements of Good Clinical Practice (according to ICH) and local regulation which ever affords the greater subject protection.

## 10.4 Indemnity and Compensation

The Sponsor will adhere to local regulations and guidelines regarding clinical trial compensation of patients who health is adversely affected by taking part in the study. The Sponsor carries Product Liability and Clinical Trial insurance in compliance with current law in the United States.

## 10.5 Incentives

Patients will be compensated for participation in the study. Details of the compensation are set forth in the ICF. This compensation may be reduced in the case of failure of patients to comply with treatment, as outlined in Sections 4.6.

## 10.5 Protocol Amendments

Neither the Investigator nor the Sponsor will alter this study protocol without obtaining the written agreement of the other. Once the study has started, amendments should be made only in exceptional cases.

# 11.0 STUDY ADMINISTRATION

## 11.1 Clinical Monitoring

Monitoring procedures will be followed in order to comply with GCP guidelines. Auditing procedures, if applicable, will be followed in order to comply with GCP guidelines. On-site checking of the case report forms for completeness and clarity, crosschecking with source documents, and clarification of administrative matters will be performed.

ITI will select a Clinical Research Associate (CRA) to monitor the study. This clinical monitor will arrange regular visits to the investigational unit to check the progress of the study and to ensure that it is conducted in accordance with the protocol, Good Clinical Practice and appropriate regulatory local requirements. During monitoring visits the clinical monitor will:

review Case Report Forms (CRFs) for omission of data, errors and compliance,

perform source data verification where appropriate,

check subject informed consent forms,

ensure that study materials are correctly stored and dispensed,

check that the Investigator obligations are adhered to, and

help resolve any problems.

Regulatory authorities, the IEC/IRB, and/or quality groups from the Sponsor or clinical site may request access to all source documents, case report forms, and other study documentation for on-site audit or inspection. Direct access to these documents must be guaranteed by the Investigator, who must provide support at all times for these activities.

Supplies will be retained at this site until an authorization to destroy is received from the study Sponsor.

## 11.2 Data Quality Assurance

As a final step in the data management process, a 100% audit will be performed on the key safety parameters. In addition, a random subject sample (10%) will be selected in order to perform a partial audit of remaining data. The purpose of this audit is to detect systematic and random errors. Any errors found from this audit will be corrected and, if the error rate is greater than 50 errors per 10,000 fields (0.5%), a systematic review of the database will be undertaken, and corrections will be made to achieve an error rate of <0.5%.

## 11.3 Retention of Study Records

The Investigator and co-Investigator for a minimum of 15 years must retain the following records after completion or termination of the study:

- signed informed consent documents for all subjects,
- subject identification code list, screening log (if applicable), and enrollment log,
- composition of the IEC/IRB or other applicable statement,
- record of all communications between the Investigator and Sponsor (or Clinical Research Organization [CRO]),
- list of subinvestigators and other appropriately qualified persons to whom the Investigator has delegated significant trial-related duties, together with their roles in the study and their signatures,
- copies of case report forms and of documentation of corrections for all subjects,
- drug accountability records,
- record of any body fluids or tissue samples retained,
- all other source documents (patient records, hospital records, laboratory records, etc.),
- all other documents as listed in *Section 8* of the ICH consolidated guideline on Good Clinical Practices (GCP), *Essential Documents for the Conduct of a Clinical Trial*.

However, because of international regulatory requirements, the Sponsor may request retention for a longer period of time. The Investigator must therefore obtain approval in writing from the Sponsor prior to destruction of any records.

Normally, these records will be held in the Investigator's archives. If the Investigator is unable to meet this obligation, he or she must ask the Sponsor for permission to make alternative arrangements. Details of these arrangements should be documented.

## 11.4 Confidentiality

Subject names will not be supplied to the Sponsor. Only the subject number and subject initials will be recorded in the case report form, and if the subject name appears on any other document (e.g., pathologist report), it must be obliterated before a copy of the document is supplied to the Sponsor. Study findings stored on a computer will be stored in accordance with local data protection laws. The subjects will be told that representatives of the Sponsor, IEC/IRB, or regulatory authorities may inspect their medical records to verify the information collected, and that all personal information made available for inspection will be handled in strictest confidence and in accordance with local data protection laws.

The Investigator will maintain a personal subject identification list (subject numbers with the corresponding subject names) to enable records to be identified.

## 11.5 Documentation of Study Results

Source data are all the information in original records and certified copies of original records of clinical findings, observations, or other activities in the study, which are necessary for the reconstruction and evaluation of the study. The Investigator will permit study-related monitoring, audits, IRB/IEC reviews and regulatory inspections, with a direct access to all the required source records.

For each subject randomized, the Investigator will enter at least the following information on patient’s source records: subject’s participation in this study, concomitant therapies, clinically significant adverse events and a statement at a subject’s end of participation.

An electronic case report form (eCRF for each patient will be accessible via an Electronic Data capture (EDC) system.

The Investigator, or designated representative, must enter all protocol-required information collected during the study in the eCRF. In some cases, the source documents are the physician's subject chart. In these cases data collected in the eCRF must match the data in those charts. In other cases, the eCRF, or part of the eCRF, may also serve as source documents. In these cases, a document should be available at the Investigator's site as well as at Sponsor (or designee) that clearly identifies those data that will be recorded in the eCRF, and for which the eCRF will stand as the source document. Some data collected from third party vendors (e.g., ECG and laboratory values) will not be imported into the eCRF and will be available as separate datasets.

Details of eCRF completion and correction will be explained to the Investigator and other persons (if applicable). If the Investigator authorizes other persons to make entries in the eCRF, the names, positions, signatures, and initials of these persons must be supplied to the Sponsor (or designee) and/or the CRA.

The Investigator, or designated representative, should complete the eCRF as soon as possible after information is collected. An explanation should be given for all missing data. Deviations of procedures from the scheduled protocol time must be documented and an explanation should be provided. The completed eCRF must be reviewed and electronically signed by the Investigator. The CRO will provide the Investigator with an electronic copy of all completed eCRFs to the site upon completion of the study.

## 11.6 Annual Progress Reports

The Sponsor will submit a summary of the progress of the trial to the IND on file with the FDA once a year. In addition, the Investigator is responsible for submitting progress reports at least annually to the IRB that approved the study. Information will be provided on the date of inclusion of the first subject, numbers of subjects included, and numbers of subjects that have completed the trial, serious adverse events, other problems and amendments.

## 11.7 End of Study Report

The Investigator will notify their IRB of the end of the study and will submit a final report using the forms and/or format specified. The end of the study is defined as the last patient's last visit.

In case the study is ended prematurely, the Investigator will notify the IRB, including the reasons for the premature termination.

Within one year after the end of the study, the Investigator/Sponsor will submit a final study report with the results of the study, including any publications/abstract of the study, to the FDA

## 11.8 Use of Study Results

All information concerning the product as well as any matter concerning the operation of ITI-007, such as clinical indications for the drug, its formula, methods of manufacture and other scientific data relating to it, that have been provided by Intra-Cellular Therapies, Inc., and are unpublished, are confidential and must remain the sole property of Intra-Cellular Therapies, Inc or ITI. The Investigator and sites' personnel will agree to use the information only for the purposes of carrying out this study and for no other purpose unless prior written permission from Intra-Cellular Therapies, Inc. is obtained. The Sponsor has full ownership of the case report forms completed as part of the study.

By signing the study protocol, the Investigator agree that Intra-Cellular Therapies, Inc. may use the results of the study for the purposes of national and international registration, publication, and information for medical and pharmaceutical professionals. If necessary, the authorities will be notified of the Investigator's name, address, qualifications, and extent of involvement.

# 13.0 References

Addington D, Addington J, Schissel B. (1990) A depression rating scale for schizophrenics. Schizophr Res. 3(4):247-51.

Barnes TR. (1989) A rating scale for drug-induced akathisia. Br J Psychiatry. 154:672-6.

Bushe C, Shaw M, Peveler RC. (2008) A review of the association between antipsychotic use and hyperprolactinaemia. J Psychopharmacol. 22(2 Suppl):46-55.

Farde L, Nyberg S, Oxenstierna G, Nakashima Y, Halldin C, Ericsson B. (1995) Positron emission tomography studies on D2 and 5-HT2 receptor binding in risperidone-treated schizophrenic patients. J Clin Psychopharmacol. 15(1 Suppl 1):19S-23S.

First MB, Williams JBW, Karg RS, Spitzer RL: Structured Clinical Interview for DSM-5 Disorders – Clinical Trials Version (SCID-5-CT, Modified for ITI-007-301 Sept 2014), Arlington, VA: American Psychiatric Association.

Guy W, editor. (1976) ECDEU Assessment Manual for Psychopharmacology. Rockville, MD, U.S. Department of Health Education, and Welfare.

Kay SR, Fiszbein A, Opler LA. (1991) The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull. 13(2):261-76.

Meltzer H. (1989) Clinical studies on the mechanism of action of clozapine: the dopamine-serotonin hypothesis of schizophrenia. Psychopharmacology (Berl). 99 Suppl:S18-27. Review.

Morosini P-L, Magliano L, Brambilla L, Ugolini S, Pioli R (2000) Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social functioning. Acta Psychiatr Scand 101:323-329.

Nordström AL, Farde L, Nyberg S, Karlsson P, Halldin C, Sedvall G. (1995) D1, D2, and 5-HT2 receptor occupancy in relation to clozapine serum concentration: a PET study of schizophrenic patients. Am J Psychiatry. 52(10):1444-9.

Patrick DL, Burns T, Morosini P, Rothman M, Gagnon DD, Wild D, Adriaenssen I (2009) Reliability, validity and ability to detect change of the clinician-rated Personal and Social Performance scale in patients with acute symptoms of schizophrenia. Curr Med Res Opin 25:325-338.

Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, Currier GW, Melvin G, Greenhill L, Shen S, & Mann JJ (2011) The Columbia-Suicide Severity Rating Scale: Initial Validity and Internal Consistency Findings From Three Multisite Studies With Adolescents and Adults. American Journal of Psychiatry 168:1266-1277.Rummel-Kluge C, Komossa K, Schwarz S, Hunger H, Schmid F, Kissling W,

Davis JM, Leucht S (2012) Second-Generation Antipsychotic Drugs and Extrapyramidal Side Effects: A Systematic Review and Meta-analysis of Head-to-Head Comparisons. Schizophrenia Bulletin 38:167-177.

Simpson GM and Angus JWS (1970). A rating scale for extrapyramidal side effects. *Acta Psychiat Scand*. 212:11–19.

Svenningsson P and Greengard P (2007) p11 (S100A10) — an inducible adaptor protein that modulates neuronal functions. *Curr Op Pharmacol* 7:27-32.

# APPENDICES

## Appendix A: Declaration of Helsinki

**Introduction**

It is the mission of the physician to safeguard the health of the people. His or her knowledge and conscience are dedicated to the fulfillment of this mission.

The Declaration of Geneva of the World Medical Association binds the physician with the words, “The health of my subject will be my first consideration", and the International Code of Medical Ethics declares, "A physician shall act only in the subject's interest when providing medical care which might have the effect of weakening the physical and mental condition of the subject".

The purpose of biomedical research involving human subjects must be to improve diagnostic, therapeutic and prophylactic procedures and the understanding of the aetiology and pathogenesis of disease.

In current medical practice most diagnostic, therapeutic or prophylactic procedures involve hazards. This applies especially to biomedical research.

Medical progress is based on research that ultimately must rest in part on experimentation involving human subjects.

In the field of biomedical research a fundamental distinction must be recognized between medical research in which the aim is essentially diagnostic or therapeutic for a subject, and medical research, the essential object of which is purely scientific and without implying direct diagnostic or therapeutic value to the person subjected to the research.

Special caution must be exercised in the conduct of research that may affect the environment, and the welfare of animals used for research must be respected.

Because it is essential that the results of laboratory experiments be applied to human beings to further scientific knowledge and to help suffering humanity, the World Medical Association has prepared the following recommendations as a guide to every physician in biomedical research involving human subjects. They should be kept under review in the

future. It must be stressed that the standards as drafted are only a guide to physicians all over the world. Physicians are not relieved from criminal, civil and ethical responsibilities under the laws of their own countries.

## I. Basic principles

1. Biomedical research involving human subjects must conform to generally accepted scientific principles and should be based on adequately performed laboratory and animal experimentation and on a thorough knowledge of the scientific literature.

2. The design and performance of each experimental procedure involving human subjects should be clearly formulated in an experimental protocol which should be transmitted for consideration, comment and guidance to a specially appointed committee independent of the Investigator and the Sponsor, provided that this independent committee is in conformity with the laws and regulations of the country in which the research experiment is performed.

3. Biomedical research involving human subjects should be conducted only by scientifically qualified persons and under the supervision of a clinically competent medical person. The responsibility for the human subject must always rest with a medically qualified person and never rest on the subject of the research, even though the subject has given his or her consent.

4. Biomedical research involving human subjects cannot legitimately be carried out unless the importance of the objective is in proportion to the inherent risk to the subject.

5. Every biomedical research project involving human subjects should be preceded by careful assessment of predictable risks in comparison with foreseeable benefits to the subject or to others. Concern for the interests of the subject must always prevail over the interests of science and society.

6. The right of the research subject to safeguard his or her integrity must always be respected. Every precaution should be taken to respect the privacy of the subject and to minimize the impact of the study on the subject's physical and mental integrity and on the personality of the subjects.

7. Physicians should abstain from engaging in research projects involving human subjects unless they are satisfied that the hazards involved are believed to be predictable. Physicians should cease any investigation if the hazards are found to outweigh the potential benefits.

8. In publication of the results of his or her research, the physician is obliged to preserve the accuracy of the results. Reports on experimentation not in accordance with the principles laid down in this Declaration should not be accepted for publication.

9. In any research on human beings, each potential subject must be adequately informed of the aims, methods, anticipated benefits and potential hazards of the study and the discomfort it may entail. He or she should be informed that he or she is at liberty to abstain from participation in the study and that he or she is free to withdraw his or her consent to participation at any time. The physician should then obtain the subject's freely given informed consent, preferably in writing.

10. When obtaining informed consent for the research project the physician should be particularly cautious if the subject is in a dependent relationship to him or her or may consent under duress. In that case the informed consent should be obtained by a physician who is not engaged in the investigation and who is completely independent of this official relationship.

11. In the case of legal incompetence, informed consent should be obtained from the legal guardian in accordance with national legislation. Where physical or mental incapacity makes it impossible to obtain informed consent, or when the subject is a minor, permission from the responsible relative replaces that of the subject in accordance with national legislation.

Whenever the minor child is in fact able to give consent, the minor’s consent must be obtained in addition to the consent of the minor’s legal guardian.

12. The research protocol should always contain a statement of the ethical considerations involved and should indicate that the principles enunciated in the present Declaration are complied with.

## II. Medical research combined with professional care (clinical research)

1. In the treatment of the sick person, the physician must be free to use a new diagnostic and therapeutic measure, if in his or her judgment it offers hope of saving life, reestablishing health or alleviating suffering.

2. The potential benefits, hazards and discomfort of a new method should be weighed against the advantages of the best current diagnostic and therapeutic methods.

