Statistical Analysis Plan

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Risperidone Extended-Release Injectable Suspension (TV-46000) for Subcutaneous Use as Maintenance Treatment in Adult and Adolescent Patients With Schizophrenia

Study Number TV46000-CNS-30072

NCT03503318

SAP Am 01 Addendum 01 Approval date: 16 December 2020

Statistical Analysis Plan with Amendment 01

Addendum 01

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Risperidone Extended-Release Injectable Suspension (TV-46000) for Subcutaneous Use as Maintenance Treatment in Adult and Adolescent Patients with Schizophrenia

Phase 3

Study TV46000-CNS-30072

IND number 124384; NDA number: 213586;
EudraCT Number: 2018-001619-65

Sponsor

Teva Branded Pharmaceutical

Products R&D, Inc.

145 Brandywine Parkway

West Chester, Pennsylvania 19380

United States

Statistical Analysis Plan Am 01 approval date:	22 April 2020
Addendum approval date:	16 December 2020

This study was performed in compliance with Good Clinical Practice (GCP), including the archiving of essential documents.

STATISTICAL ANALYSIS PLAN AM 01, ADDENDUM 01 APPROVAL

Study Number: TV46000-CNS-30072

Study Title: A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Risperidone Extended-Release Injectable Suspension (TV-46000) for Subcutaneous Use as Maintenance Treatment in Adult and Adolescent Patients with Schizophrenia

Author: [redacted]
Assoc. Director, Clinical Development, SCD Statistics

Addendum: Statistical Analysis Plan Amendment 01 Addendum 01

I have read this addendum and approve.

Sivan Weiss
[redacted]

Approver: [redacted] **Date**

Director, Statistical Therapeutic Area Head of Neurology, Psychiatry and Biosimilar

[redacted]
Date & Time: 16 Dec 2020 13:53 +02:00

Approver: [redacted] **Date**

Vice President, Therapeutic Head, Neurology and Psychiatry, Global Specialty Clinical Development

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AIMS	Abnormal Involuntary Movement Scale
AR(1)	Autoregressive(1)
BARS	Barnes Akathisia Rating Scale
CAR	Censoring-At-Random
CDSS	Calgary Depression Scale for Schizophrenia
CGI-I	Clinical Global Impression–Improvement
CI	confidence interval
CNAR	Censoring-Not-At-Random
COVID-19	Coronavirus disease 2019
CSH	Heterogeneous Compound Symmetry
DAI-10	Drug Attitudes Inventory 10-item version
eITT	Extended Intent-to-treat
EoT	End of Treatment
ET	Early Termination
FU	Follow-up
HR	hazard ratio
ITT	Intent-to-treat
MI	Multiple Imputation
ML	Maximum-Likelihood
PANSS	Positive and Negative Syndrome Scale
q1m	once-monthly
q2m	once every 2 months
REML	Restricted Maximum Likelihood
SAP	Statistical Analysis Plan
SAS	Simpson Angus Scale

ADDENDUM 01 TO STATISTICAL ANALYSIS PLAN WITH AMENDMENT 01

Dated 22 April 2020

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Risperidone Extended-Release Injectable Suspension (TV-46000) for Subcutaneous Use as Maintenance Treatment in Adult and Adolescent Patients with Schizophrenia

Study TV46000-CNS-30072

The purpose of this statistical analysis plan (SAP) addendum is to outline the additional planned analyses to be completed for Study TV46000-CNS-30072 to support the analyses described in the SAP with Amendment 01 for Study TV46000-CNS-30072. After the approval of the SAP Am 01 (on 22 April 2020) and complying with the statements in it, this addendum provides a detailed description of the following:

(a) Tipping point and imputation details (Section 6.2.3.3 and Appendix A in the SAP Am 01)

(b) Analyses to evaluate the effect of Coronavirus disease 2019 (COVID-19) situation (Sections 6.2.5, 6.4)

(c) clarification of SAP language for clinical programming purposes.

A general guidance to the use of this document:

The reader of this SAP addendum is encouraged to read the study protocol for details on the conduct of this study, the operational aspects of clinical assessments, and the timing for completing the participation of a patient in this study, and the complete SAP for details regarding the planned data analyses. This document does not replace the SAP for this study in any manner. Its purpose is to clarify and provide further explanations and details regarding issues that were unknown or unresolved at the time of the design of the study, or during the issuance of the SAP amendment, such as the imputation model for handling suspected informative censoring, and the potential impact of COVID-19 on the study conduct, as well as a few additional minor issues that emerged at a later stage.

1. CLARIFICATIONS TO STATISTICAL ANALYSIS PLAN SECTIONS AND PROGRAMMING INSTRUCTIONS

This section describes minor changes, corrections and clarifications to the SAP. Where applicable, the clarifications and modifications to the respective sections are shown in underlined text. Only paragraphs that were modified or elaborated are depicted here; paragraphs that remained unchanged were not copied from the SAP.

1.1. Section 3.4 - Per-Protocol Analysis Set

The per-protocol (PP) analysis set will include all patients in the ITT analysis set who have no important protocol deviations that were deemed to be related to the primary analysis and to be excluded from the PP analysis; a case-by-case evaluation of which important protocol deviations deemed to be related to the primary endpoint will be conducted before unblinding.

In this analysis set, treatment will be assigned based on the treatment patients actually received. If patients erroneously received a wrong drug assignment in some visits, a decision to which treatment group they will be assigned will be determined in a case-by-case manner before unblinding. The PP analysis set will be discussed before unblinding and findings will be documented in the study data review document.

1.2. Section 4.4 - Study Days and Visits

For summary tables of safety, unscheduled visits will be mapped to the closest scheduled visit according to the allowed time window specified in Tables 1 and 2 of the study protocol. Unscheduled visits that occurred after Early Termination/End-of-Treatment visit will be mapped to follow-up visits. If both Early Termination/End-of-Treatment and follow-up 1 visit occurred on the same day and the Early Termination/End-of-Treatment results are missing, the results from the follow-up 1 visit will be mapped to the Early Termination/End-of-Treatment visit. For CGI-I, CGI-S and CGI-SS scales, unscheduled visits will be mapped as efficacy mapping.

For last safety assessment, the last available non-missing value will be used. For last efficacy assessment, the last available non-missing value not later than Early Termination/End-of-Treatment visit will be used. Unscheduled visits that could not be mapped will not be displayed in the by-visit summaries, but they will be considered for the last assessment visit. There will be 2 types of last assessment for the CGI-I, CGI-S and CGI-SS scales: (1) while on treatment, excluding FU and (2) last assessment including FU.

For summary tables of efficacy, unscheduled visits will be mapped to windows between the scheduled visits according to the date of the visit. The visit windows will be labeled 'Unsc visit after visit X' and will be presented sequentially in summaries with the scheduled visits.