3. In any medical study, every subject - including those of a control group, if any should be assured of the best proven diagnostic and therapeutic method.

4. The refusal of the subject to participate in a study must never interfere with the physician-subject relationship.

5. If the physician considers it essential not to obtain informed consent, the specific reasons for this proposal should be stated in the experimental protocol for transmission to the independent committee.

6. The physician can combine medical research with professional care, the objective being the acquisition of new medical knowledge, only to the extent that medical research is justified by its potential diagnostic or therapeutic value for the subject.

## III. Non therapeutic biomedical research involving human subjects (non clinical biomedical research)

- In the purely scientific application of medical research carried out on a human being, it is the duty of the physician to remain the protector of the life and health of that person on whom biomedical research is carried out.

- The subjects should be volunteers - either healthy persons or subjects for whom the experimental design is not related to the subject's illness.

- The Investigator or the investigating team should discontinue the research if in his, her or their judgment it may, if continued, be harmful to the individual.

- In research on man, the interest of science and society should never take precedence over considerations related to the well-being of the subject.

# Appendix B: SAFETY PARAMETERS

**Vitals signs**

Systolic and diastolic blood pressure, and heart rate measures are performed. Orthostatic blood pressure and pulse will be measured after at least 10 min rest in the supine position, 3 min after sitting position and 1 min after standing.

- Systolic blood pressure (SBP): 95 ≤ SBP ≤ 140 mm Hg.

- Diastolic blood pressure (DBP): 50 ≤ DBP ≤ 90 mm Hg.

- HR: 50 ≤ HR ≤ 100 bpm.

- Diagnosis of orthostatic hypotension: Orthostatic hypotension is a reduction of SBP of at least 20 mm Hg or DBP of at least 10 mm HG within 3 min of standing ("Consensus statement on the definition of orthostatic hypotension, pure autonomic failure, and multiple system atrophy"; Neurology 1996;46:1470). Orthostasis will be defined comparing standing measures to supine measures.

**ECG recordings**

An international standardized 12-lead ECG is performed after at least a 10 min rest (sitting), using an automatic recorder. Traces are recorded in the case report form, with comments in case of abnormality as well as:

- Rhythm (sinusal -other)

- Heart rate (50 ≤ N ≤ 100 bmp)

- PQ interval (120 ≤ N ≤ 210 msec)

- QRS duration (70 ≤ N ≤ 120 msec)

- QTc B (N ≤ 450 msec)

- QTc F (N ≤ 450 msec)

- QT < 500 ms

Clinical laboratory parameters should be provided to the Sponsor and trial master file by PPD Central laboratory.

# Appendix C: Additional risks from associated procedures

***Blood Volumes***

A maximum of approximately 169 mL of blood will be taken during the entire study. Taking blood may cause pain, bleeding or bruising where the needle enters the body; in rare cases, it may result in fainting.

# Appendix D: List of Example Inhibitors and Inducers of CYP450 3A4

## Strong Inhibitors (≥5-fold increase in AUC or >80% decrease in CL):

Boceprevir
Clarithromycin
Conivaptan
Grapefruit juice
Indinavir
Itraconazole
Ketoconazole
Lopinavir/ritonavir
Mibefradil
Nefazodone
Nelfinavir
Posaconazole
Ritonavir
Saquinavir
Telaprevir
Telithromycin
Voriconazole

## Strong Inducers (≥80% decrease in AUC):

Avasimibe
Carbamazepine
Phenytoin
Rifampin
St. John's wort

**Reference:**
**http://www.fda.gov/drugs/developmentapprovalprocess/developmentresources/druginteractionslabeling/ucm093664.htm**

# STATISTICAL ANALYSIS PLAN

## ITI-007-301

## A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTI-CENTER STUDY TO ASSESS THE ANTIPSYCHOTIC EFFICACY OF ITI-007 IN PATIENTS WITH SCHIZOPHRENIA

**AUTHOR:** [redacted]

**VERSION NUMBER AND DATE:** V1.5, 20JUL2015

## STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan V1.5 (Dated 20Jul2015) for Protocol ITI-007-301.

| | Name | Signature | Date |
|---|---|---|---|
| Author: | [redacted] | | |
| Position: | | | |
| Company: | | | |

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan, authorizing that the content is acceptable for the reporting of this study.

| | Name | Signature | Date |
|---|---|---|---|
| Approved By: | [redacted] | | |
| Position: | | | |
| Company: | | | |

| | |
|---|---|
| Approved By: | [redacted] |
| Position: | |
| Company: | |

| | |
|---|---|
| Approved By: | [redacted] |
| Position: | |
| Company: | |

| | |
|---|---|
| Approved By: | [redacted] |
| Position: | |
| Company: | |

## STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan V1.5 (Dated 20Jul2015) for Protocol ITI-007-301.

| | Name | Signature | Date |
|---|---|---|---|
| **Author:** | [redacted] | | |
| **Position:** | | | |
| **Company:** | | | |

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan, authorizing that the content is acceptable for the reporting of this study.

| | Name | Signature | Date |
|---|---|---|---|
| **Approved By:** | [redacted] | | |
| **Position:** | | | |
| **Company:** | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |

# MODIFICATION HISTORY

| Unique Identifier for this Version | Date of the Document Version | Author | Significant Changes from Previous Authorized Version |
|---|---|---|---|
| 1.0 | 08Jan2015 | [redacted] | Not Applicable – First Version |
| 1.1 | 25Mar2015 | | Updated for client comments |
| 1.2 | 20Apr2015 | | Updated for client comments |
| 1.3 | 01May2015 | | Updated for client comments |
| 1.4 | 02Jul2015 | | Updated for consistency of shells and multiplicity adjustment |
| 1.5 | 20Jul2015 | | Update for client comments |

# TABLE OF CONTENTS

# 1. LIST OF ABBREVIATIONS

| | |
|---|---|
| AE | Adverse Event |
| AIMS | Abnormal Involuntary Movement Scale |
| ANCOVA | Analysis of Covariance |
| AR(1) | Autoregressive(1) |
| ARH(1) | Heterogeneous autoregressive(1) |
| ATC | Anatomic Therapeutic Chemical classification |
| BARS | Barnes Akathisia Rating Scale |
| BLQ | Lower limit of quantification |
| BMI | Body mass index |
| BPRS | Brief Psychiatric Rating Scale |
| C-SSRS | Columbia – Suicide Severity Rating Scale |
| CDSS | Calgary Depression Scale for Schizophrenia |
| CGI-S | Clinical Global Impression Scale – Severity |
| CI | Confidence interval |
| CMH | Cochran-Mantel-Haenszel |
| CPK | Creatine phosphokinase |
| CS | Compound symmetry |
| CSH | Heterogeneous compound symmetry |
| DMC | Data Monitoring Committee |
| ECG | Electrocardiogram |
| eCRF | Electronic Case Report Form |
| ENR | All Subjects Enrolled |
| EPS | Extrapyramidal symptoms |
| FA0 | No Diagonal Factor Analytic |
| FWER | Familywise error rate |
| HR | Heart rate |
| IEC | Independent Ethics Committee |
| IRB | Institutional Review Board |
| ITI | Intra-Cellular Therapies, Inc. (Sponsor) |
| ITT | Intent-to-treat |
| LOCF | Last observation carried forward |
| LS | Least squares |
| MAR | Missing at random |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MI | Multiple imputation |
| MMRM | Mixed-Effect Model Repeated Measure |
| MNAR | Missing not at random |
| MTP | Multiple testing procedure |
| PANSS | Positive and Negative Syndrome Scale |
| PK | Pharmacokinetic(s) |
| PP | Per protocol |
| PR | PR interval of the electrocardiogram; time duration between the P and R waves |
| PSP | Personal and Social Performance Scale |

| | |
|---|---|
| PT | Preferred Term |
| QD | Once daily |
| QRS | QRS interval of the electrocardiogram; duration of the QRS complex |
| QT | QT interval of ECG, duration between the Q and T waves |
| QTc | QT interval of ECG corrected for heart rate |
| QTcB | QT interval of ECG corrected for heart rate using Bazett's formula |
| QTcB | QT interval of ECG corrected for heart rate using Fridericia's formula |
| RND | All Subjects Randomized |
| RR | Time duration between two consecutive R waves of the electrocardiogram |
| SAE | Serious Adverse Event |
| SAF | Safety Analysis |
| SAP | Statistical Analysis Plan |
| SAS | Simpson-Angus Scale |
| SAS® | Statistical Analysis Software |
| SCID-CT | Structured Clinical Interview for DSM Disorders - Clinical Trial Version |
| SD | Standard deviation |
| SMQs | Standard MedDRA Queries |
| SOC | System Organ Class |
| TEAE | Treatment-emergent adverse event |
| TOEP | Toeplitz structure |
| TOEPH | Toeplitz structure |
| ULQ | Upper limit of quantification |

# 2. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of efficacy, safety, and pharmacokinetic (PK) data for Protocol ITI-007-301. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

This statistical analysis plan (SAP) is based on protocol version 2.1, dated 23Oct2014.

# 3. STUDY OBJECTIVES

## 3.1. PRIMARY OBJECTIVE

The primary objective is to determine whether ITI-007 administered to patients with acutely exacerbated schizophrenia demonstrates antipsychotic efficacy compared to placebo, as measured by change from baseline to Day 28 on the Positive and Negative Syndrome Scale (PANSS) total score.

## 3.2. SECONDARY OBJECTIVES

The secondary objectives are to determine whether, compared to placebo, ITI-007 administered to patients with acutely exacerbated schizophrenia:

- demonstrates enhanced social function as measured by change from baseline on a PANSS-derived Social Factor;
- demonstrates enhanced social function as measured by change from baseline on the Personal and Social Performance Scale (PSP);
- demonstrates efficacy as assessed by change from baseline on any of the individual PANSS subscales: Positive, Negative, and General Psychopathology;
- demonstrates efficacy as assessed by change from baseline on the negative symptoms subscale of the PANSS in a subgroup of patients with prominent negative symptoms at baseline (a score of 4 or higher on at least 3 negative subscale symptom items) and as assessed by change from baseline on a PANSS derived Negative Symptom Factor in a subgroup of patients with prominent negative symptoms at baseline (a score of 4 or higher on at least 3 Negative Symptom Factor items);
- demonstrates an improvement in symptoms of depression as measured by the Calgary Depression Scale for Schizophrenia (CDSS) in a subgroup of patients with co-morbid depressive symptoms at baseline (CDSS total score > 6);
- demonstrates an improvement in symptoms of psychosis as measured by the total PANSS and PANSS subscales in a subgroup of patients with co-morbid depressive symptoms at baseline (CDSS total score > 6);
- demonstrates therapeutic efficacy as assessed by change from baseline on the Clinical Global Impression Scale-Severity (CGI-S).

## 3.3. EXPLORATORY OBJECTIVES

The exploratory objectives of the study include, but are not limited to, the evaluation of the association between response to treatment and any genetic trait or change from baseline in protein biomarkers. Gene and protein biomarker analyses will be specified in a separate plan and reported separately.

## 3.4. SAFETY OBJECTIVES

The safety objectives are to determine whether ITI-007 is safe and well tolerated.

# 4. STUDY DESIGN

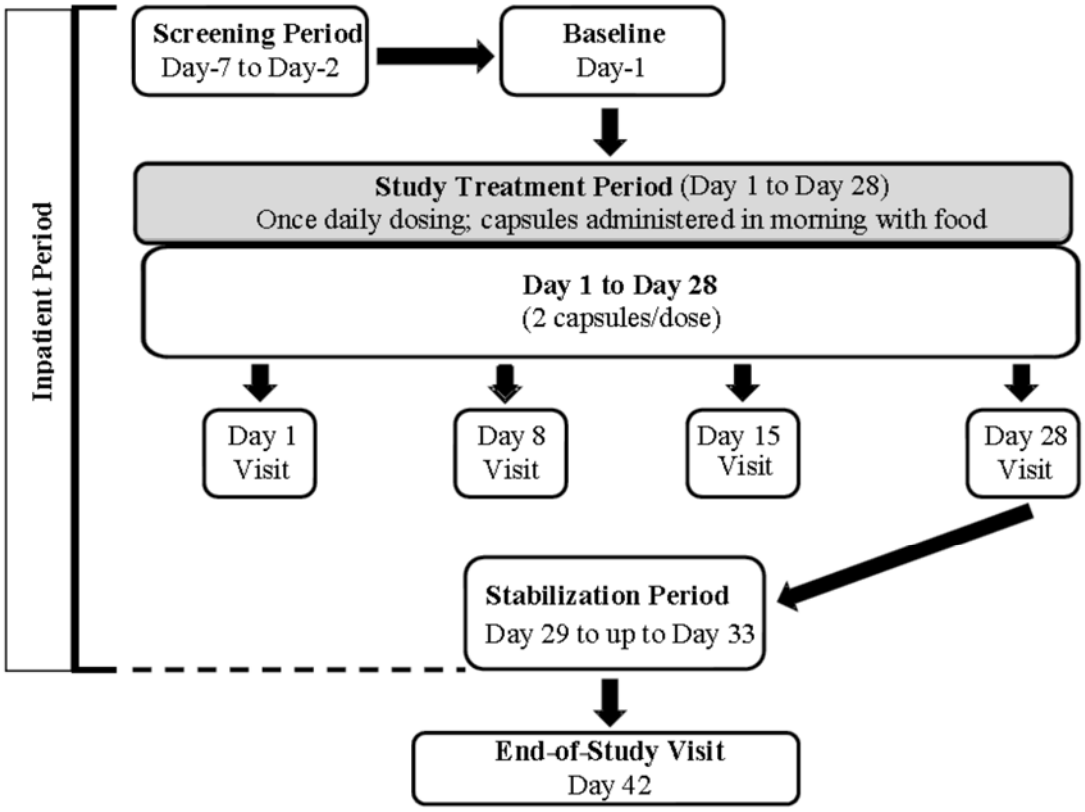
## 4.1. GENERAL DESCRIPTION

The study will be conducted as a randomized, double-blind, parallel-group, placebo-controlled, multi-center study in patients diagnosed with schizophrenia having an acute exacerbation of psychosis.

Approximately 440 subjects will be randomized in a 1:1:1 ratio to receive one of three study treatments: ITI-007 40 mg QD, ITI-007 60 mg QD, and placebo QD.

Study participation will last approximately 7 weeks and include an inpatient *Screening Period* (2 to 7 days prior to Day 1), an inpatient *Study Treatment Period* with 4 weeks of daily administration of study treatment, up to a 5-day inpatient *Stabilization Period* during which subjects will be stabilized on standard antipsychotic medication, and an outpatient *Final Safety Follow-up* approximately 2 weeks after the end of the Treatment Period. The study design is shown schematically in Figure A. The timing and assessments during each study period are described in the Schedule of Assessments (Section 6.1 of the protocol and Appendix 1).

**Figure A: Schematic of Study Design**




## 4.2. CHANGES TO ANALYSIS FROM PROTOCOL

Section 7.1.3 of the protocol specifies a summary of the number and percentage of screen failures by reason. Specific screen fail reasons other than inclusion/exclusion criteria not met were not collected for this study and therefore only the number and percentage of screen failures will be provided. Any inclusion/exclusion criteria not met will be provided in a listing.

Section 7.1.6 of the protocol describes a hierarchical testing approach to preserve the alpha level at 0.05 when testing the primary endpoint for each dose group (ITI-007 60 mg and ITI-007 40 mg) versus placebo. In this SAP, we introduce a key secondary endpoint and a gatekeeping procedure to address the multiplicity issue (see Section 7.2 of this SAP for further detail).

Section 7.1.6 of the protocol states the sensitivity analysis will be conducted on the ITT and PP analysis sets. The sensitivity analysis (detailed Section 15.1.4 of this SAP) is a pattern-mixture model using a placebo-based multiple imputation method for subjects who discontinue early. The PP set (defined in Section 6.4 of this SAP) will not contain any subjects who discontinue early and therefore the sensitivity analysis will only be performed for the ITT analysis set.

The PANSS responder analysis is defined in Section 7.1.6 of the protocol and states that missing data will be imputed using LOCF method. To apply a more conservative approach, any missing data at Day 28 will be defined as a non-responder (see Section 15.3.2 of this SAP).

# 5. FINAL AND PLANNED ANALYSES

There are no planned Data Monitoring Committees (DMCs) or interim analyses for this study and only final analyses will be performed.

All final, planned analyses detailed in this SAP will be performed by [redacted] following sponsor authorization of this Statistical Analysis Plan, Database Lock, Sponsor Authorization of Analysis Sets and Unblinding of Treatment.

# 6. ANALYSIS SETS

Analysis of efficacy and safety endpoints will be performed based on the analysis sets defined in this section and as specified for each endpoint throughout this SAP. Inclusion/exclusion of subjects from each analysis set will be determined prior to the unblinding of the study based on blinded data review.

## 6.1. ALL SUBJECTS ENROLLED [ENR] SET

The all subjects enrolled (ENR) set will contain all subjects who provide informed consent for this study.

## 6.2. ALL SUBJECTS RANDOMIZED [RND] SET

The all subjects randomized (RND) set will contain all subjects who signed informed consent and were randomized to study medication.

For analyses and displays based on RND, subjects will be classified according to randomized treatment.

## 6.3. INTENT-TO-TREAT [ITT] SET

The intent-to-treat (ITT) set will contain all randomized subjects who received at least one dose of study medication, and had data recorded for at least one valid baseline and at least one valid post baseline PANSS measurement. Subjects will be classified according to randomized treatment.

## 6.4. PER PROTOCOL [PP] SET

The per-protocol (PP) set will contain all randomized subjects who completed study treatment per protocol (i.e., completed 28 days of treatment) and who did not experience any major protocol deviations, as specified in Section 10. Deviations from the protocol which, based on clinical input, may have an impact on the efficacy analyses will be identified prior to database lock and considered major protocol deviations. Subjects will be classified according to randomized treatment.

## 6.5. SAFETY ANALYSIS [SAF] SET

The safety analysis (SAF) set will contain all randomized subjects who received at least one dose of study medication.

For analyses and displays based on the SAF, subjects will be classified according to treatment received.

## 6.6. PHARMACOKINETIC [PK] SET

Pharmacokinetic (PK) set will contain all randomized subjects who received at least one dose of the study medication, have a baseline assessment as well as at least one valid post baseline PANSS measurement, and for whom a valid assay result (according to laboratory guidelines) has been obtained. Subjects with partial data will be evaluated on a case-by-case basis to determine if sufficient data are available for meaningful analyses.

# 7. GENERAL CONSIDERATIONS

Relative Study Day will be calculated from the date of Day 1, which is the day of first treatment with study medication, and will be used to show the start and/or stop days of treatments, study procedures, and events.

- If the date of the treatment, procedure, or event is on or after Day 1 date then:

  Relative Study Day = (date of event – Day 1 date) + 1.

- If the date of the treatment, procedure, or event is prior to the Day 1 date then:

  Relative Study Day = (date of event – Day 1 date).

In the situation where the date is partial or missing, Relative Study Day, and any corresponding event durations will appear missing in the listings.

Analyses presented by visit or study day will be based on the scheduled visits as planned in the protocol. Visit windows for unscheduled visits or early discontinuation visits are defined in Table A, which provides the mapping of relative day ranges to the scheduled target days and the study periods. If more than one assessment is available in the same 'Range of Relative Study Days' (window), the

assessment closest to the Scheduled Target Day will be selected and assigned to the Scheduled Target Day. If two or more assessments are available in the same window and are equidistant from the Scheduled Target Day, the latest assessment will be selected.

**Table A: Mapping of Relative Day Ranges to Schedule Target Day**

| Study Phase | Study Period | Range of Relative Study Days | Scheduled Target Day |
|---|---|---|---|
| Pre-Treatment | Screening | 2-7 days before Day 1 date | Day -7 |
| Pre-Treatment | Baseline | 1 day before Day 1 date (definition of baseline varies by assessment – see descriptions below) | Day -1 |
| Study Treatment | Treatment | 2 to 4 days relative to Day 1 date | Day 1 |
| Study Treatment | | 5 to 11 days relative to Day 1 date | Day 8 |
| Study Treatment | | 12 to 18 days relative to Day 1 date | Day 15 |
| Study Treatment | | 19 to 25 days relative to Day 1 date | Day 22 |
| Study Treatment | | ≥26 days relative to Day 1 date and before the start of the Stabilization Period. | Day 28 |
| Post-treatment | Stabilization | ≤ 35 days relative to Day 1 date and after the start of the Stabilization Period. | Day 33 |
| Post-treatment | End of Safety Follow-up | > 35 days relative to Day 1 date and after the start of the Stabilization Period. | Day 42 |

First treatment with study medication is scheduled for Visit 3 on Relative Study Day 1 and baseline assessments are scheduled for Visit 2 on Relative Study Day -1. For analysis purposes, baseline is defined as the last non-missing pre-treatment measurement.

Assessments for which their time (hours and minutes) are recorded will be considered baseline if the assessment date and time is before the date and time of the first treatment. Assessments for which only their date is available will be considered baseline if the assessment date is before the date of the first treatment or if the assessment was done on Study Day 1 and, according to the Study Schedule of Assessments, was supposed to be collected on Day 1 prior to treatment.

ECG assessment will be collected in triplicates, once to three times a day (pre-dose, 3-4 hours post-dose, and 6-7 hours post-dose), on pre-specified Study Days. The average of three recordings corresponding to each triplicate collected at Visit 2 on Day -1 at time points approximately 7:00 am, 11:00 am – noon, and 2:00 – 3:00 pm will be considered baseline for time-matched changes from baseline at post-baseline visits at time points pre-dose, 3-4h post-dose, and 6-7h post-dose respectively.

Vital signs assessment are scheduled to be collected multiple times on pre-specified Study Days at the same time points as the ECGs described above. The assessment at Visit 2 on Day -1 at time points approximately 7:00 am, 11:00 am – noon, and 2:00 – 3:00 pm will be considered baseline for time-matched changes from baseline at post-baseline visits at time points pre-dose, 3-4h post-dose, and 6-7h post-dose respectively.

Unless otherwise specified, the following calculations will be used for change from baseline and percent change from baseline:

For quantitative measurements, change from baseline will be calculated as:

- Value at Visit X – Baseline Value

Except where otherwise specified, percent change from baseline will be calculated as:

- $\frac{\text{Value at Visit X} - \text{Baseline Value}}{\text{Baseline Value}} \times 100$

The default significant level for statistical tests will be (5%); confidence intervals will be 95% and all tests will be two-sided, unless otherwise specified in the description of the analyses. All analyses will be conducted using SAS version 9.4 or higher.

## 7.1. MISSING DATA

The total and subscale scores of any assessment with more than one item (PANSS, BARS, SAS, AIMS, etc.) will be derived from individual items. Any individual missing item in any scale will not be imputed. If one or more items are missing at a visit, then the associated total score or subscale score will be set to missing.

The main objective of the analyses in this study is to evaluate treatment effect of ITI-007 compared to placebo if treatment is administered for the planned study duration. In order to evaluate this estimand in the presence of subjects that may discontinue treatment prematurely, the main analyses will be performed based on an assumption of data being missing at random (MAR) so that subjects discontinuing from treatment early are considered to have unobserved values in line with observed outcomes in their treatment arm taking into account the available data prior to discontinuation.

The Mixed Effects Model for repeated Measures (MMRM) will be used for the primary analysis, main analyses of the secondary efficacy endpoints, and exploratory analyses of safety data. The MMRM approach does not impute missing data but is based on all subjects included in the analysis set using all available longitudinal data (either complete or partial). This approach is based on the MAR assumption as described above.

A sensitivity analysis will also be performed to evaluate impact of missing data on study conclusions. Placebo-based multiple imputation sensitivity analysis (see Section 15.1.4) will consider Missing Not At Random (MNAR) mechanism for the missing data, which will provides an estimand of efficacy attributable to the active treatment if received through the time point of interest, while limiting efficacy after early discontinuation to that of the placebo and trial effect.

As an exploratory analysis of the primary efficacy variable, the change from baseline in PANSS total score to Day 28 will be analyzed, using the last observation carried forward imputation method (see Section 15.4.1). Analysis of the LOCF Day 28 endpoint will provide an estimate of efficacy attributable to ITI-007 compared to placebo at the end of period of adherence to treatment without taking into account the ability of subjects to adhere to treatment for the full planned period of 28 days.

For more detail on handling missing efficacy data, see Sections 15.1.2, 15.1.4, 15.2.1.2, 15.2.1.4, 15.3.2, and 15.4.1.

## 7.2. MULTIPLE COMPARISONS/ MULTIPLICITY

The study is designed to evaluate the efficacy of two doses of ITI-007, 60 mg and 40 mg, which requires

multiple comparisons. Testing multiple null hypotheses may result in substantial inflation of the familywise error rate (FWER), i.e., the probability of erroneously rejecting at least one true null hypothesis. It is essential to ensure that a multiple testing procedure (MTP) controls the FWER in the strong sense (Hochberg and Tamhane, 1987), i.e., under any configuration of the true and false null hypotheses, at a preassigned level of significance.

A mixture-based gatekeeping procedure will be used to adjust for multiplicity. The procedure was proposed by Dmitrienko, Kordzakhia and Brechenmacher (2015) and is a modification of the standard method introduced by Dmitrienko and Tamhane (2011, 2013). The key property of the modified mixture-based gatekeeping procedure is that it results in a power gain over the standard one.

There are two hypotheses associated with the primary efficacy endpoint:

- LP1 represents ITI-007 40 mg vs. Placebo for the change from baseline in PANSS total score at Day 28
- HP1 represents ITI-007 60 mg vs. Placebo for the change from baseline in PANSS total score at Day 28

There are two hypotheses associated with the key secondary endpoint:

- LP2 represents ITI-007 40 mg vs. Placebo for the change from baseline in central CGI-S score at Day 28
- HP2 represents ITI-007 60 mg vs. Placebo for the change from baseline in central CGI-S score at Day 28

The following clinically meaningful logical restrictions will be placed on the four tests:

- Test LP2 will be carried out only if LP1 is significant
- Test HP2 will be carried out only if HP1 is significant

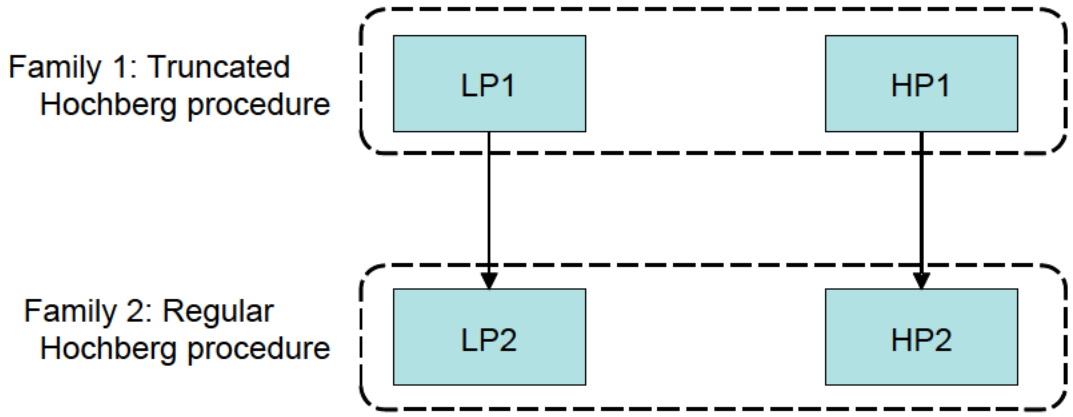
The logical restrictions reflect the well-known fact that carrying out a dose-placebo test on a secondary endpoint is only relevant if a significant treatment effect was established at this dose level on the primary endpoint. The importance of accounting for clinically meaningful restrictions of this kind when setting up multiplicity adjustments was emphasized in O'Neill (1997) and Hung and Wang (2009).

Given these logical restrictions, the four tests can be grouped into two families:

- Family 1: Primary Endpoint - Test LP1 and Test HP1.
- Family 2: Secondary Endpoint - Test LP2 and Test HP2.

The families and logical restrictions are depicted in Figure B. The arrows indicate the dependence between the tests within each dose. For example, testing the primary endpoint on the lower dose (LP1) will be carried out first and, if this test produces a significant outcome, testing the secondary endpoint on the same dose (LP2) will be performed.

**Figure B: Family Specific Tests**




In addition, this gatekeeping procedure accounts for the fact that the test statistics for the two dose-placebo contrasts within each family are positively correlated (the correlation coefficient in each family is equal to 0.5 due to a balanced design in this clinical trial). Specifically, it applies Hochberg-type test that performs well in settings with positively correlated test statistics within Families 1 and 2. Formal definitions of the regular and truncated Hochberg tests are given in Appendix 4. For the truncated Hochberg test, we will set $\gamma$=0.8.

The general decision rules used in this gatekeeping procedure are defined in Figure B. As indicated in Figure B, the hypothesis are tested using the truncated Hochberg test in Family 1 (primary endpoint) because it serves as a gatekeeper for testing family 2 (secondary endpoint) in the sequence. The regular Hochberg test is applied in Family 2 since this is the last family in the testing sequence. Using the truncated Hochberg test allows proceeding to testing Family 2 even if only one dose is significantly different from placebo in Family 1. It is shown in Dmitrienko and Tamhane (2011, 2013) that the Hochberg-based gatekeeping procedure controls the overall Type I error rate in the strong sense at a two-sided $\alpha$=0.05.

The regular and truncated Hochberg tests control the local Type I error rate within each family of hypotheses if the test statistics within each family follow a bivariate normal distribution with a positive correlation and thus the positive dependence condition (MTP2 condition), which guarantees local familywise error rate control, is met (Sarkar and Chang, 1997; Sarkar, 1998). Further, the Hochberg-based gatekeeping procedure does not make any assumptions about the correlation between the primary and key secondary endpoint which are typically difficult to justify in practice.

In general, gatekeeping procedures with semiparametric components (e.g., based on Hochberg or Hommel procedures) control the global FWER in multiplicity problems with several sequences of null hypotheses provided the joint distribution of the test statistics is multivariate normal with non-negative pairwise correlation coefficients (Sarkar, 2008).