1.3. Section 6.1 - General

The ITT analysis set (Section 3.2 in the SAP Am 01) will be used for efficacy summaries during Stage 2 of the study (double blind treatment stage) unless otherwise specified. In addition, summaries using the eITT subset of adolescents will not be presented, since only 1 adolescent

was enrolled and randomized. Summaries will be presented by treatment group, for overall TV46000 (i.e. including q1m + q2m), and for all patients, as applicable.

The pharmacokinetic analysis set will be used for the concentration data descriptive summary of risperidone, 9-OH-risperidone, and the total active moiety.

1.4. Section 6.2.3.2 - Interval Censoring

A sensitivity analysis will be conducted to assess the impact of large intervals between the previous assessment and the assessment when the first relapse was observed via the interval censoring method.

To avoid possible time bias in the right censoring rule, interval censoring rule will be performed as a sensitivity analysis for the primary endpoint (the primary endpoint of this study is the time to impending relapse, for further details see section 6.2 in the SAP Am 01). The time from randomization to impending relapse is unlikely to precisely coincide with the clinical visit, and thus will fall within an interval between 2 consecutive assessments with an unknown precise occurrence. This phenomenon is referred to as interval censoring. The analysis will define the left and right boundaries (relative to the first day of injection) of the time interval in the following way:

1. For patients who experienced an impending relapse based on criteria 2 or 4, as explained in Section 6.2.1 of the SAP Am 01, with the exact date of the event known, both left and right boundaries will be set to the event day and the interval length will be equal to 0. Otherwise, proceed to the next item.
2. For patients with impending relapse based on criteria 1 or 3, the left boundary will be defined as the last assessment with negative results (no relapse identified) immediately preceding the positive (relapse identified) assessment (including unscheduled visits), or the first day of injection if such an assessment does not exist. The right boundary of the time interval will be defined as the day of documented impending relapse.
3. For patients that were censored (ie, completed or withdrew early from the study without experiencing impending relapse), the left boundary will be set to the EoT/ET visit and the right boundary will be set to a missing value.

1.5. Section 6.2.4 - Sub-Group Analyses

Subgroup analysis will be performed for the primary efficacy endpoint according to the following categories, if applicable. The following analyses were slightly modified:

- Time from diagnosis (≤10 years, ≥10 years and ≤20 years, ≥20 years). Since the time from diagnosis in this study was found to be longer than anticipated, following the Independent Data Monitoring Committee (IDMC) request the categories were updated to more correctly reflect the patient characteristics and incorporate a better understanding of the data.
- COVID-19 status (pre, during and post COVID-19 pandemic outbreak, each patient will be classified into one of the levels), if data permit; However, since only 48 patients were affected by the COVID-19 pandemic, this sub-group analysis will not be performed.

1.6. Section 6.3.1 Key Secondary Efficacy Endpoints and Analyses

The eITT analysis set (Section 3.3 of the SAP Am 01) was planned to be used for all summaries in the key secondary endpoints. However, since only 1 adolescent patient was enrolled and randomized to the study, and since this will not impact the key secondary endpoints analyses, these analyses will be performed only on the ITT analysis set and not the eITT, as applicable.

1.7. Section 6.3.1.1 - Time to Impending Relapse in the eITT Analysis Set

The key secondary endpoint of time to impending relapse on eITT analysis set was planned to be analyzed similarly to the primary efficacy analysis, as described in Section 6.2.2 in the SAP Am 01, using the eITT analysis set. However, since only 1 adolescent was randomized in the study, the expected result will be highly correlated to the primary analysis (adding 1 patient, 1 relapse). Therefore, this analysis will not be part of the key secondary endpoints, and all key secondary endpoints that planned to be analyzed on eITT analysis set will be conducted on the ITT analysis set, as applicable.

1.8. Section 6.3.1.6.2 – DAI-10 Analysis

By visit change from baseline to each visit and endpoint in the DAI-10 total score will be summarized. Least squares means and 95% CI of change from baseline to endpoint, if applicable, will be performed using a repeated measures analysis of covariance (repeated measures ANCOVA) model with treatment arm (trt), subject id (subjid) and study visit (avisit), stratification variable (rand_strata) as factors and baseline DAI-10 score (Baseline) as a covariate. If for some visits there are less than 15 observations per arm, they will not be incorporated into the model, as convergence problems might occur. For these cases only summary statistics will be presented.

1.9. [redacted]

[redacted]

1.14. [redacted]

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]

[redacted]

2. [redacted]

[redacted]

Table 2: [redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	

Table 3: [redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]				[redacted]			[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]				[redacted]			[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]				[redacted]			[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
[redacted]	[redacted]				[redacted]			[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	

Table 4: [redacted]

[redacted]								
[redacted]								
[redacted]								
[redacted]								
[redacted]								
[redacted]								
[redacted]								
[redacted]								

[redacted]

Table 5: [redacted]

[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							
[redacted]							

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 6: [redacted]

[redacted]							
[redacted]							

2.2. [redacted]

[redacted]

2.2.1. [redacted]

[redacted]

[1] [redacted]

[redacted]

2.2.2. [redacted]

2.3. [redacted]

a. [redacted]

3. [redacted]

4. REFERENCES

Ballotpedia [Homepage]. States that issued lockdown and stay-at-home orders in response to the coronavirus (COVID-19) pandemic, 2020. Available at: https://ballotpedia.org/States_that_issued_lockdown_and_stay-at-home_orders_in_response_to_the_coronavirus (COVID-19) pandemic, 2020 [updated 09 November 2020; accessed 12 December 2020].

Bartlett Jonathan. London School of Hygiene & Tropical Medicine [Internet]. DIA working group, Imputation based approaches. Multiple imputation for time to event data under Kaplan-Meier, Cox or piecewise-exponential frameworks – SAS® macros [updated: 02 April 2019; cited 07 December 2020]. Available from: http://www.missingdata.org.uk/. The macros can be downloaded from: http://missingdata.lshtm.ac.uk/files/2019/04/Package_Release_V3-final.zip

In Your Pocket [Homepage]. The Coronavirus (covid-19) in Bulgaria: updates. Available at: https://www.inyourpocket.com/sofia/the-coronavirus-covid-19-in-bulgaria-updates_77571f [updated 26 November 2020; accessed 12 December 2020].

Lipkovich I, Ratitch B, O'Kelly M. Sensitivity to censored-at-random assumption in the analysis of time-to-event endpoints. Pharm Stat. 2016;15:216-29.Peto R, Pike MC, Armitage P, et al. "Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design". Br J Cancer. 1976;34(6):585–612.

Statistical Analysis Plan

Study TV46000-CNS-30072 with Protocol Amendment 03

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Risperidone Extended-Release Injectable Suspension (TV-46000) for Subcutaneous Use as Maintenance Treatment in Adult and Adolescent Patients with Schizophrenia

A Randomized, Double-Blind, Placebo-Controlled Study on Efficacy, Safety, and Tolerability of TV-46000 in Adults and Adolescents with Schizophrenia

A Study to Test if TV-46000 is Effective for Maintenance Treatment of Schizophrenia

Efficacy, Safety, and Tolerability Study (Phase 3)

IND number: 124384; NDA number: 213586;
EudraCT number: 2018-001619-65

Approval Date: 22 April 2020

Sponsor

Teva Branded Pharmaceutical
Products R&D, Inc.
145 Brandywine Parkway
West Chester, Pennsylvania 19380United States

Prepared by: [redacted]
Mgr Biostatistics, Teva Global Statistics
[redacted]
Assoc Dir, Clinical Development, SCD Statistics

STATISTICAL ANALYSIS PLAN APPROVAL

Study No.: TV46000-CNS-30072

Study Title: A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Risperidone Extended-Release Injectable Suspension (TV-46000) for Subcutaneous Use as Maintenance Treatment in Adult and Adolescent Patients with Schizophrenia

Statistical Analysis Plan for:

- ☐ Interim Analysis
- ☒ Final Analysis
- ☐ Integrated Summary of Efficacy
- ☐ Integrated Summary of Safety

Amendment: Statistical Analysis Plan with Amendment 01

Author: [redacted]
Assoc Dir, Clinical Development, SCD Statistics
[redacted]
Mgr Biostatistics, Teva Global Statistics

[redacted] 21 April, 2020

Approver: [redacted] **Date**
Director, Statistical Head of Neurology, Psychiatry, Biosimilar and Oncology

[redacted] 22-APR-2020

Approver: [redacted] **Date**
Vice President, Therapeutic Area Head, Neurology and Psychiatry, Specialty Clinical Development

TABLE OF CONTENTS

LIST OF TABLES

AMENDMENT HISTORY

The Statistical Analysis Plan for study TV46000-CNS-30072 (original study protocol dated 14 December 2017) has been amended and reissued as follows:

Amendment number	Date	Author(s)	Summary of changes	Reason for amendment
01	22 April 2020	[redacted] Mgr Biostatistics, Teva Global Statistics [redacted] Assoc Dir, Clinical Development, SCD Statistics	- Assumption for sample size calculation: Hazard ratio updated from 1.82 to 2.50. - Instead of 2 interim analyses, no interim analysis will be performed. - Study completion when at least 90 events are observed; - Updated enrollment projections - Clarification that if the total number of adolescent patients in the study will be less than or equal to 5, separate summary tables for adolescents and summary tables of the eITT analysis set will not be presented. - adding instruction to the least-squares mean calculations for cases when the number of observations per visit is small. - In Appendix A, update of the interval partition in the Bayesian piecewise exponential model to avoid convergence problems - adding subgroup analysis of before, during and after the COVID-19 pandemic outbreak - adding to a sensitivity analysis to evaluate the effect of the COVID-19	Removal of the interim analyses, and revision of the final statistical analysis of the study data to be conducted when at least 90 relapse events are observed. Inclusion of sensitivity and supplementary analyses to evaluate the impact of COVID-19 pandemic.

Amendment number	Date	Author(s)	Summary of changes	Reason for amendment
			pandemic on the treatment effect using Cox model with time-dependent covariate - adding a supplementary analysis analyzing the primary endpoint with multiple imputation for patient early terminated due to COVID-19 pandemic - Editorial changes were made throughout the document for the purpose of clarification.	
Original Statistical Analysis Plan	07 November 2019	[redacted] Senior Biostatistician, Teva Global Statistics [redacted] Assoc Dir, Clinical Development, SCD Statistics	Not applicable	Not applicable

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation	Term
AE	adverse event
AIMS	Abnormal Involuntary Movement Scale
ANCOVA	analysis of covariance
BARS	Barnes Akathisia Rating Scale
bpm	beats per minute
BMI	body mass index
CDSS	Calgary Depression Scale for Schizophrenia
CGI-I	Clinical Global Impression–Improvement
CGI-S	Clinical Global Impression of Severity
CGI-SS	Clinical Global Impression-Severity of Suicidality
CI	confidence interval
CMH	Cochran–Mantel–Haenszel
CNAR	Censoring-not-at-random
COVID-19	Coronavirus disease 2019
CRF	case report form
CSR	Clinical study report
C-SSRS	Columbia Suicide Severity Rating Scale
CYP2D6	cytochrome P450 2D6
DAI-10	Drug Attitudes Inventory 10-item version
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
ECG	Electrocardiogram/Electrocardiography
eITT	extended intent-to-treat
EoT	End-of-Treatment Visit
EQ-5D-5L	5-Level EuroQol Five Dimensions Questionnaire
ER	Emergency room
ET	Early Termination
GFR	Glomerular filtration rate
IDMC	Independent Data Monitoring Committee
ICF	Informed consent form
IMP	investigational medicinal product
INN	international nonproprietary name

Abbreviation	Term
IRT	Interactive Response Technology
ITT	Intent-to-Treat
MedDRA	Medical Dictionary for Regulatory Activities
msec	milliseconds
MI	Multiple imputation
MNAR	Missing-not-at-random
PANSS	Positive and Negative Syndrome Scale
PopPK	Population pharmacokinetics
PP	Per-Protocol
PSP	Personal and Social Performance Scale
PT	Preferred term
q1m	every month
q2m	every 2 months
QTcF	QT interval corrected using Fridericia's formula
R&D	Research and Development
SAS	Simpson-Angus Scale
SCID-5	Structured Clinical Interview for DSM-5
SD	Standard Deviation
SE	Standard Error
SI	standard international
SOC	system organ class
SOP	standard operating procedure
SQLS	Schizophrenia Quality of Life Scale
ULN	upper limit of normal
USA	United States of America
VAS	visual analogue scale
WHO	World Health Organization

INTRODUCTION

This Statistical Analysis Plan describes the planned analysis and reporting for Teva Branded Pharmaceutical Products R&D, Inc. study TV46000-CNS-30072 (A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Risperidone Extended-Release Injectable Suspension (TV-46000) for Subcutaneous Use as Maintenance Treatment in Adult and Adolescent Patients with Schizophrenia), and was written in accordance with GSD-SOP-704 (Final Statistical Analyses for Clinical Studies).

The reader of this Statistical Analysis Plan is encouraged to read the study protocol for details on the conduct of this study, the operational aspects of clinical assessments, and the timing for completing the participation of a patient in this study.