Operating characteristics of the mixture-based gatekeeping procedure for this trial were assessed via simulations. The details and results of these simulations are described in Appendix 4.

There will be no multiplicity adjustments made for the remaining secondary endpoints.

## 7.3. EXAMINATION OF SUBGROUPS

The following subgroups will be assessed and described as part of the exploratory or secondary analyses:

**Secondary Analyses**

- Patients with prominent negative symptoms at baseline defined by
  - a score of 4 or higher on at least 3 PANSS Negative subscale symptom items.
  - a score of 4 or higher on at least 3 PANSS Negative Symptom Factor items.
- Patients with co-morbid depressive symptoms at baseline defined as CDSS total score > 6

**Exploratory Analyses**

- Patients with co-morbid depressive symptoms at baseline defined as SCID-CT diagnosis of co-morbid depression at screening
- Age category (≤ 40, > 40 years)
- Gender
- Race (Black or African American or Not Black or African American)
- Ethnicity (Hispanic or Latino or Not Hispanic or Latino
- Weight (≥ 7% weight gain or < 7% weight gain; ≥ 7% weight loss or < 7% weight loss)
- Study Site (Sites with small number of subjects may be combined)
- Years since first diagnosis
- Weeks since start of current exacerbated episode
- Time since last dose of antipsychotic medication
- Severity of disease
  - Number of the positive symptom subscale items with a baseline score ≥ 4 (< 2 or ≥ 2 items)
  - Number of the negative symptom subscale items with a baseline score ≥ 4 (< 3 or ≥ 3 items)

Other subgroup analyses may be performed as deemed appropriate.

## 8. OUTPUT PRESENTATIONS

The templates provided with this SAP describe the presentations for this study and therefore the format and content of the summary tables, figures and listings to be provided by [redacted].

Continuous variables will be summarized using descriptive statistics (number of subjects [n], mean, standard deviation [SD], median, minimum, and maximum, unless otherwise stated). The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean and median will be reported to one more decimal place than the raw data recorded in the database. The standard deviation (SD) will be reported to two more decimal places than the raw data recorded in the database.

Categorical variables will be summarized using frequency counts and percentages.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. P-values less than 0.001 will be presented as "<0.001".

Confidence intervals (CIs) will be presented to two more decimal place than the raw data.

# 9. DISPOSITION AND WITHDRAWALS

Subject disposition and withdrawals, will be presented for the ENR set.

Subject disposition and withdrawals will be presented by treatment group, when applicable, and overall. The number and percentage of subjects who were screened, screen failed, randomized, completed or discontinued the 28-day treatment period, reasons for early treatment discontinuation, completed or withdrew from the study, and reasons for study withdrawal will be presented. The reasons for treatment discontinuation will also be presented by time to treatment discontinuation (≤ Day 8, > Day 8 - ≤ Day 15, > Day 15 - ≤ Day 22, > Day 22 - ≤ Day 28) and by treatment group. The reasons for study withdrawal are listed in Table B. The same reasons apply to treatment discontinuation except for lost to follow-up and screen failure.

A subject is defined to have completed the study treatment if the subject has completed the 28-day treatment period and procedures.

A subject is defined to have completed the study if the subject has completed the End-of-Study assessments on Study Day 42.

**Table B: Reasons for Study Withdrawal and Study Medication Discontinuation Terminology**

| CRF Terminology |
|---|
| Adverse event<br>Adverse event associated with worsening of schizophrenia<br>Adverse event not associated with worsening of schizophrenia |
| Death |
| Lack of efficacy |
| Lost to follow-up |

| |
|---|
| Protocol violation |
| Physician decision |
| Screen Failure |
| Study terminated by sponsor |
| Subject withdrew consent:<br>Personal or family reasons<br>Refused to provide a reason and refused all End of study procedures<br>Self-reported adverse event<br>Self-reported lack of efficacy |
| Other |

Adverse event preferred terms associated with worsening of schizophrenia will be identified by a Medical Monitor review prior to database lock. The number and percentage of randomized subjects who discontinued due to an adverse event associated with worsening of schizophrenia will be summarized. The number and percentage of randomized subjects discontinued due to an adverse event not associated with worsening of schizophrenia will also be presented.

The time to treatment discontinuation due to adverse events, lack of efficacy, or due to any other reason of special interest will be evaluated separately using the Kaplan-Meier method. Time to treatment discontinuation will be defined in days for those subjects who are randomized and receive at least one dose of study medication but discontinue study medication prior to Day 28 as the date of discontinuation minus date of first dose of study medication plus 1. Subjects who complete the treatment period or who discontinue for a reason other than the one being evaluated will be censored. The Log-rank test will be used to compare the time to discontinuation between each treatment group and the placebo group. The same analysis will be repeated for time to treatment discontinuation for any reason, where only subjects who complete the treatment period will be censored.

The number randomized by site, as well as a table of the number and percentage discontinued by visit, will be summarized by treatment group and overall for all randomized subjects.

# 10. PROTOCOL DEVIATIONS

Protocol deviations are any changes to, deviations, or departures from the study design or procedures that are under the investigators' control and that have not been reviewed and approved by the Institutional Review Board (IRB) / Independent Ethics Committee (IEC). Major protocol deviations are any unapproved changes or deviations that may affect the subject's rights, safety or well-being, or are likely to have an impact on the completeness, accuracy and reliability of the study data.

Prior to database lock, a listing of subjects with major protocol deviations will be produced and sent to

Intra-Cellular Therapies, Inc. (ITI) for review and approval. The approved listing of major deviations will be used in conjunction with the treatment information to define the Per Protocol analysis set. Randomized subjects with major protocol deviations will be listed and summarized by treatment group and overall and by deviation category.

Protocol deviations that are considered major include but are not limited to the following:

**Table C: Major Protocol Deviations**

| Protocol Deviation Category | Deviation |
|---|---|
| Informed Consent | The signature date is more recent than the date of screening-related procedures or the date of first treatment with study medication; |
| Inclusion / Exclusion Criteria | Subject did not meet inclusion criteria but randomized and received treatment (For example, subject PANSS score smaller than 70 at baseline);<br>Subject met exclusion criteria but randomized and received treatment (For example, subject used antipsychotic medications during the screening period); |
| Treatment Discontinuation Criteria | Subject met study treatment discontinuation criteria but continued to receive study medication; |
| Visit / Schedule | The time of a pre-dose ECG was later than the treatment time for the same Study Day; |
| Prohibited Medications during the 28-day Treatment Period | Subject took psychotropic medications other than lorazepam, benztropine, or propranolol;<br>Subject took lorazepam, benztropine, or propranolol but failed to follow the regimen as specified in the Table 3 of the protocol; |
| Treatment Unblinded | Subject was unblinded during the treatment period; |
| Incorrect Treatment | Subject took incorrect treatment; |
| Treatment Compliance | < 75% non-missing medication compliance |
| Procedure / Test | Subject has a positive drug screen for substance abuse during the treatment period; |

# 11. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic data and other baseline characteristics will be presented for the ITT analysis set, PP analysis set and SAF.

No formal statistical testing will be carried out for comparing demographic or other baseline characteristics between treatment groups.

The following demographic and other baseline characteristics will be reported for this study:

Demographics

- Age (years) calculated as
  (Date of Informed Consent – Date of Birth)/365.25
- Gender
- Race
- Ethnicity

Other Baseline Characteristics

- Waist circumference (cm)
- Weight (kg)
- Height (cm)
- Body Mass Index (BMI) (kg/$m^2$), ), where
  BMI (kg/ $m^2$) = weight (kg)/ height $(m)^2$

Consumption Habits

- Alcohol consumption
- Caffeine consumption
- Tobacco use

Schizophrenia Diagnosis

- Time since first diagnosis of Schizophrenia (years) calculated as
  (Date of first dose of study medication – Date of First Diagnosis)/365.25
- Time since start of current exacerbated episode (weeks) calculated as
  (Date of first dose of study medication – Start Date of Current Exacerbated Episode)/7
- Time since last dose of antipsychotic medication (days) calculated as
  (Date of first dose of study medication – End Date of Antipsychotic Medication)

Schizophrenia Baseline Efficacy

- Brief Psychiatric Rating Scale (BPRS) total score
- Presence of depression (SCID-CT confirmed co-morbid depression; CDSS score >6)
- Baseline efficacy parameters, including PANSS total score, PANSS pro-social factor score, PANSS positive score, PANSS negative score, PANSS general psychopathology score, CGI-S score (site based (at screening) and central (at baseline)), PSP score, CDSS depression score

Schizophrenia Baseline Safety

- Baseline safety parameters, including BARS total score, BARS global assessment of akathisia, SAS, and AIMS total score

# 12. MEDICAL HISTORY

Medical History information will be presented for the SAF.

Medical History will be coded using MedDRA version 17.1 or higher.

Medical History conditions are defined as those conditions which stop prior to or at Screening.

The number and percentage of subjects with at least one pre-existing condition as well as the number of subjects having pre-existing conditions will be summarized by system organ class (SOC), preferred term (PT), and treatment group. Within each subject, multiple reports of pre-existing conditions that map to a common PT and SOC will be condensed into a single medical history for incidence counts.

# 13. MEDICATIONS

Medications will be presented for the SAF and coded to preferred names using Who Drug Dictionary.

See Appendix 2 for handling of partial dates for medications, in the case where it is not possible to define a medication as prior, concomitant, or post treatment, the medication will be classified by the worst case; i.e. concomitant.

A medication is considered to have started prior to the date and time of first dose of study drug if indicated to have started prior to the first dose of study medication on the eCRF.

A medication is considered to have started after the date and time of the last dose of study drug if indicated to have started after the last dose of study medication on the eCRF.

- 'Prior' medications are medications which started and stopped prior to date and time of first dose of study medication.
- 'Prior concomitant' medications are medications which started prior to and stopped after the date and time of first dose of study medication.
- 'Concomitant' medications are medications which:
  - started after the date and time of first dose of study medication and started prior to the date and time of last dose of study medication,
  - AND stopped after the date and time of first dose of study medication or were ongoing at the completion of Day 28 post-dose assessments or after the last dose of study medication (if planned assessments are not performed).
- 'Post' medications are medications which started after completion of Day 28 assessments which are planned to be conducted after the last dose of study medication.

Prior, prior concomitant, concomitant, and post medication use will be summarized by preferred term using frequencies and percentages by treatment group. Medications will be sorted alphabetically by

preferred term in summaries. Subjects with multiple occurrences of a medication in the same preferred term will only be counted once within the preferred term.

During the study, a subject may be treated with lorazepam for agitation, anxiety, or to aid sleep. In addition, a subject may be treated with benztropine for treatment of extrapyramidal side effects and propanolol may be given for treatment of akathisia. Lorazepam, benztropine and propanolol are allowed according to the instructions outlined in Protocol Table 3, which specify the doses, study days and the time relative to study procedures scheduled for that days. The number and percent of subjects in the ITT set receiving lorazepam, benztropine, or propranolol, the total dose of each of these drugs, and the total number of days on drug will be summarized by treatment group for the screening period and for each week on study treatment (Day 1-8, Day 9-15, Day 16-22, and Day 23 – 28).

# 14. STUDY MEDICATION EXPOSURE AND TREATMENT COMPLIANCE

Exposure to study medication and treatment compliance will be presented for the SAF and ITT analysis sets.

The date of first and last study medication administration will be taken from the eCRF 'Dosing' form. If multiple entries occurred on this form, the earliest 'start of study medication' and the latest 'stop date of study medication' will be used.

Duration of exposure (days) will be calculated as (date of last dose of study medication) – (date of first dose of study medication) + 1. Duration of exposure will be summarized by treatment group as a continuous variable, using mean, SD, median, minimum and maximum. In addition, the number and percentage of subjects within each duration category (1-7 days, 8-14 days, 15-21 days, 22-27 days, 28 days), will be presented.

For each subject, dosing compliance (%) will be calculated as 100 x (number of compliant days) / (number of days in the Treatment Period). A day within the treatment period will be considered a compliant day if the subject had 2 capsules. The "Number of Days in the Treatment Period" is equivalent to the duration of exposure as defined above.

# 15. EFFICACY OUTCOMES

## 15.1. PRIMARY EFFICACY

### 15.1.1. PRIMARY EFFICACY VARIABLE AND DERIVATION

The primary efficacy variable is change from baseline to Day 28 in PANSS total score. The PANSS is an interview-based measure of psychopathology severity in adults with psychotic disorders. Thirty items are rated using a Likert scale, from 1='Absent' to 7='Extreme.' The PANSS will be conducted and scored centrally by a third-party vendor, MedAvante. The PANSS total score is the sum of all 30 items and ranges from 30 through 210. If one or more items are missing at a visit, the total score will also be set to missing.

In general, higher values of PANSS scores represent greater severity of illness.

### 15.1.2. MISSING DATA METHODS FOR PRIMARY EFFICACY ENDPOINT

The use of MMRM as the primary analysis method inherently implies, for example, that the treatment effect on PANSS total score will be similar for those who withdraw and for those who complete the study in their respective treatment arms. To challenge the robustness of this assumption, a sensitivity analysis which utilizes multiple imputations and a different assumption about unobserved outcomes will be performed, as detailed in Section 15.1.4.

### 15.1.3. PRIMARY ANALYSIS OF PRIMARY EFFICACY VARIABLE

The primary objective of this study is to test the hypothesis that ITI-007 administered to patients with acutely exacerbated schizophrenia demonstrates antipsychotic efficacy compared to placebo. The null hypothesis is that there is no difference between treatment and placebo in the change from baseline in PANSS total score at Day 28. The alternative hypothesis is that treatment has greater reduction than the placebo in change from baseline in PANSS total score at Day 28.

The treatment effect on the primary efficacy endpoint, change from baseline to Day 28 in PANSS total score, will be evaluated using an MMRM model. The model will include the change from baseline at each pre-specified timepoint as the response variable and study visit, baseline PANSS total score, baseline PANSS total score-by-study visit interaction, treatment (ITI-007 40 mg, ITI-007 60 mg, placebo) and treatment-by-study visit interaction. An unstructured covariance matrix will be used to model the correlation among repeated measurements within subject. In the event the convergence cannot be attained with the unstructured correlation matrix, the following alternative structures will be attempted in the specified order: heterogeneous Toeplitz structure (TOEPH), heterogeneous autoregressive(1) (ARH(1)), heterogeneous compound symmetry (CSH), No Diagonal Factor Analytic (FA0(q), with q equal to the number of time points), Toeplitz structure (TOEP), autoregressive(1) (AR(1)), and compound symmetry (CS). If either ARH(1) or AR(1) structure is used, a random subject intercept will also be included in the model. The Kenward-Rogers method will be used to estimate the denominator degrees of freedom. The treatment and treatment-by-time interaction terms allow for comparisons of the treatment groups at each of the following time points: Days 8, 15, 22, and 28. Treatment differences will be evaluated via contrasts for the time-by-treatment factor. This methodology will be used to compare the treatment groups versus placebo for change from baseline at each time point. The change from baseline to Day 28 will be used for the primary efficacy analysis.

A mixture-based gatekeeping procedure will be implemented to preserve the Type 1 error rate at the two-sided 0.05 level for the multiple dose group comparisons of the primary and key secondary efficacy endpoints. The key secondary endpoint, central CGI-S score, is described in Section 15.2.1.5. Refer to Section 7.2 and Appendix 4 for more details on the gatekeeping procedure. Adjusted p-values will be presented by treatment group for each endpoint.

The primary efficacy analysis will be performed for the ITT analysis set. The same analysis will also be conducted for the PP analysis set.

Estimates of model parameters will be presented, as well as least squares mean (LSM) estimates for change from baseline in PANSS score, standard errors and 95% confidence intervals (CI) for LSMs will be presented by treatment group and time point. Contrast estimates (LSMs) for between-group comparisons (e.g., ITI-007 40 mg vs. placebo and ITI-007 60 mg vs. placebo), the corresponding

standard errors, 95% CIs, effect sizes, and p-values will be presented for each visit. Effect size will be calculated for each ITI dose group as $\frac{LS\ Mean\ Difference}{LS\ Mean\ Difference\ Standard\ Error \sqrt{\frac{1}{n1}+\frac{1}{n2}}}$, where n1 and n2 are the number of patients in the model for the ITI dose group and placebo, respectively.

### 15.1.4. SENSITIVITY ANALYSIS OF PRIMARY EFFICACY VARIABLE

A pattern-mixture model using a placebo-based multiple imputation method will be performed as a sensitivity analysis to explore the robustness of the MMRM results for the primary efficacy analysis. The primary analysis assumes a MAR mechanism for missing data. This placebo-based multiple imputation sensitivity analysis will consider a Missing Not At Random (MNAR) mechanism for the missing data, where it will be assumed that subjects who discontinue early from the ITI-007 treatment will follow the trajectory of outcomes similar to the one in the placebo arm after discontinuation of the ITI-007 treatment, taking into account observed values prior to discontinuation (Little and Yau, 1996; Ratitch et al., 2013). Subjects discontinuing early from the placebo arm will be assumed to have unobserved outcomes similar to placebo subjects who remain on placebo treatment. The assumption that efficacy profiles of dropouts after discontinuation of active treatment are similar to those of placebo subjects provides an estimand of efficacy attributable to the active treatment if received through the time point of interest, while limiting efficacy after early discontinuation to that of the placebo and trial effect. This strategy is conservative with respect to the experimental treatment because this methodology tends to minimize the difference between treatment and placebo groups. If the results of this analysis are in line with the primary efficacy results, then it can be concluded that the primary analysis results are robust.

A subject with complete data will have measurements at baseline, Day 8, Day 15, Day 22, and Day 28. We say that a patient with missing data at Day 8, 15, or 22 and a non-missing PANSS score at Day 28 has intermittent missing data. A subject with missing data at a post baseline visit and all subsequent visits is said to have monotone missing data.

Intermittent missing data will be multiply imputed using the MAR assumption within each treatment arm. Monotone missing data for all subjects (from placebo and ITI-007 treatment arms) will be multiply imputed using a placebo-based imputation model.

The steps to implement this sensitivity analysis are detailed in Appendix 3.

The sensitivity analysis will be conducted for the ITT analysis set.

## 15.2. SECONDARY EFFICACY

The secondary efficacy analyses will be performed for the ITT analysis set.

### 15.2.1. SECONDARY EFFICACY VARIABLES, DERIVATIONS, AND ANALYSES

#### 15.2.1.1. PANSS subscales

The PANSS Pro-social subscale, also referred to as the derived Social Factor, is a subscale of the PANSS which assess social functioning. It includes six PANSS items related to active and passive social

avoidance, emotional withdrawal, stereotyped thinking, and suspiciousness. PANSS Pro-social subscale score = P6+P3+N4+N2+N7+G16. The possible range for the pro-social PANSS subscore is 6 to 42.

The PANSS Positive subscale assesses hallucinations, delusions and related symptoms and consists of the following 7 items: Delusions (P1), Conceptual disorganization (P2), Hallucinatory behavior (P3), Excitement (P4), Grandiosity (P5), Suspiciousness (P6), and Hostility (P7). PANSS Positive subscale score = P1+P2+P3+P4+P5+P6+P7. The possible range for the positive PANSS subscore is 7 to 49.

The PANSS Negative subscale assesses emotional withdrawal and similar symptoms. It consists of the following 7 items: Blunted affect (N1), Emotional withdrawal (N2), Poor rapport (N3), Passive/apathetic social withdrawal (N4), Difficulty in abstract thinking (N5), Lack of spontaneity and flow of conversation (N6), and Stereotyped thinking (N7). PANSS Negative subscale score = N1+N2+N3+N4+N5+N6+N7. The possible range for the negative PANSS subscore is 7 to 49.

The PANSS General Psychopathology subscale addresses other symptoms such as anxiety, somatic concern and disorientation. The General Psychopathology PANSS subscore consists of the following 16 items: Somatic concern (G1), Anxiety (G2), Guilt feelings (G3), Tension (G4), Mannerisms and posturing (G5), Depression (G6), Motor retardation (G7), Uncooperativeness (G8), Unusual thought content (G9), Disorientation (G10), Poor attention (G11), Lack of judgment and insight (G12), Disturbance of volition (G13), Poor impulse control (G14), Preoccupation (G15), and Active social avoidance (G16). PANSS General Psychopathology subscale score = G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G13+ G14+G15+G16.The possible range for the general psychopathology PANSS subscore is 16 to 112.

All of the above PANSS subscores will be calculated centrally by a third-party vendor, MedAvante. If one or more items for a given subscale are missing, then the corresponding subscore will be set to missing.

Change from baseline for each PANSS subscale described above will be evaluated using a similar MMRM methodology as specified for the primary analysis while substituting baseline PANSS total score in the model with an appropriate baseline score for the variable being analyzed. See the primary analysis Section 15.1.3 for more details.

Change from baseline in the PANSS negative subscale will be evaluated using a similar MMRM approach for patients with prominent negative symptoms at baseline. A patient will be classified as having prominent negative symptoms at baseline if at least 3 PANSS Negative subscale items have a score of 4 or higher.

A similar MMRM approach will also be used to evaluate the change from baseline in the PANSS total score and subscales in a subgroup of patients with co-morbid depressive symptoms at baseline (CDSS total score >6).

Summary statistics, including number of subjects, mean, SD, median, minimum and maximum, of each PANSS item at baseline and on Days 8, 15, 22 and 28 will be presented by Study Day and treatment group. An analysis of change from baseline to Day 8, 15, 22 and 28 of each PANSS item separately will also be conducted. A t-test will be used for comparison of mean change from baseline between each ITI-007 treatment group versus placebo. Mean differences, corresponding 95% CIs, and p-values will be presented.

#### 15.2.1.2. PSP total score

The PSP assesses the severity of personal and social dysfunction. The total score is calculated by the

investigator and ranges from 1 to 100 with higher scores representing better personal and social functioning.

Change from baseline to Day 28 in PSP total score will be calculated. For subjects discontinuing early, a multiple imputation model under the assumption of MAR mechanism, using baseline value as a predictor, will be used to impute missing PSP total scores at Day 28. The observed and imputed data will be analyzed using an ANCOVA model, with effects for baseline score, and treatment. Least squares means for each treatment group, LSM differences between treatment groups, the associated standard errors and two-sided 95% confidence intervals for the LSMs, and p-values for between-treatment tests of differences at Day 28 will be presented.

#### 15.2.1.3. PANSS Negative Symptom Factor score

The PANSS Marder Negative Factor score is defined as the sum of PANSS items N1, N2, N3, N4, N6, G7, G16. The Van Der Gaag Negative Factor score is defined as the sum of PANSS items N1, N2, N3, N4, N6, G7, G8, G16.

Change from baseline in each of the PANSS Negative Factor score for patients with prominent negative symptoms at baseline will be evaluated using a similar MMRM methodology as specified for the primary analysis while substituting baseline PANSS total score in the model with baseline score for the variable being analyzed. See the primary analysis Section 15.1.3 for more details. A patient will be classified as having prominent negative symptoms at baseline if at least 3 PANSS Negative Factor subscale items have a score of 4 or higher.

#### 15.2.1.4. CDSS total score

The CDSS assess the level of depression in schizophrenia. There are nine items rated on a scale from 0='Absent' to 3='Severe'. The CDSS total score is calculated by summing the nine items, resulting in a range from 0 to 27. A higher score is associated with greater severity of depression. If one or more items are missing, then the CDSS total score will be set to missing.

Change from baseline in CDSS total score at Day 28 will be calculated. For subjects discontinuing early, a multiple imputation model under the assumption of MAR mechanism, using baseline value as a predictor, will be used to impute missing CDSS total scores at Day 28. The observed and imputed data will be analyzed using an ANCOVA model, with effects for baseline score, and treatment. Least squares means for each treatment group, LSM differences between treatment groups, the associated standard errors and two-sided 95% confidence intervals for the LSMs, and p-values for between-treatment tests of differences at Day 28 will be presented.

Change from baseline in CDSS total score at Day 28 will be evaluated using a similar imputation and ANCOVA approach for those subject with co-morbid depression at baseline (CDSS total baseline score > 6).

#### 15.2.1.5. CGI-S score

The change from baseline to Day 28 in central CGI-S is the key secondary endpoint.

The CGI-S is a single value assessment of illness severity and ranges from 1= 'Normal, not at all ill' to 7= 'Among the most extremely ill subjects'. A higher score is associated with greater illness severity, and a baseline score of at least 4 is required to be eligible to participate in the study.

The CGI-S will be assessed locally at the site and/or centrally by a third-party vendor, MedAvante. Site based CGI-S assessments are conducted at Screening, Days 8, 15, 22, 28, 33, and 42. Central CGI-S assessments are conducted at Baseline, Days 8, 15, 22, and 28. The site based and central CGI-S assessments will be analyzed separately. For the site based assessment, Days 33 and 42 will not be included in the efficacy analyses but will be summarized descriptively.

Change from baseline for each CGI-S score described above will be evaluated using a similar MMRM methodology as specified for the primary analysis while substituting baseline PANSS total score in the model with an appropriate baseline score for the variable being analyzed. See the primary analysis Section 15.1.3 for more details.

## 15.3. OTHER EFFICACY

### 15.3.1. MARDER AND VAN DER GAAG PANSS FACTORS

There are five PANSS Factor scores defined separately by Marder and Van Der Gaag. The Negative Factor score was previously defined for both in Section 15.2.1.3. The remaining four Marder factor scores are derived as follows:

- Positive: Sum of items P1, P3, P5, P6, N7, G1, G9, G12
- Disorganization: Sum of items P2, N5, G5, G10, G13, G15
- Excitement: Sum of items P4, P7, G8, G14
- Emotional Distress: Sum of items G2, G3, G4, G6

Similarly, the four remaining PANSS Van Der Gaag Factor scores derived as follows:

- Positive: Sum of items P1, P3, P5, P6, N7, G9
- Disorganization: Sum of items P2, N5, N7, G10, G11
- Excitement: Sum of items P4, P7, G8, G14
- Emotional Distress: Sum of items G2, G3, G4, G6

For each factor, if any one of the component items has a missing value then the corresponding factor score is considered missing. Higher scores indicate greater severity of illness.

Baseline values, post baseline values and change from baseline in each PANSS factor score for each of the Five Factors will be summarized by visit and treatment using descriptive statistics including number of subjects, mean, SD, median, minimum, and maximum. The efficacy variables for each of the Five Factors are the change from baseline to Days 8, 15, 22, and 28. The difference between the treatment groups (ITI-007 40 mg and ITI-007 60 mg) and placebo will be analyzed using the MMRM as detailed for the primary efficacy variable. In addition to the summary statistics described above, changes from baseline to Days 8, 15, 22, and 28, will also be summarized using LSMs, difference between LSMs, the associated standard errors and 95% CIs for the LSMs and differences, effect size, and p-value as obtained from the MMRM analysis.

### 15.3.2. PANSS RESPONDER

The PANSS total score will be used for the responder rate analysis. A PANSS response is defined as at least 20, 30, or 40 percent reduction in PANSS total score from baseline to Day 28. Percent reduction will be calculated by first subtracting 30 from the PANSS total scores (baseline values) resulting in a score range starting from 0 and a 'complete cure' that is represented by a 100% decrease from baseline. The modified values will then be used for calculating the percent change from baseline in PANSS total score at Days 8, 15, 22 and 28. A subject will be considered a responder based on three different percent levels: 20%, 30% and 40% reduction in PANSS total score. A subject with missing PANSS assessment on Day 28 will be considered a non-responder, so that the inability to complete the treatment for the planned duration of 28 days is considered as treatment failure.

The number and percentages of responders will be calculated and compared between the treatment groups at each planned visit using the Fisher's Exact Test, separately for each percent level of response.

The time to first response, measured in days from the date of first treatment with study medication to the date that the earliest 20%, 30% or 40% or greater reduction is observed, will be analyzed using the Kaplan-Meier method. Subjects who do not meet the specified thresholds will be censored. The corresponding censored time will be calculated using last visit where the PANSS is assessed. The log-rank test will be used to compare the time to the first response among all treatment groups, as well as for the comparisons between each of the ITI-007 groups and the placebo.

## 15.4. EXPLORATORY EFFICACY

The exploratory efficacy analyses will comprise of a supportive analysis of the primary efficacy variable, and subgroup analyses.

### 15.4.1. PANSS TOTAL SCORE AT THE LAST OBSERVATION CARRIED FORWARD: ANALYSIS OF COVARIANCE

As a supportive analysis of the primary efficacy variable, the change from baseline in PANSS total score to the LOCF Day 28 endpoint will be evaluated, where the LOCF Day 28 endpoint is defined as the last post-baseline measurement collected within the study treatment period. Analysis of the LOCF Day 28 endpoint will provide an estimate of efficacy attributable to ITI-007 compared to placebo at the end of period of adherence to treatment without taking into account the ability of subjects to adhere to treatment for the full planned period of 28 days.

The analysis of change from baseline in PANSS total score to the LOCF Day 28 time point will be performed using an analysis of covariance (ANCOVA), with effects for baseline total PANSS score and treatment. LSMs for each treatment group, the LSM difference between treatment groups, the associated standard errors and two-sided 95% confidence intervals for the differences between the treatment groups, and p-values for between-treatment tests of differences will be presented.

The observed and change from baseline in PANSS total scores will also be summarized at baseline, Days 8, 15, and 22 using LOCF imputation method at each visit by treatment group. This analysis will be conducted on the ITT and PP populations.

### 15.4.2. EXPLORATORY SUBGROUP ANALYSES

Refer to Section 7.3 for a complete list of subgroups.

15.4.2.1. Patients with co-morbid depressive symptoms at screening based on SCID-CT diagnosis

A patient will be classified as having co-morbid depressive symptoms at screening if the patient had SCID-CT diagnosis of co-morbid depression at screening.

The analysis of PANSS total score as described in Section 15.1.3 will be repeated for this subgroup. The analysis of PANSS subscales and CDSS total score described in Sections 15.2.1.1 and 15.2.1.4 respectively will also be repeated for this subgroup.