The Statistical Analysis Plan is intended to be in agreement with the protocol, especially with regards to the primary and key secondary endpoints and their respective analyses. However, the Statistical Analysis Plan may contain more details regarding these particular endpoints of interest, or other types of analyses (eg, other endpoints).When differences exist in descriptions or explanations provided in the study protocol and this Statistical Analysis Plan, the Statistical Analysis Plan prevails; the differences will be explained in the Clinical Study Report.

The following changes are described in this amendment:

1. Removal of the interim analyses and revision of the final statistical analysis of the study data to be conducted when at least 90 relapse events are observed,
2. Inclusion of sensitivity and supplementary analyses to evaluate the impact of COVID-19 pandemic.

1. STUDY OBJECTIVES AND ENDPOINTS

1.1. Primary and Secondary Study Objectives and Endpoints

The primary and secondary study objectives and endpoints are:

Objectives	Endpoints
The **primary objective** of this study is to evaluate the efficacy of TV-46000 during maintenance treatment in adult patients with schizophrenia.	The primary efficacy endpoint is time to impending relapse. Relapse is defined as 1 or more of the following items: • Clinical Global Impression–Improvement (CGI-I) of ≥5 (greater than or equal to minimally worse, ie, minimally worse, much worse or very much worse), **AND** – an increase of any of the following individual Positive and Negative Syndrome Scale (PANSS) items: conceptual disorganization, hallucinatory behavior, suspiciousness, and unusual thought content, to a score of >4 with an absolute increase of ≥2 on that specific item since randomization, **OR** – an increase in any of the following 4 individual PANSS items: conceptual disorganization, hallucinatory behavior, suspiciousness, and unusual thought content, to a score of >4 and an absolute increase of ≥4 on the combined score of these 4 PANSS items (conceptual disorganization, hallucinatory behavior, suspiciousness, and unusual thought content) since randomization • hospitalization due to worsening of psychotic symptoms (including partial hospitalization programs), excluding hospitalization for psychosocial reasons • Clinical Global Impression-Severity of Suicidality (CGI-SS) of 4 (severely suicidal) or 5 (attempted suicide) on Part 1 and/or 6 (much worse) or 7 (very much worse) on Part 2 • violent behaviour resulting in clinically significant self-injury, injury to another person, or property damage
The **key secondary objective** of this study is to evaluate the efficacy of TV-46000 during maintenance treatment in a total population (adults and adolescents) and in adolescent patients with schizophrenia.	Key secondary endpoints are: • time to impending relapse (as defined under the primary objective) in the total population (adults and adolescents) • impending relapse rate at week 24 • percentage of patients who maintain stability at endpoint • percentage of patients achieving remission at endpoint • observed rate of impending relapse at endpoint • Drug Attitudes Inventory 10-item version **(adult patients**

Objectives	Endpoints
	only) • Schizophrenia Quality of Life scale (SQLS) **(adult patients only)** • time to impending relapse in adolescent patients with schizophrenia
A **secondary objective** of this study is to evaluate the safety and tolerability of TV-46000 in the total population.	The safety variables include adverse events, extrapyramidal symptoms (EPS), risk of suicide events, depression symptoms, injection pain and other injection site reactions (local tolerability), vital signs, laboratory tests, physical examination, electrocardiogram (ECG) measurements, body weight, rescue medication use, time to all-cause discontinuation, all-cause discontinuation rates and discontinuation rates due to adverse events (tolerability), and the following rating scales: • Abnormal Involuntary Movement Scale (AIMS) • Simpson-Angus Scale • Barnes Akathisia Rating Scale • Columbia Suicide Severity Rating Scale (C-SSRS) • Calgary Depression Scale for Schizophrenia (CDSS) • CGI-SS
A **secondary objective** of this study is to evaluate the pharmacokinetics of oral risperidone and TV-46000 after administration of multiple doses in adults, adolescents, and the total population.	• The pharmacokinetic endpoints are the plasma concentrations of risperidone, 9-OH-risperidone, and total active moiety (sum of risperidone and 9-OH-risperidone).

1.2. [redacted]

Objectives	Endpoints
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Objectives	Endpoints
	[redacted]
[redacted]	[redacted]

2. STUDY DESIGN

2.1. General Design

This is a double-blind, randomized, relapse prevention study comparing a therapeutic dose of TV-46000 subcutaneous (sc) once-monthly (q1m) and every 2 months (q2m) with placebo sc (q1m) in a 1:1:1 ratio.

The duration of patient participation in the study will include up to 4 weeks of screening, 12 weeks of the oral conversion/stabilization stage (Stage 1), and a double-blind maintenance stage. The double-blind maintenance stage will be variable in duration, with patients continuing until they experience a relapse event; meet 1 or more of the study discontinuation or withdrawal criteria; or remain relapse-free during the double-blind phase until the study is terminated because at least 90 relapse events are recorded in the study adult population.

Patients that require a stabilization dose below 2 mg/day will not be randomized in the study. Also, as a precautionary measure, adolescent patients that will require a stabilization dose of more than 4 mg during the stabilization stage will not be randomized.

Patients in the TV-46000 treatment groups will receive a dose of TV-46000 (q1m or q2m) that is equivalent to the oral dose on which they were stabilized in Stage 1. The maximal dose administered to adult patients will be equivalent to an oral risperidone dose of 5 mg/day, and the maximal dose administered to adolescents will be equivalent to 4 mg/day. Patients randomized to TV-46000 q1m or placebo sc will receive a sc injection of TV-46000 or placebo, respectively, at baseline and every 4 weeks (q4w) thereafter. Patients randomized to TV-46000 q2m will receive a TV-46000 sc injection at baseline and every 8 weeks (q8w) thereafter, and a placebo sc injection 4 weeks after baseline and q8w thereafter, to ensure blinding of the doses and durations of the TV-46000 injections and the placebo injections.

Per definition, an exacerbation in symptoms during Stage 1 cannot be defined as a relapse event, since relapse events can only occur following stabilization and randomization. Randomized patients who relapse or meet 1 or more of the withdrawal criteria should be invited to perform the Early Termination (ET) visit as soon as possible, within 4 weeks of the last injection. Patients who remain relapse-free when at least 90 relapse events in adults are recorded in the study should be invited to perform the End-of-Treatment visit within 4 weeks of the last injection. Therefore, a patient is considered a study completer if he or she experienced impending relapse or remained relapse-free at the time of study termination.

End of study is defined as the date when the last patient in Stage 2 has completed all efficacy and safety assessments at the final visit per protocol.

Study procedures and assessments with their timing for the pre-treatment period (screening and Stage 1) are summarized in Table 1 of the study protocol. Study procedures and assessments starting from baseline, through the double-blind maintenance stage (Stage 2), end of treatment (EoT)/early termination (ET) and follow-up are summarized in Table 2 of the study protocol.

In the event of an emergency situation (eg, the Coronavirus disease 2019 [COVID-19] pandemic), in case a patient cannot return to the clinic for the scheduled visits (eg, due to quarantine, isolation, patient's concern, or closure of the site clinic), remote assessment of

efficacy and safety scales via TC and/or videoconference (VC), with VC being the preferred method, may be allowed. The results of the scale rating will be directly entered into the eCRF per the usual process.

During the conduct of this study, an Independent Data Monitoring Committee (IDMC) will review accumulating unblinded safety and PK data on a regular basis to ensure the continuing safety of the study patients and study conduct issues (see Section 2.3). There will be no interim analysis in this study.

2.2. Randomization and Blinding

Patients will be randomized to receive doses of TV-46000 q1m sc injection, TV-46000 q2m sc injection, or placebo q1m sc injection in a 1:1:1 ratio. Randomization will be stratified by gender (male or female) and the dose of oral risperidone on which the patient was stabilized during Stage 1 (2/3 mg, 4 mg, or for adults only 5 mg). The doses of TV-46000 will be equivalent to 2 to 5 mg/day of oral risperidone (equivalent to the oral dose on which the patient was stabilized in Stage 1). Patients randomized to TV-46000 q1m or placebo sc will receive sc injection of TV-46000 or placebo, respectively, at baseline and q4w thereafter. Patients randomized to TV-46000 q2m will receive a TV-46000 injection at baseline and q8w thereafter and a placebo sc injection 4 weeks after baseline and q8w thereafter to ensure blinding of the doses and durations of the TV-46000 injections and the placebo injections.

Patients and investigators will remain blinded to the identity of the treatment administered to each patient. Due to the differences between the TV-46000 product and placebo, an unblinded nurse, not associated with rating the patient (including assessment of the injection site if needed) and independent from the study team, will be required at each site to administer the study drug.

Additional measures to mitigate the risk of unblinding are described in Section 5.9 of the study protocol.

Patients will be randomly assigned to treatment groups by means of a computer-generated randomization list. The randomization list will be assigned to the relevant treatment groups through a qualified service provider (eg, via the Interactive Response Technology [IRT] system). The generation of the randomization list and management of the IRT system will be done by a qualified service provider under the oversight of the responsible function at Teva. The specifications for randomization will be under the responsibility and oversight of Teva Global Statistics.

The sponsor's clinical personnel (and delegates) involved in the study will be blinded to the identity of the investigational medicinal product (IMP) until the database is locked for final analysis, and the IMP assignment is known. However, if a prioritized sample analysis is needed, bioanalytical and IDMC clinical pharmacology representatives are unblinded.

In the event of an emergency (ie, where knowledge of the study drug assignment is needed to make treatment decisions for the patient), the treatment group and dose to which the patient has been allocated can be determined by accessing the IRT system. All investigational centers will be provided with details on how to access the system for code breaking at the start of the study. The Medical Monitor or equivalent should be notified following unblinding. Any unblinding of the IMP performed by the investigator must be recorded in the source documents.

The pharmacist at each investigational center who will dispense the IMPs will know the IMP assignments for each patient. In addition, up to 2 other individuals from each investigational center will know the IMP assignments to provide quality assurance and oversight in their preparation and administration, as necessary. These individuals will not be involved in conduct of any study procedures and assessment of any adverse events and should not discuss the medication assignment or appearance with others (except for the unblinded nurse administering the medication).

Additional details regarding maintenance of randomization and blinding can be found in Section 5.10 of the study protocol.

2.3. Data Monitoring Committee

During the conduct of this study there will be an IDMC that will review accumulating unblinded safety and pharmacokinetic data on a regular basis (as detailed in the IDMC charter) to ensure the continuing safety of the study patients and any study conduct issues.

The IDMC will be composed of independent physicians with expertise in the relevant therapeutic field, a statistician and a pharmacokinetic specialist.

The IDMC chairperson will communicate with the sponsor in regard to issues resulting from the conduct and clinical aspects of the study. The sponsor will work closely with the committee to provide the necessary data for review.

The IDMC will provide recommendations about modifying, stopping, or continuing the study. The conduct and specific details regarding the IDMC sessions will be outlined in the IDMC charter.

2.4. Sample Size and Power Considerations

In order to calculate the sample size, the following assumptions were used:

- a placebo effect (namely, median time to impending relapse of 7 months in the placebo group) similar to that observed in a similarly-designed study (Kane et al 2012);
- a hazard ratio of 2.50 (placebo vs each TV-46000 arm);
- a randomization ratio of 1:1:1 (q1m:q2m:placebo), with 2 primary hypotheses to be tested (q1m vs placebo and q2m vs placebo) at a 2-sided alpha of 0.050.

To control for the family-wise type-I error (multiplicity) of the primary endpoint family, a fixed sequential (hierarchical) testing procedure will be used. The total number of events was calculated to attain a statistical power of at least 90% in both comparisons of q1m and q2m versus placebo.

Assuming the above, a total of at least 90 relapse events will need to be observed during Stage 2 of the study in the ITT analysis set (adult patients) in the 3 treatment groups (combined) in order to attain at least 90% power to meet success criteria in both comparisons (East 6 [Version 6.3] manual, 2014). The sample size rationale is based on the adult patients. There is no estimation of the sample size in the adolescent population, and their number is not known at this time.

Assuming an approximate accrual time of 18 months and a maximal treatment duration during the double blind stage of approximately 24 months, approximately 173 adult patients will need to be randomized to each treatment group for a total of approximately 520 adult patients randomized. Assuming that 40% of the patients enrolled in Stage 1 will not be randomized to the double-blind phase (Stage 2), a total of approximately 860 adult patients will need to be enrolled into Stage 1. As an event-driven study, depending on the actual recruitment rate and percentage of patients who are enrolled in Stage 1 and are randomized to Stage 2, it may be possible to randomize more than 520 adult patients, as long as Stage 2 of the study ends when the number of relapse events in the ITT analysis set reaches at least 90.

2.5. Sequence of Planned Analyses

All analyses identified in this Statistical Analysis Plan will be performed after the end of study as defined in the study protocol.

This Statistical Analysis Plan and any corresponding amendments will be approved before database lock, in accordance with GSD-SOP-702: Statistical Analysis Plan (SAP).

The randomization codes for the final analysis will not be unblinded until this Statistical Analysis Plan has been approved and issued and the database has been locked.

Any results such as exploratory analyses completed to support study analyses, which were not identified in this Statistical Analysis Plan, may be documented in the clinical study report (CSR).