15.4.2.2. Weight

The primary efficacy endpoint analysis using MMRM and the analysis for PANSS Positive and Negative subscales described in Section 15.2.1.1 will be repeated for weight subgroups. The subgroups are defined as weight gain at Day 28 (≥ 7%, < 7%) or weight loss at Day 28 (≥ 7%, < 7%).

15.4.2.3. Age, Sex, Race, and Ethnicity

The primary efficacy endpoint analyses using MMRM will be repeated with added terms for demographic subgroup, treatment-by-subgroup interaction term, subgroup-by-study visit interaction term, as well as a three-way interaction term for treatment-by-subgroup-by study visit to explore the consistency of treatment effect across certain subgroups. A model will be specified for each of the following demographic subgroups: age category (≤40, >40 years), gender, race (Black or African American or Not Black or African American), and ethnicity (Hispanic or Latino or Not Hispanic or Latino).

Summary statistics and LSMs for change from baseline to Day 28 in PANSS total score will be presented for each treatment group. The p-value for treatment-by-demographic interaction term at Day 28 will also be presented. A Forest plot will be produced presenting LSM difference estimates and their corresponding 95% CIs for each subgroup level of each explored demographic factor, with a vertical line representing the LSM estimate of the treatment difference from the primary analysis model.

15.4.2.4. Other subgroups

The number of weeks since start of the current exacerbated episode will be categorized into four levels: < 1 week, ≥ 1 - < 2 weeks, ≥ 2 - < 3 weeks, ≥ 3 - < 4 weeks, ≥ 4 weeks.

The time since last antipsychotic medication will be categorized into < 2 days, ≥ 2 days - < 5 days, ≥ 5 – < 8 days and ≥ 8 days.

Years since first diagnosis will be categorized as follows: < 2 years, ≥ 2 - < 5 years, ≥ 5 years.

Site and severity of disease will also be analyzed.

The primary efficacy endpoint analyses using MMRM will be repeated for each of these subgroups with added terms for subgroup, treatment-by-subgroup interaction term, subgroup-by-study visit interaction term, as well as a three-way interaction term for treatment-by-subgroup-by study visit to explore the consistency of treatment effect across certain subgroups. Data will be presented and

summarized as outlined in the analysis described in Section 15.4.2.3.

Other subgroup analyses may be performed as deemed appropriate.

# 16. SAFETY OUTCOMES

All outputs for safety outcomes will be based on the Safety Analysis Set.

There will be no statistical comparisons between the treatment groups for safety data, unless otherwise specified within the relevant sections.

Selected safety endpoints will be compared between each of the treatment groups and placebo using the MMRM, whenever applicable, and ANCOVA, for those variables that are measured only twice during the study, once at screening or baseline and once post baseline. These comparisons are considered strictly exploratory.

## 16.1. ADVERSE EVENTS

Adverse Events (AEs) will be coded using Medical Dictionary for Regulatory Activities (MedDRA) central coding dictionary, Version 17.1.

Treatment emergent adverse events (TEAEs) are defined as AEs that started or worsened in severity on or after the first dose of study medication and on or before the last date of study medication.

All AEs with an onset date after the last study medication dose but on or prior to the last dose + 7 days will be listed in the AE data listing and labelled as 'Stabilization Period Adverse Event'. All AE's with an onset date after the last dose of study medication + 7 days will be labelled as 'Follow-up Adverse Event'. Neither will be included in summary tables for TEAEs.

See Appendix 2 for handling of partial dates for AEs. In the case where it is not possible to define an AE as treatment emergent or not, the AE will be classified as the worst case, i.e. treatment emergent.

An overall summary of number of subjects within each of the categories described in the sub-sections 16.1.1 and 16.1.2 below will be provided as specified in the templates.

Listings will include TEAEs and Non-TEAEs.

### 16.1.1. AEs AND TEAEs

Incidence of TEAEs will be presented by System Organ Class (SOC) and Preferred Term (PT) and also broken down further by maximum severity and relationship to study medication.

The number and percentage of subjects with at least one TEAE and total number of subjects having events for each PT and SOC will be summarized. A summary of TEAEs will be provided with only preferred terms and a separate summary for only system organ classes. An additional summary of TEAEs will be provided for preferred terms occurring in at least 5% of subjects in any treatment group (ITI-007 60 mg, ITI-007 40 mg, or placebo). Within each subject, multiple reports of events that map to a common MedDRA PT and/or SOC will be condensed into a single AE for incidence counts. Summaries will be presented by treatment group and in decreasing frequency by decreasing dose group (ITI-007

60 mg, ITI-007 40 mg) and placebo group.

Relative risk of at least one TEAE and for each preferred term and SOC will also be presented along with 95% confidence intervals and p-values obtained by the Chi-square test for association.

Intensity is classed as not specified/mild/ moderate/ severe/ life-threatening (increasing severity). AEs and TEAEs with a missing severity will be classified as "not specified". If a subject reports a TEAE more than once within the same PT and SOC, the event with the worst case severity will be used in the corresponding severity summaries.

Relationship, as indicated by the Investigator, is classed as "not specified", "unrelated", "unlikely to be related", "possibly related", "probably related", "definitely related" (increasing severity of relationship). A "related" TEAE is defined as a TEAE with a relationship to study medication as "possibly related", "probably related" or "definitely related" to study medication. TEAEs with a missing relationship to study medication will be regarded as "not specified". If a subject reports the same AE more than once within the same PT and SOC, the AE with the worst case relationship to study medication will be used in the corresponding relationship summaries.

AEs leading to early withdrawal or discontinuation of study medication will be identified by using the Adverse Events page of the (e)CRF, where item 'Action taken with study drug' indicates permanent discontinuation of study medication, i.e., "Drug withdrawal".

For AEs leading to early withdrawal or discontinuation of study medication, summaries of incidence rates (frequencies and percentages) by SOC and PT will be prepared.

### 16.1.2. SERIOUS ADVERSE EVENTS AND DEATH

Serious adverse events (SAEs) are those events recorded as "Serious" on the Adverse Events page of the (e)CRF. A summary of SAEs by SOC and PT will be prepared.

AEs leading to Death are those events which are recorded as "Fatal" on the Adverse Events page of the (e)CRF. A listing AEs leading to death will be prepared.

### 16.1.3. EXTRAPYRAMIDAL TEAES

Treatment-emergent extrapyramidal adverse events will be defined by the Standard MedDRA Queries (SMQs) shown in the table below. The number and percentage of subjects with at least one AE mapped to a PT contained in SMQ will be presented by treatment. Separate tables will be presented for the narrow and broad interpretation of the SMQ.

**Table D: Extrapyramidal TEAE SMQ**

| Category | Applicable SMQs |
|---|---|
| EPS | SMQ Extrapyramidal Syndrome (narrow and broad) |
| Akathisia | SMQ Akathisia (narrow and broad) |
| Dyskinesia | SMQ Dyskinesia (narrow and broad) |

| | |
|---|---|
| Dystonia | SMQ Dystonia (narrow and broad) |
| Parkonsonism | SMQ Parkinson-like events (narrow and broad) |

## 16.2. LABORATORY EVALUATIONS

Results from the central laboratory will be included in the reporting of this study for Hematology, Clinical Chemistry, and Urinalysis. A list of laboratory assessments to be included in the outputs is included in the protocol, Section 6.3.10. Summary statistics will be presented in SI and US conventional units. Shift tables and abnormalities based on normal range criteria will be presented using SI units only.

Local laboratory tests will be utilized if necessary for this study. Local laboratory results will be converted to the same units as central laboratory tests and summarized together with the central results. In cases where both central and local laboratory results are available for the same time point, the central result will be used for by visit summaries. When central results are unavailable, local results will be summarized.

Quantitative laboratory measurements reported as "< X", i.e. below the lower limit of quantification (BLQ), or "> X", i.e. above the upper limit of quantification (ULQ), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e. as "< X" or "> X" in the listings.

Since prolactin values tend to be elevated in subjects being treated with antipsychotic drugs, prolactin laboratory values will remain blinded to the study team during the trial. Prolactin values will be summarized by treatment group. All laboratory samples are to be collected after an overnight fast. Results for glucose and lipid tests total cholesterol, LDL cholesterol, HDL cholesterol and triglycerides collected non-fasting will be excluded from descriptive statistics but will appear in the data listings.

Change from baseline to Day 28 in total cholesterol, HDL, LDL, glucose, insulin, triglycerides, and prolactin will be analyzed using the MMRM . The model will include terms for study visit, baseline value, baseline value-by-study visit interaction, treatment (ITI-007 40 mg, ITI-007 60 mg, placebo) and treatment-by-study visit interaction. Least squares means for each treatment group, the LSM differences between treatment groups, the associated standard errors and two-sided 95% confidence intervals, and p-values for between-treatment tests of differences will be presented.

The following summaries will be provided for laboratory data:

- Actual and change from baseline by visit (for quantitative measurements)
- Incidence of abnormal values according to normal range criteria
- Shift from baseline according to normal range criteria (for quantitative measurements and categorical measurements)
- Shift from baseline according to markedly abnormal criteria defined in Section 16.2.1 (for quantitative measurements and categorical measurements)

- Listing of subjects meeting markedly abnormal criteria
- Incidence of liver function related values meeting pre-defined as defined in Section 16.2.2 criteria during the treatment period and the entire study

### 16.2.1. LABORATORY REFERENCE RANGES AND MARKEDLY ABNORMAL CRITERIA

Quantitative laboratory measurements will be compared with the relevant laboratory reference ranges in SI units and US Conventional units and categorized as:

- Low: Below the lower limit of the laboratory reference range.
- Normal: Within the laboratory reference range (upper and lower limit included).
- High: Above the upper limit of the laboratory reference range.

In addition to the high and low quantitative laboratory assignments (as identified by means of the laboratory reference ranges), markedly abnormal quantitative safety (and other) laboratory assessments will also be identified in accordance with the predefined markedly abnormal criteria as presented in the table below.

**Table E: Markedly Abnormal Values for Laboratory Evaluations**

| Haematology Parameter | Markedly Abnormal Range |
|---|---|
| Hemoglobin | Male: ≤ 11.5 g/dL |
| | Female: ≤ 9.5 g/dL |
| Hematocrit | Male: ≤ 37% |
| | Female: ≤ 32% |
| WBC | ≤ 2.8 x $10^3$ cells/µL |
| | ≥ 16 x $10^3$ cells/µL |
| Neutrophils (percent) | ≤ 15% |
| Eosinophils (percent) | ≥ 10% |
| Platelet Count | ≤ 75 x $10^3$ cells/µL |
| | ≥ 700 x $10^3$ cells/µL |
| **Chemistry Parameter** | **Markedly Abnormal Range** |
| Alkaline Phosphatase | ≥ 2 x ULN |

| | |
|---|---|
| GGT | ≥ 3 x ULN |
| ALT | ≥ 3 x ULN |
| AST | ≥ 3 x ULN |
| Total Bilirubin | ≥ 2 x ULN |
| Albumin | < 2.5 g/dL |
| Glucose | < 45 mg/dL |
| | >160 mg/dL |
| Sodium | < 130 mmol/L |
| | > 150 mmol/L |
| Potassium | < 3 mmol/L |
| | > 5.5 mmol/L |
| Chloride | < 90 mmol/L |
| | > 115 mmol/L |
| Calcium | < 7 mg/dL |
| | > 12 mg/dL |
| Blood Urea Nitrogen | ≥ 30 mg/dL |
| Creatinine | ≥ 2.0 mg/dL |
| Creatine Phosphokinase (CPK) | ≥ 5 x ULN |
| HbA1c | ≥ 7.5% |
| Prolactin | ≥ 5 x ULN |
| Total Cholesterol (fasting) | ≥ 300 mg/dL |
| LDL Cholesterol (fasting) | ≥ 200 mg/dL |
| Triglycerides (fasting) | ≥ 300 mg/dL |
| **Urinalysis Parameter** | **Markedly Abnormal Range** |
| RBC | > 10 hpf |
| WBC | > 20 hpf |

### 16.2.2. LIVER FUNCTION RELATED CRITERIA

Liver function related laboratory tests will be summarized in accordance to the criteria listed below in Table F.

**Table F: Liver Function Pre-defined Criteria**

| Liver Function Parameter | Criteria |
|---|---|
| ALT | ≥ 3 x ULN |
| | ≥ 5 x ULN |
| AST | ≥ 3 x ULN |
| | ≥ 5 x ULN |
| Total Bilirubin (in combination with ALT or AST criteria) | > 1.5 ULN |
| | > 2 ULN |
| CPK | ≥ 5 ULN |

## 16.3. ECG EVALUATIONS

Results from the central ECG (Electrocardiogram) Reading Centre will be included in the reporting of this study.

The following ECG parameters will be reported for this study:

- PR Interval (msec)
- QRS Interval (msec)
- QT Interval (msec)
- RR Interval (msec)
- QTcF Interval (msec) [derived by central ECG]
- QTcB Interval (msec) [derived by central ECG]
- HR (bpm)

- Overall assessment of ECG:
    - Normal
    - Abnormal

ECGs will be collected in triplicates and will be analyzed as an average of the measures at each time point. For the overall assessment, there are five possible results for each triplicate from the central cardiologist: 'Abnormal, Significant', 'Abnormal Insignificant', 'Incomplete Analysis', 'Normal', and 'Uninterpretable'. Overall assessments with a result of 'Incomplete Analysis' or 'Uninterpretable' will be considered missing. Triplicates will be analyzed according to the worst assessment and categorized in the following order, 'Abnormal Significant', 'Abnormal, Insignificant' and 'Normal'.

The average of three ECG recordings corresponding to each triplicate collected at Visit 2 on Day -1 at time points approximately 7:00 am, 11:00 am – noon, and 2:00 – 3:00 pm will be considered baseline for time-matched changes from baseline at post-baseline visits at time points pre-dose, 3-4h post-dose, and 6-7h post-dose respectively.

The following summaries will be provided for ECG data:

- Actual and change from baseline by visit and time point (for quantitative measurements)
- Incidence of abnormal and normal ECGs by visit and time point.
- Shift from baseline to each visit and time point according to markedly abnormal criteria (for quantitative measurements and categorical measurements)
- Incidence of markedly abnormal results by categories defined in Section 16.3.1
- Listing of subjects meeting markedly abnormal criteria

### 16.3.1. ECG MARKEDLY ABNORMAL CRITERIA

Markedly abnormal quantitative ECG measurements will be identified in accordance with the following predefined markedly abnormal criteria:

- Absolute values for QT interval, QTc interval, QTcB interval and QTcF will be classified as:
    - > 450 msec
    - > 480 msec
    - > 500 msec
- Change from Baseline for QT interval, QTc interval, QTcB interval and QTcF will be classified as:
    - >30 msec increase from baseline
    - >60 msec increase from baseline

## 16.4. VITAL SIGNS

The following Vital Signs measurements will be reported for this study:

- Sitting / Immediately upon Standing/ Supine/ 3 minutes after Standing Systolic Blood Pressure (mmHg)

- Sitting / Immediately upon Standing/ Supine/ 3 minutes after Standing Diastolic Blood Pressure (mmHg)
- Sitting / Immediately upon Standing/ Supine/ 3 minutes after Standing Pulse Rate (bpm)
- Respiratory Rate (breaths/min)
- Oral Temperature ($^{0}$C)
- Weight (kg)
- BMI (kg/m$^2$) [derived – see Section 11]
- Waist Circumference (cm)

Vital signs including systolic blood pressure, diastolic blood pressure, pulse rate, and respiratory rate are collected once to three times a day (pre-dose, 3-4 hours post-dose, and 6-7 hours post-dose), on pre-specified Study Days. The assessment at Visit 2 on Day -1 at time points approximately 7:00 am, 11:00 am – noon, and 2:00 – 3:00 pm will be considered baseline for time-matched changes from baseline at post-baseline visits at time points pre-dose, 3-4h post-dose, and 6-7h post-dose respectively.