3. ANALYSIS SETS

3.1. Enrolled Patients Set

The enrolled patients set will include all patients who have met study eligibility requirements for Stage 1 and received oral risperidone.

3.2. Intent-to-Treat Analysis Set

The intent-to-treat (ITT) analysis set will include adult patients randomized to the double-blind maintenance stage (Stage 2), regardless if they have received treatment or not. In the ITT analysis set, treatment will be assigned based on the treatment to which patients were randomized, regardless of which treatment they actually received.

The purpose of the primary analyses is to estimate the effect of randomized treatment only, omitting any effect of additional treatment that may be administered *after* discontinuation of the randomized treatment. Accordingly, the ITT analysis set complies with the while-on-treatment strategy[1] that will be undertaken with regard to the intercurrent event of treatment discontinuation, with censoring at this intercurrent event (unless the treatment discontinuation is due to relapse). This study estimand will be the difference in time to relapse (survival) under the treatment to which the patient was initially randomized until his/her last treatment or early termination of all adult patients that were successfully stabilized on oral risperidone on daily dose rage of 2 mg to 5 mg.

3.3. Extended Intent-to-Treat Analysis Set

The extended intent-to-treat (eITT) analysis set will include all patients (adults and adolescents) randomized to the double-blind maintenance stage (Stage 2), regardless if they have received treatment or not. In the eITT analysis set, treatment will be assigned based on the treatment to which patients were randomized, regardless of which treatment they actually received.

3.4. Per-Protocol Analysis Set

The per-protocol (PP) analysis set will include all patients in the ITT analysis set who have no important protocol deviations (see Section 5.7 for definition). In this analysis set, treatment will be assigned based on the treatment patients actually received. If patients erroneously received a wrong drug assignment in some visits, a decision to which treatment group they will be assigned will be determined in a case-by-case manner before unblinding. The PP analysis set will be discussed before unblinding and findings will be documented in the study data review document.

3.5. Safety Analysis Set

The safety analysis set will include all randomized patients who receive ≥1 dose of study treatment or placebo in the double-blind maintenance stage (Stage 2). In the safety analysis set,

[1] ICH E9 (R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials. EMA/CHMP/ICH/436221/2017. European Medicines Agency. 30 August 2017

treatment will be assigned based on the treatment patients actually received, regardless of the treatment to which they were randomized, unless otherwise specified. If patients erroneously received a wrong drug assignment in some visits, a decision to which treatment group they will be assigned will be determined in a case-by-case manner before unblinding.

3.6. Pharmacokinetics Analysis Set

The pharmacokinetics analysis set will include all patients from the safety analysis set who also have ≥1 plasma concentration measured.

4. GENERAL ISSUES FOR DATA ANALYSIS

4.1. General

Descriptive statistics for continuous variables include n, mean, standard deviation (SD), standard error (SE), median, minimum, and maximum. Descriptive statistics for categorical variables include patient counts and percentages; a "missing" category will be displayed as appropriate.

For presentations by oral risperidone dose, the patient's last dose up to and including visit 5 will be considered the oral risperidone dose

If the total number of adolescent patients in the study will be less than or equal to 5, separate summary tables for adolescents and summary tables of the eITT analysis set (Section 3.3) will not be presented. For the following sections, the presentation for eITT will be replaced with presentation for ITT, in case of less/equal than 5 adolescents.

All tables will be presented separately for adult and adolescent populations. However, the adolescent population summaries will be presented only if their final number is large enough according to the clinical judgement. In this case, the adolescent data will only be listed.

4.2. Specification of Baseline Values

The baseline value of Stage 1 (Oral Conversion and Stabilization Stage), is the last observed value before the first oral risperidone administration as part of the conversion and stabilization stage, unless otherwise noted.

For the double-blind maintenance stage (Stage 2: Relapse Prevention), baseline is defined as the last observed data before the first study drug injection. However, for efficacy endpoints, if the patient was already injected and some of the efficacy data was collected on the same day but after the injection, then this data will still be used as baseline.

4.3. Handling Withdrawals and Missing Data

Despite the best efforts to obtain complete data, missing data is unavoidable. For all variables, only observed patient data will be used in the statistical analyses, ie, there is no plan to estimate missing data, unless otherwise specified. The only estimation of missing data is planned for the sensitivity analysis for the primary endpoint, using the tipping point approach (see Section 15 and Appendix A), with the exception outlined below. Data from patients who did not relapse will be censored at the last valid relapse assessment date.

Ad-hoc imputation for safety summary tables will be performed only for the specific cases described below. There will be no imputation in the data listings (including ad-hoc imputation); all values will be displayed as recorded in the clinical database.

4.3.1. Ad-hoc Imputation for Safety Data

Any adverse event with unknown severity will be considered as 'severe' for the tabulations. During Stage 2 of the study, for any adverse event with unknown relationship, the treatment will be considered as 'related'.

Any laboratory values given as '<x' or '>x' in the database will be assigned with the absolute value of x without the sign for the descriptive statistics and the calculation of changes from baseline (eg, a value of <0 .1 will be imputed as 0.1 for the calculations).

4.3.2. Missing items in the Quality of Life Questionnaires

Missing items will not be imputed. However, for the SQLS, CDSS, AIMS, SAS and BARS questionnaires, if the number of missing items in a scale (or sub-scale) is less than half, then the missing items will be assigned with the value of the arithmetic mean of the non missing items in the scale within the patient and visit.

4.3.3. Handling of Adverse Events with Missing Dates

The date of First Study Treatment, referred to below, corresponds to the first injection of TV-46000 (IMP) that the patient received. For the purpose of this study, oral risperidone is not considered an IMP. Any adverse event with incomplete start dates will be handled as described below, end dates will not be imputed.

- If only the day is missing, then the day will be imputed to the first day of the month, unless the adverse event month-year corresponds to the month-year of the Date of First Study Treatment. In that case, the day will be imputed with the Date of First Study Treatment, unless the adverse event end date occurred prior to Date of First Study Treatment.
- If both the day and month are missing and the year is the year of the Date of First Study Treatment then the date will be imputed to Date of First Study Treatment, unless the adverse event end date occurred prior to Date of First Study Treatment, in which case the start day-month will not be imputed. In addition, if the Date of First Study Treatment year < year of ADVERSE EVENTS start year then start day-month will be imputed to 01JAN.

4.3.4. Handling of Concomitant Medication with Missing Dates

For determination of prior vs concomitant medication, missing start or end month are imputed to JAN or DEC respectively and missing start or end day are imputed to day 01 and 30 respectively, unless the month is not missing then missing day is imputed to 28.