BMI measurements classify a subject's weight status as underweight, normal, overweight or obese using Table G. Shifts from Baseline to Day 28 will be produced by treatment group for the Safety population to show the percentage of subjects which fall into each BMI category combination.

The analysis of change from baseline in all vital sign measurements to the LOCF Day 28 time point will be performed using an analysis of covariance (ANCOVA), with effects for baseline value and treatment. LSMs for each treatment group, the LSM difference between treatment groups, the associated standard errors and two-sided 95% confidence intervals for the differences between the treatment groups, and p-values for between-treatment tests of differences will be presented.

**Table G: BMI Weight Status Categories**

| BMI (kg/m$^2$) | Weight Status |
|---|---|
| < 18.5 | Underweight |
| 18.5 <= BMI < 25.0 | Normal |
| 25.0 <= BMI < 30 | Overweight |
| 30.0 <= BMI | Obese |

The following summaries will be provided for vital signs data:

- Actual and change from baseline by visit and time point
- Actual and change from screening by visit for weight and waist circumference
- BMI shift summaries
- Incidence of markedly abnormal values
- Shift from baseline by visit and time point according to markedly abnormal criteria

- Listing of subjects meeting markedly abnormal criteria

### 16.4.1. VITAL SIGNS MARKEDLY ABNORMAL CRITERIA

Markedly abnormal quantitative Vital Signs measurements will be identified in accordance with the following predefined markedly abnormal criteria in the table below.

**Table H: Markedly Abnormal Criteria for Vital Signs**

| Variable | Unit | Low | High |
|---|---|---|---|
| SBP | mmHg | ≤ 90 mmHg AND<br>change from baseline ≤ -20 mmHg | ≥ 180 mmHg AND<br>change from baseline ≥ 20 mmHg |
| DBP | mmHg | ≤ 50 mmHg AND<br>change from ≤ -15 mmHg | ≥ 105 mmHg AND<br>change from baseline ≥ 15 mmHg |
| Pulse rate | Bpm | ≤ 50 bpm AND<br>change from baseline ≤ -15 bpm | ≥ 120 bpm AND<br>change from baseline ≥ 15 bpm |
| Weight | Kg | percentage change from baseline ≤ -7.0 % | percentage change from baseline ≥ 7.0 % |

## 16.5. PHYSICAL EXAMINATION

The following summaries will be provided for physical examination data:

- Incidence of abnormalities at screening
- Incidence of abnormalities post baseline (Day 28 or discontinuation)

## 16.6. OTHER SAFETY ASSESSMENTS

### 16.6.1. BARNES AKATHISIA RATING SCALE (BARS)

The BARS is a rating scale, measuring the observable, restless movements that characterize akathisia. It consists of 4 items: objective restlessness, awareness of restlessness, distress related to restlessness and global clinical assessment of akathisia. Each item is on a 4-point scale 0 to 3, except for the global clinical assessment which is on a 6-point scale 0 to 5, both using low values to represent absence of akathisia and high representing severe akathisia. The BARS total score is the sum of items 1 through 3

and ranges from 0 to 9. Higher values of the BARS total score indicate akathisia is higher in severity. If one or more items at a visit are missing the total will not be calculated.

The observed and change from baseline in BARS total scores will be summarized by visit.

### 16.6.2. ABNORMAL INVOLUNTARY MOVEMENT SCALE (AIMS)

The AIMS is a clinician rated assessment of abnormal movements. It contains items related to: facial and oral movements; extremity movements; trunk movements; global judgments and dental status. The individual items on the AIMS range from 0= "None" to 4= "Severe". The global severity score is the response to "Severity of abnormal movements" found within the global judgments section. The (non-global) total AIMS score is the sum of items 1 through 7. The possible range for the total AIMS score is 0 to 28. Higher values of the total AIMS score indicate increased severity in abnormal movement. If one or more of the AIMS total score items are missing at a visit, the score will be set to missing.

The observed and change from baseline in AIMS total scores will be summarized by visit.

### 16.6.3. SIMPSON-ANGUS RATING SCALE (SAS)

The SAS is a clinician-rated assessment of neuroleptic-induced parkinsonism consisting of 10 items. Items are anchor-based, rated on a 5-point scale and address rigidity, gait (bradykinesia), tremor, glabellar tap, salivation and akathisia. Each individual item on the SAS ranges from 0= "Normal" to 4= "Extreme Symptoms". The total SAS score is defined as the sum of all 10 items and ranges between 0 and 40. Lower values of the SAS total score indicate milder symptoms. If one or more items are missing at a visit the SAS total score will be set to missing.

The observed and change from baseline in SAS total scores will be summarized by visit.

### 16.6.4. COLUMBIA SUICIDE SEVERITY RATING SCALE (C-SSRS)

The C-SSRS is an instrument designed to systematically assess and track suicidal adverse events (suicidal behavior and suicidal ideation) throughout the trial. The C-SSRS includes the following four sections: Suicidal Ideation, Ideation Intensity, Suicidal Behavior and Actual Suicide Attempts. The strength of this suicide classification system is in its ability to accurately and comprehensively assess suicidality, while limiting the over-identification of suicidal behavior. The C-SSRS will be administered by a trained rater at the site.

Suicidal Ideation is rated on a 6-point scale from 0="No ideation present" to 5="Active ideation with plan and intent". A score of 4 or 5 on this scale indicates serious suicidal ideation. Any score greater than 0 will be counted as having suicidal ideation.

The Ideation Intensity total score is the sum of five items from the Ideation Intensity scale: frequency, duration, controllability, deterrents, and reasons for ideation. The possible range for the Intensity total score is 0 to 25. If a subject did not endorse any suicidal ideation, a score of 0 for the intensity total score will be given.

Suicidal Behavior is collected as actual attempt, interrupted attempt, and aborted attempt. Any

attempt will be defined as suicidal behavior.

The number and percentage of subjects with suicidality as measured by the C-SSRS will be summarized, where suicidality is defined as having at least one occurrence of suicidal ideation or at least one occurrence of suicidal behavior. The suicidality indicator will be set to 1 if the subject exhibits suicidality for each visit, 0 if the subject does not exhibit suicidality, and missing otherwise. This data will be summarized at each scheduled visit and overall by study period (treatment and post-treatment (includes re-stabilization and follow-up)).

The number and percentage of subjects with each type of suicidal ideation or any suicidal ideation during each study period will be summarized similarly as above. The most severe ideation, the ideation intensity items (frequency, duration, controllability, deterrents, and reasons for ideation), and the ideation intensity total score will also be summarized descriptively.

The number and percentage of subjects with each type of suicidal behavior (actual attempts, interrupted attempts, aborted attempts, and preparatory acts or behavior) will be summarized for each scheduled visit and overall by study period. The number and percentage of subjects with any suicidal behavior and those completing suicide will also be summarized for each scheduled visit and overall by study period.

Data collected on actual suicide attempts (lethality of actual attempts and potential lethality of attempts) will be presented in a data listing.

An overall summary of C-SSRS data (post-Baseline data across all scheduled visits and by study period, treatment re-stabilization and follow-up) will include the frequency and percentage of the following:

- At least one suicidal ideation post-Baseline
- Emergence of suicidal ideation (no suicidal ideation at Baseline, and any type of suicidal ideation post-Baseline)
- Emergence of serious suicidal ideation (no suicidal ideation at Baseline, and any serious suicidal ideation [ideation score of 4 or 5] post-Baseline)
- Most severe type of ideation post-Baseline
- Worsening of suicidal ideation (most severe suicidal ideation post-Baseline was more severe than it was at Baseline)
- At least one suicidal behavior post-Baseline
- Emergence of suicidal behavior (no suicidal behavior at Baseline, and any type of suicidal behavior post-Baseline)
- At least one actual attempt post-Baseline
- At least one interrupted attempt post-Baseline
- At least one aborted attempt post-Baseline
- At least one preparatory acts or behaviors post-Baseline

- At least one instance of suicidality [any ideation or behavior] post-Baseline
- Emergence of suicidality (no suicidality at Baseline, and any suicidality post-Baseline)
- Any completed suicides post-Baseline

# 17. PHARMACOKINETICS

All PK summaries will be performed on the PK analysis set.

Blood samples will be collected pre-dose and 3-4 h post-dose on Study Days 1, 8, 28, at 6-7 h post-dose on Day 28, and once on Day 33 for determination of ITI-007 (IC200056) and metabolite (IC200161, IC200131, and IC200565) concentrations. Samples will be collected from all subjects in order to maintain the blind. All samples will be analyzed from ITI-007 recipients; selected samples from placebo recipients will be analyzed for presence/absence of active drug to verify if any subjects randomized to placebo took an incorrect (active) drug. Descriptive statistics for trough and peak ITI-007/metabolite levels will be calculated.

Subjects with partial data will be evaluated on a case-by-case basis to determine if sufficient data are available for meaningful analysis. Values excluded from the pharmacokinetic analysis will be flagged with an asterisk and concentration values reported as below the level of quantification (BLQ) will be listed with the lower limit of quantification in parentheses.

Descriptive statistics (N, arithmetic mean, SD, arithmetic CV%, median, minimum and maximum) will be used to summarize ITI-007 and metabolite concentration data at each planned sampling time point for each study part and treatment (ITI-007 60 mg and ITI-007 40 mg). BLQ values will be set to zero prior to calculation of descriptive statistics for the plasma concentration-time profile.

The correlation between plasma concentration and change from baseline in PANSS total score to Day 28 will be summarized by ITI-007/metabolite and each planned sampling time point. Similarly, the correlation between plasma concentration and change from baseline in PANSS total score and PANSS subscales (Positive, Negative, and General Psychopathology) to Day 28 will be summarized by treatment (ITI-007 60 mg and ITI-007 40 mg) and ITI-007/metabolite and each planned sampling time point. Correlation will be measured using Spearman's Rho and Pearson's R and p-values will be presented with each correlation measurement.

# 18. REFERENCES

Allison P. Missing data. In: Handbook of Survey Research, edited by Marsden, P.V. and Wright, J.D. Bingley, UK: Emerald Group Publishing Ltd. 2010; 631-657.

Brechenmacher, T., Xu, J., Dmitrienko, A., Tamhane, A.C. (2011). A mixture gatekeeping procedure based on the Hommel test for clinical trial applications. Journal of Biopharmaceutical Statistics. 21, 748-767.

Dmitrienko, A., D'Agostino, R.B., Huque, M.F. (2013). Key multiplicity issues in clinical drug development. Statistics in Medicine. 32, 1079-1111.

Dmitrienko, A., Kordzakhia, G., Brechenmacher, T. (2015). Mixture-based gatekeeping procedures for multiplicity problems with multiple sequences of hypotheses. Journal of Biopharmaceutical Statistics. To appear.

Dmitrienko, A., Tamhane, A.C., Wiens, B. (2008). General multistage gatekeeping procedures. Biometrical Journal. 50, 667-677.

Dmitrienko, A., Tamhane, A.C. (2011). Mixtures of multiple testing procedures for gatekeeping applications in clinical trials. Statistics in Medicine. 30, 1473-1488.

Dmitrienko, A., Tamhane, A.C. (2013). General theory of mixture procedures for gatekeeping. Biometrical Journal. 55, 402-419.

Dobson, A. J. (2002). An introduction to generalized linear models, 2nd ed. Chapman & Hall/CRC.

EMA. (n.d.). Guideline on Missing Data in Confirmatory Clinical Trials. Retrieved December 15, 2011, from European Medicines Agency: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500096793.pdf

Hochberg, Y., Tamhane, A. C. (1987). Multiple Comparison Procedures. New York: Wiley.

Hung, J., Wang, S.J. (2009). Some controversial multiple testing problems in regulatory applications. Journal of Biopharmaceutical Statistics. 19, 1-11.

Little R, Yau L (1996) Intention-to-treat analysis for longitudinal studies with drop-outs. Biometrics 52: 1324-1333.

Meltzer, H.Y. et al. (2011). Lurasidone in the treatment of schizophrenia: a randomized, double-blind, placebo- and olanzapine-controlled study. American Journal of Psychiatry. 168, 957-967.

O'Neill, R.T. (1997). Secondary endpoints can not be validly analyzed if the primary endpoint does not demonstrate clear statistical significance. Controlled Clinical Trials. 18, 550-556.

Ratitch B, O'Kelly M, Tosiello R (2013) Missing Data in Clinical Trials: from Clinical Assumptions to Statistical Analysis using Pattern Mixture Models. Pharmaceutical Statistics, early view DOI: 10.1002/pst.1549, http://onlinelibrary.wiley.com/doi/10.1002/pst.1549/pdf, accessed 20 June 2013.

Sarkar, S., Chang, C. K. (1997). Simes' method for multiple hypothesis testing with positively dependent test statistics. Journal of the American Statistical Association. 92, 1601-1608.

Sarkar, S. K. (1998). Some probability inequalities for censored MTP2 random variables: A proof of the Simes conjecture. Annals of Statistics. 26, 494-504.