4.4. Study Days and Visits

Study days are numbered relative to the first day of study drug administration, namely the beginning of Stage 2. The start of treatment (Day 1) is defined as the date on which a patient receives the first dose of study drug, as recorded on the Case Report Form (CRF). Days will be numbered relative to treatment start (ie, ..., –2, –1, 1, 2, ...; with Day 1 being the first day of study drug administration and day –1 being the day before the first day of study drug administration).

For summary tables of safety, unscheduled visits will be mapped to the closest scheduled visit according to the allowed time window specified in Tables 1 and 2 of the study protocol. Unscheduled visits that occurred after Early Termination/End-of-Treatment visit will be mapped to follow-up visits. If both Early Termination/End-of-Treatment and follow-up 1 visits occurred

on the same day and the Early Termination/End-of-Treatment results are missing, the results from the follow-up 1 visit will be mapped to the Early Termination/End-of-Treatment visit, and vice versa.

For by-visit summaries, if there are multiple assessments at a post-baseline visit day, then the last non-missing assessment at that visit day will be used for the summary. If the multiple assessments include a schedule assessment, the scheduled will be used.

For last safety assessment, the last available non-missing value will be used. For last efficacy assessment, the last available non-missing value not later than Early Termination/End-of-Treatment visit will be used. Unscheduled visits that could not be mapped will not be displayed in the by-visit summaries, but they will be considered for the endpoint/last assessment visit.

5. STUDY POPULATION

5.1. General

The eITT analysis set (Section 3.3) will be used for all study population summaries of the double-blind treatment stage unless otherwise specified. Summaries will be presented by treatment group and for all patients. The primary analysis will be conducted on the ITT analysis set, and the key secondary analyses will be conducted on the eITT analysis set (Section 3.3), unless otherwise specified. For the endpoints that are evaluated only in the adult population, the analysis will be conducted on the ITT analysis set. The analysis set on which analysis for other end-points will be conducted will be specified on a case-by-case basis. The enrolled patients set will be used for data summaries before the double-blind treatment stage.

It is intended to also present summary tables by age subgroups (adults/adolescents) as applicable.

However since the number of adolescent patients may be too sparse (below 5 randomized adolescents), these summary tables might not be presented.

In addition, and if data permit, COVID-19 Subgroups (eg., COVID-19-impacted patients/ COVID-19-non-impacted patients and/or home visits etc.) of the eITT analysis set will be presented.

5.2. Patient Disposition

The following will be summarized using descriptive statistics of data for

- patients screened,
- screened and not enrolled in the Oral Conversion and Stabilization Stage (Stage 1) and reason for screen failure/not enrolled,
- screened and enrolled in Stage 1; patients who were enrolled in Stage 1 but not randomized for the double blind stage (Stage 2) and reason for not randomized;
- patients who were randomized;
- patients who were randomized but not treated;
- patients in the ITT, or adolescents in the eITT , PP, safety, and pharmacokinetics analysis sets;
- patients who completed the treatment and/or study;
- patients who withdrew from the treatment or the study and the reason for withdrawal, including due to COVID-19.

Adolescent patient disposition may be summarized separately, if applicable.

5.3. Demographics and Baseline Characteristics

Patient demographic and baseline characteristics, including age, age group (adolescents [ages 13-17] and adults [18 years of age and above]), gender, race, ethnicity, baseline weight (kg),

baseline height (cm), baseline body mass index (kg/m^2) and categorization, and injection site will be summarized using the descriptive statistics.

The randomization stratification factors levels (see Section 2.2), will be summarized using the descriptive statistics for each variable category.

Psychiatric history will be assessed using months since the initial diagnosis of schizophrenia and months since the most recent relapse.

The ITT analysis set, the eITT analysis set, and the eITT subset of adolescents will be used for the summaries.

In addition, some of the demographic and screening characteristics will be summarized using the descriptive statistics for the enrolled patients set.

5.4. Medical History

All medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The incidence of medical history abnormalities will be summarized using descriptive statistics by system organ class (SOC) and preferred term (PT). Patients are counted only once in each PT and SOC category. Summaries will be presented by treatment group and for all patients.

5.5. Prior Therapy and Medication

Any prior therapy, medication, or procedure a patient has had within 30 days before study drug administration will be recorded on the CRF. Trade name or international nonproprietary name (INN), indication, and dosage will be recorded. The sponsor will encode all therapy and medication according to the World Health Organization (WHO) drug dictionary (WHO Drug).

The incidence of prior therapies and medications will be summarized using descriptive statistics by therapeutic class and PT. Patients are counted only once in each therapeutic class category, and only once in each PT category. Prior therapies and medications will include all medications taken and therapies administered before the first day of study drug administration. This will be collected twice during the study: at screening (before Stage 1) and Stage 2 (at baseline) of the study. Note that the prior therapy and medication for Stage 2 will include the concomitant medications in Stage 1.

5.5.1. Pre Study Risperidone Exposure

Pre-study risperidone treatment use will be defined based on patients' or investigator’s report on the first and last days of pre-study risperidone treatment. Summary of the number and percentage of patients exposed to risperidone will be summarized by treatment and overall.

5.6. Childbearing Potential and Methods of Contraception

For female patients, information related to childbearing potential and menopause will be collected at each visit in both study periods (see Section 8.6.2.2). Data will be listed.

For female and male patients, methods of contraception will be collected at screening. Data will be listed.

5.7. Study Protocol Deviations

Data from patients with any important protocol deviations as defined in Appendix C of the study protocol, and as recorded in protocol deviation CRF during the study will be summarized overall and for each category using descriptive statistics.

Data pertaining to deviations will be reviewed and important protocol deviations will be determined by the study team on ongoing basis and documented in the Statistical Data Review meeting minutes prior to the unblinding of the treatment codes.

Following unblinding and the subsequent availability of additional results, such as pharmacokinetic data and prolactin results, patients that are identified as non-compliant with regard to the study medication may be removed from the PP analysis set. This includes, for example, placebo patients taking oral risperidone or prohibited neuroleptics during Stage 2.

6. EFFICACY ANALYSIS

6.1. General

The ITT analysis set (Section 3.2) will be used for efficacy summaries during Stage 2 of the study (double blind treatment stage) unless otherwise specified. In addition, summaries using the eITT subset of adolescents will be presented, where relevant. Summaries will be presented by treatment group and for all patients.

The enrolled patients analysis set (Section 3.1) will be used for data summaries during Stage 1 of the study; summaries will be presented for all patients in this set, and by the ITT or eITT as appropriate. Unless otherwise specified, for the enrolled analysis set, data will be presented by randomized, not-randomized patients and overall, as appropriate.

The pharmacokinetic analysis set will be used for all pharmacokinetic/pharmacodynamics analyses which will be presented in a separate statistical analysis plan and report.

The PP analysis set will be used for the sensitivity of the primary efficacy variable.