# APPENDIX 1. SCHEDULE OF ASSESSMENTS

| Event | Screening Period | Baseline | Study Treatment Period | | | | | End-of-Re-Stabilization Period | End-of-Study Visit |
|---|---|---|---|---|---|---|---|---|---|
| | -7 to -2 [a] | -1 | 1 | 8 | 15 | 22 | 28 | 33 [b] | Day 42 ± 2 |
| Visit Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Informed consent | Before Any Study-Specific Procedures are Conducted | | | | | | | | |
| Admission to Inpatient Unit | X | | | | | | | | |
| Medical History | X | | | | | | | | |
| Inclusion/Exclusion Criteria Review | X | X | | | | | | | |
| SCID-CT [c] | X | | | | | | | | |
| BPRS [c] | X | | | | | | | | |
| Screening Adjudication Form(s) and Recordings [c] | X | | | | | | | | |
| Modified Physical Exam | X | | | | | | X | | |
| Hepatitis/HIV Testing | X | | | | | | | | |
| Urine Drug and Alcohol Screen [d] | X | | X | | | | | | |
| Urine pregnancy Test | X | | X | | | | X | X | |
| Clinical Laboratories [e] | X | | X | X | | | X | X | |
| 12-Lead ECG [f] | X | X | X | X | | | X | X | |
| Vital Signs [g] | X | X | X | X | X | | X | X | |

## SCHEDULE OF ASSESSMENTS (CONTINUED)

| Event | Screening Period | Baseline | Study Treatment Period | | | | | End-of-Re-Stabilization Period | End-of-Study Visit |
|---|---|---|---|---|---|---|---|---|---|
| | -7 to -2 [a] | -1 | 1 | 8 | 15 | 22 | 28 | 33 [b] | Day 42 ± 2 |
| Visit Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Randomization | | | X | | | | | | |
| Study Drug Dosing | | | X | X | X | X | X | | |
| PK Sample Collection [h] | | | X | X | | | X | X | |
| Samples for protein biomarkers [i] | | | X | | | | X | | |
| Sample for genetic testing [j] | | | X | | | | X | | |
| PANSS [k] | | X | | X | X | X | X | | |
| Central CGI-I [k,l] | X | X | | X | X | X | X | | |
| Site-based CGI-I [k] | | | | X | X | X | X | X | X |
| PSP | | X | | | | | X | | |
| CDSS | | X | | | | | X | | |
| SAS [m] | | X | | X | | X | X | | |
| BARS [m] | | X | | X | | X | X | | |
| AIMS [m] | | X | | X | | X | X | | |
| C-SSRS | X | X | | X | X | | X | X | X |
| Concomitant medication | X | X | X | X | X | X | X | X | X |
| Adverse events [n] | X | X | X | X | X | X | X | X | X |
| Restart Standard Antipsychotics | | | | | | | X | | |

a) Upon signing of the ICF, patients are to be admitted to the inpatient unit, if they are not inpatient already.
b) The End-of-Stabilization period is up to 5 days
c) These assessments are to be recorded and submitted to an independent psychiatrist or clinical psychologist (or panel of psychiatrists and/or clinical psychologists) for review.
d) The Urine Drug and Alcohol Screen will be also performed upon patient's return to the inpatient unit from the day pass.
e) Clinical Laboratory samples are to be taken after an overnight fast of at least 10 hours.
f) ECGs are to be triplicate 10 second epochs with 5 min between recordings. One triplicate ECGs will be performed during Screening and one triplicate ECG will be performed on Day 33. On Days 1, 8, and 28, triplicate ECGs will be taken pre-dose (trough, before breakfast) and at 3-6 hours post-dose (peak, before lunch). An additional triplicate ECG will be performed on Day 28 at 6-7 hours post-dose. Triplicate ECGs on Day -1 are to be performed at the same time of day planned for Day 28: once before breakfast (approximately at 7:00 am corresponding to the same time of day for the trough measure), again before lunch (approximately at 11:00 am – noon, corresponding to the same time of day for the 3-4 hours post-dose peak measure), and again approximately at 2:00-3:00 pm corresponding to the same time of day of the 6-7 hours post-dose measure. In all cases ECGs are conducted before other assessments scheduled in the same time window; for example, when ECG, vital signs and blood sample collection for PK measure are scheduled for the same time window, ECG measures should be conducted first, followed by vital signs and then blood alcohol collection.
g) Respiratory rate, oral temperature, and 3-positional blood pressure and pulse (after at least 10 minutes lying down, after approximately 1 minute sitting, immediately upon standing, and after approximately 3 minutes standing) will be taken at least once at all scheduled visits. On Days 1, 8, 15, 22, and 28 vitals will be taken pre-dose (trough) and 3-6 hours post-dose (peak). Additionally, vital signs will be assessed on Days -1 and 28 at 6-7 hours post-dose. Vital signs are always taken after conducting the ECGs, as applicable, and prior to any other assessments scheduled in the same time window. Height, weight and waist circumference should be collected during screening. Weight and waist circumference should be collected again on Day -1 for baseline and again on Day 28.
h) Blood samples for PK collection are to be collected prior to dosing (trough) and then 3-6 hours post-dose (peak) on Days 1, 8, and 28, again at 6-7 hours post-dose only on Day 28, and then once on Day 33. Samples are collected after ECG and vital signs.
i) Samples for protein biomarkers on Day 1 has be collected pre-dose.
j) Blood sample is to be collected once for genetic testing, if the optional ICF addendum is signed. If the addendum is signed prior to Day 1 then the sample will be taken on Day 1. If it is signed after Day 1, then the sample will be taken on Day 28. It is not to be taken twice.
k) PANSS and CGI-S are to be conducted by a remote central rater. Efforts should be made to schedule the remote interviews at approximately the same time of day for each visit. Due to potential limitations around scheduling of the remote rater, the PANSS may be conducted within a 1-day window provided for this assessment (see Section 6.2). PANSS and CGI scores at Baseline Visit are reviewed for inclusion of subject (subject will be included with the baseline PANSS score of 70 or higher and baseline CGI-S score of 4 or higher).
l) CGI-S is conducted by a qualified site rater at screening (in addition to the CGI-S conducted by the remote central rater) at baseline through Day 28. Site-based CGI-S is conducted by qualified site rater; during the treatment period, the site rater will interview the subject and review baseline and current PANSS and CGI-S scores from the remote centralized rater prior to rating the site-based CGI-S for the visit.
m) On Days 8, 15, and 28, the SAS, BARS, and AIMS assessments are to be conducted 3-6 hours after the administration of study treatment
n) The recording of adverse events is to start immediately after the ICF is signed and continue through until End-of-Study, Visit 9.
o) On Day 28, after all study assessments have been completed, or on Day 29 patients are to be started on standard antipsychotic medication as prescribed by the Investigator

# APPENDIX 2. PARTIAL DATE CONVENTIONS

Imputed dates will NOT be presented in the listings.

## ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS:

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Known | Known | If start date < study med start date or start date>study med end date , then not TEAE<br>If study med start date <=start date <= study med end date , then TEAE |
| | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then:<br>If start date < study med start date or start date>study med end date, then not TEAE<br>If study med start date <=start date <= study med end date , then TEAE |
| | Missing | If start date < study med start date or start date>study med end date, then not TEAE<br>If study med start date <=start date <= study med end date, then TEAE |
| | | |
| Partial, but known components show that it cannot be on or after study med start date or cannot be before study med end date | Known | Not TEAE |
| | Partial | Not TEAE |
| | Missing | Not TEAE |
| | | |
| Partial, could be on or after study med start date but before study med end date | Known | If stop date < study med start date, then not TEAE<br>If stop date >= study med start date, then TEAE |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then:<br>If stop date < study med start date, then not TEAE<br>If stop date >= study med start date, then TEAE |
| | Missing | Assumed TEAE |
| | | |
| Missing | Known | If stop date < study med start date, then not TEAE<br>If stop date >= study med start date, then TEAE |
| | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then:<br>If stop date < study med start date, then not TEAE<br>If stop date >= study med start date, then TEAE |
| | Missing | Assumed TEAE |

## ALGORITHM FOR PRIOR / CONCOMITANT MEDICATIONS:

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Known | Known | If stop date < study med start date, assign as prior<br>If stop date >= study med start date and start date < study med start date, assign as prior concomitant<br>If stop date >= study med start date and start date <= end of treatment, assign as concomitant<br>If stop date >= study med start date and start date > end of treatment, assign as post study |
| | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then:<br>If stop date < study med start date, assign as prior<br>If stop date >= study med start date and start date < study med start date, assign as prior concomitant<br>If stop date >= study med start date and start date <= end of treatment, assign as concomitant<br>If stop date >= study med start date and start date > end of treatment, assign as post study |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| | Missing | If stop date is missing could never be assumed a prior medication<br>If start date <= end of treatment, assign as concomitant<br>If start date > end of treatment, assign as post treatment |
| | | |
| Partial | Known | Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown), then:<br>If stop date < study med start date, assign as prior<br>If stop date >= study med start date and start date < study med start date, assign as prior concomitant<br>If stop date >= study med start date and start date <= end of treatment, assign as concomitant<br>If stop date >= study med start date and start date > end of treatment, assign as post study |
| | Partial | Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown) and impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then:<br>If stop date < study med start date, assign as prior<br>If stop date >= study med start date and start date < study med start date, assign as prior concomitant<br>If stop date >= study med start date and start date <= end of treatment, assign as concomitant<br>If stop date >= study med start date and start date > end of treatment, assign as post study |
| | Missing | Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown), then:<br>If stop date is missing could never be assumed a prior medication<br>If start date <= end of treatment, assign as concomitant<br>If start date > end of treatment, assign as post treatment |
| | | |
| Missing | Known | If stop date < study med start date, assign as prior<br>If stop date >= study med start date, assign as concomitant<br>Cannot be assigned as 'post treatment' |
| | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then:<br>If stop date < study med start date, assign as prior<br>If stop date >= study med start date, assign as concomitant<br>Cannot be assigned as 'post treatment' |
| | Missing | Assign as concomitant |

# APPENDIX 3. PATTERN MIXTURE MODEL IMPLEMENTATION

- Step 1: 1000 datasets will be generated where missing data at intermediate visit(s) will be imputed for each treatment group using non-missing data from all subjects within the treatment group by a Monte

Carlo Markov Chain (MCMC) imputation model using the MCMC statement in the SAS PROC MI procedure. As a result, each dataset will only have missing data at the end of subjects' records, or a monotone missing data pattern. An option to use multiple chains will be used (option CHAIN=MULTIPLE in MCMC statement). The imputation model will include baseline and all post baseline time points of the parameter being imputed. Random number generator seed of 873465 will be used at this step.

- Step 2: For each dataset from Step 1, monotone missing data will be imputed based on information from the placebo group. As a result, 1000 imputed complete datasets will be generated. This will be implemented by using SAS v9.4 PROC MI, statement MNAR MODEL (<variable to impute> / MODELOBS=(<treatment variable>='<placebo level>')). The imputation model will include baseline value and all post baseline time points prior to the one being imputed. Random number generator seed of 12676854 will be used at this step.

- Analyze each imputed complete dataset from Step 2 with an MMRM model used for the primary analysis.

- Combine estimates obtained from MMRM models applied to each imputed dataset using SAS MIANALYZE procedure.

# APPENDIX 4. MULTIPLICITY ADJUSTMENT

## REGULAR AND TRUNCATED HOCHBERG DETAILS

Consider a general problem of testing $m$ null hypotheses denoted by $H_1,\ldots,H_m$. Let $p_1,\ldots,p_m$ denote the associated raw $p$-values. Further, let $p_{(1)}<\ldots<p_{(m)}$ denote the ordered $p$-values and $H_{(1)},\ldots,H_{(m)}$ denote the hypotheses corresponding to the ordered $p$-values. Finally, let $\alpha$ denote the overall Type I error rate.

The regular Hochberg procedure is based on the following testing algorithm:

- Step 1: If $p_{(m)} > \alpha$, accept $H_{(m)}$ and go to Step 2, otherwise reject all null hypotheses and stop.

- Step $i = 2,\ldots,m$-1: If $p_{(m-i+1)} > \alpha/i$, accept $H_{(m-i+1)}$ and go to Step $i$+1, otherwise reject all remaining null hypotheses and stop.

- Step $m$: If $p_{(1)} > \alpha/m$, accept $H_{(1)}$, otherwise reject $H_{(1)}$.

The truncated Hochberg procedure is defined as a convex combination of the Bonferroni procedure and regular Hochberg procedure based on a pre-specified truncation parameter $0 \le \gamma < 1$ (Dmitrienko, Tamhane and Wiens, 2008). The truncated Hochberg procedure is based on the following testing algorithm:

- Step 1: If $p_{(m)} > (\gamma+(1-\gamma)/m)\alpha$, accept $H_{(m)}$ and go to Step 2, otherwise reject all null hypotheses and stop.

- Step $i = 2,\ldots,m$-1: If $p_{(m-i+1)} > (\gamma/i+(1-\gamma)/m)\alpha$, accept $H_{(m-i+1)}$ and go to Step $i$+1, otherwise reject all remaining null hypotheses and stop.

- Step $m$: If $p_{(1)} > \alpha/m$, accept $H_{(1)}$, otherwise reject $H_{(1)}$.

With $\gamma = 0$, the truncated Hochberg procedure simplifies to the Bonferroni procedure and, with $\gamma = 1$, the truncated Hochberg procedure simplifies to the regular Hochberg procedure.

## ASSESSMENT OF OPERATING CHARACTERISTICS

Operating characteristics of the proposed gatekeeping procedure were assessed via simulations. Power evaluation was performed based on the following commonly used criteria:

- Marginal power of each test after a multiplicity adjustment.
- Disjunctive power in Family 1 (Tests LP1 and HP1) and Family 2 (Tests LP2 and HP2) after a multiplicity adjustment. Disjunctive power is defined as the probability of establishing a significant treatment effect at the low dose or high dose within a family, e.g., Test LP1 is significant or Test HP1 is significant in Family 1. This criterion is the most commonly used measure of the overall probability of success in Family 1 since the trial's overall outcome will be declared successful if it least one dose demonstrates a significant improvement over placebo on the primary endpoint.
- Expected number of significant tests in Family 1 (Tests LP1 and HP1) and Family 2 (Tests LP2 and HP2) after a multiplicity adjustment. This criterion tends to be more informative than disjunctive power since it takes into account the number of significant outcomes within a family whereas disjunctive power cannot differentiate between settings where only one test is significant or both tests are significant within a family.

The assumptions listed below were used for the simulations. Even though the mixture-based gatekeeping procedure does not assume correlation between the primary and secondary endpoints, a correlation coefficient is defined below to generate the data for the simulations.

**Design parameters: True effect sizes of the four tests**

| Test | Effect size |
|---|---|
| Test HP1 | 0.4 |
| Test LP1 | 0.4 |
| Test HP2 | 0.3 |
| Test LP2 | 0.3 |

**Design parameters: Correlation between the primary and key secondary endpoint**

| Parameter | Value |
|---|---|
| Correlation | 0.3 |

**Design parameters: Common sample size across the treatment arms**

| Parameter | Value |
|---|---|
| Sample size | 132 |

**Analysis parameters: Truncation parameter used in Family 1**

| Parameter | Value |
|---|---|
| Truncation parameter | 0.8 |

**Analysis parameters: Two-sided significance level (alpha)**

| Parameter | Value |
|---|---|
| Alpha | 0.05 |

**Simulation parameters**

| Parameter | Value |
|---|---|
| Number of simulation runs | 10000 |

The following are the results from the simulations:

**Unadjusted and multiplicity-adjusted marginal power of the individual tests**

| Test | No multiplicity adjustment | Gatekeeping procedure |
|---|---|---|
| Test HP1 | 89.6 | 87.6 |
| Test LP1 | 89.7 | 87.5 |
| Test HP2 | 68.7 | 57.5 |
| Test LP2 | 68.7 | 57.4 |

**Unadjusted and multiplicity-adjusted disjunctive power in Families 1 and 2**

| Family | No multiplicity adjustment | Gatekeeping procedure |
|---|---|---|
| Family 1 | 96.7 | 93.7 |
| Family 2 | 83.1 | 67.7 |

**Unadjusted and multiplicity-adjusted expected number of significant tests in Families 1 and 2**

| Family | No multiplicity adjustment | Gatekeeping procedure |
|---|---|---|
| Family 1 | 1.79 | 1.75 |
| Family 2 | 1.37 | 1.15 |