In the event of an emergency situation (eg, the COVID-19 pandemic), in case a patient cannot return to the clinic for the scheduled visits (eg, due to quarantine, isolation, patient's concern, or closure of the site clinic), remote assessment of efficacy scales via TC and/or VC, with VC being the preferred method, may be allowed. The results of the scale rating will be directly entered into the eCRF per the usual process.

These measures will be implemented on a case-by-case basis, and only when and where they are warranted due to the emergency situation. Preferably, the original protocol instructions will be followed whenever the new instructions are not required.

6.2. Primary Efficacy Endpoint and Analysis

6.2.1. Definition

The primary efficacy endpoint is time to impending relapse. Patients meeting any 1 or more of the below impending relapse criteria are considered relapsed. Relapse is defined as 1 or more of the following items:

1. CGI-I of ≥5 (greater than or equal to minimally worse, ie, minimally worse, much worse or very much worse) **AND**

    a) an increase of any of the following individual PANSS items: conceptual disorganization, hallucinatory behavior, suspiciousness, and unusual thought content, to a score of >4 with an absolute increase of ≥2 on that specific item since randomization, **OR**

    b) an increase in any of the following 4 individual PANSS items: conceptual disorganization, hallucinatory behavior, suspiciousness, and unusual thought

content, to a score of >4 and an absolute increase[2] of ≥4 on the combined score of these 4 PANSS items (conceptual disorganization, hallucinatory behavior, suspiciousness, and unusual thought content) since randomization.

2. hospitalization due to worsening of psychotic symptoms (including partial hospitalization programs), excluding hospitalization for psychosocial reasons
3. CGI-SS of 4 (severely suicidal) or 5 (attempted suicide) on Part 1 and/or 6 (much worse) or 7 (very much worse) on Part 2
4. violent behavior resulting in clinically significant self-injury, injury to another person, or property damage.

For simplicity, the assigned numbers 1-4 of the aforementioned impending relapse criteria will be used henceforth.

It should be noted that criteria 2 and 4 could happen anytime during the study without regard to study visits, whereas criteria 1 and 3 are evaluated at the study in-clinic visits (including unscheduled visits) and thus could be recorded simultaneously at the same visit.

Per definition, an exacerbation in symptoms during Stage 1 cannot be defined as a relapse event, since relapse events can only occur following stabilization and randomization. Therefore, for efficacy evaluation, impending relapse can occur from the first injection until the end of treatment (ie, End of Treatment (EoT)/Early Termination (ET) Visit)[3]. In the follow-up time window, patients will be treated according to the investigator's judgement and may not comply with the study protocol, and therefore will not be counted for impending relapse.

6.2.2. Primary Efficacy Analysis

Time to impending relapse will be calculated as the earliest date the patient meets ≥1 of the impending relapse criteria minus the randomization date plus 1. Data from patients who did not relapse will be censored at the EoT/ET visit. Time to impending relapse for TV-46000 q1m and q2m versus placebo will be compared using the stratified log rank test at significance levels described in Section 7.

[redacted]

[redacted]

[redacted]

The stratified Cox proportional hazard model will be used as supportive evidence and for quantification of treatment effect, unless the constant hazard ratio assumption of the treatment is not justified. The hazard ratio proportionality of the 3 treatment groups will be tested using the Wald test for linear hypothesis testing and may be plotted over time to explore the treatment effect. A sample SAS® code which could be employed for this assessment is

```
proc phreg [redacted]
```

[redacted]

6.2.3. Sensitivity Analysis

6.2.3.1. Sensitivity Analysis of The Primary Endpoint Using The Per Protocol Analysis Set

The primary analysis will be repeated using the PP analysis set to assess the robustness of the principal analysis results in the subpopulation of patients that completed the study without important protocol deviations.

6.2.3.2. Interval Censoring

A sensitivity analysis will be conducted to assess the impact of large intervals between the previous assessment and the assessment when the first relapse was observed via the interval censoring method.

To avoid possible time bias in right censoring rule, interval censoring rule will be performed as a sensitivity analysis for the primary endpoint. The time from randomization to impending relapse is unlikely to precisely coincide with the clinical visit, and thus will fall within an interval between 2 consecutive assessments with unknown precise occurrence. This phenomenon is referred as interval-censoring. The analysis will define the left and right boundaries (relative to first day of injection) of the time interval in the following way:

1. For patients who experienced an impending relapse based on criteria 2 or 4, as explained in Section 6.2.1, with the exact date of the event known, both left and right boundaries will be set to the event day and the interval length will be equal to 0. Otherwise, proceed to the next item.
2. For patients with impending relapse based on criteria 1 or 3, the left boundary will be defined as the last assessment with negative results (no relapse identified) immediately preceding the positive (relapse identified) assessment (including unscheduled visits), or the first day of injection if such an assessment does not exist. The right boundary of the time interval will be defined as the day of documented impending relapse.
3. For patients that were censored (ie, completed or withdrew early from the study without experiencing impending relapse), the left boundary will be set to day of the last efficacy assessment and the right will be set to a missing value.

Frequency and percentage of the 3 types of interval censoring will be presented. Both TV-46000 q1m vs placebo and TV-46000 q2m vs placebo will be compared using a generalized log-rank test. Estimated survival distribution (Kaplan-Meier survival curve that accounts for interval censoring) will be presented graphically.

[redacted].

6.2.3.3. Sensitivity Analysis of the Primary Efficacy Analysis due to Potential Informative Censoring Using a 'Tipping Point' Approach

In time-to-event endpoints, subjects that discontinue the study prematurely are censored, so that they contribute to the number of subjects at risk of an event until the date-time of their censoring. The underlying assumption of most commonly used time-to-event methods, such as the Cox proportional hazard model and log-rank test, is that censoring is non-informative (also known as 'ignorable'), which means that the censoring is not related to the risk for relapse. This assumption complies with the while-on-treatment estimand for the primary efficacy estimand of relapse prevention (see Section 3.2). However, some of the reasons for early dropouts may potentially be related to the risk for relapse. Informative censoring (also known as non-ignorable censoring, or more specifically, censored not-at-random (CNAR)) can lead to similar biases raised in missing data problems under a missing not-at-random (MNAR) mechanism. Since the assumption of non-informative censoring cannot be verified based on observed data and may not be clinically plausible for all subjects, the following methodology in the general framework of pattern-mixture models (PPMs) will be applied to assess the sensitivity of the observed results to this assumption (see Lipkovich et al, 2016).

Patients with suspected informative censoring at withdrawal are defined as patients who discontinue the study prematurely for selected 'informative' reasons listed in Table 1 AND an increase in the PANSS score at the ET visit. Patients who withdrew for informative reasons in Table 1 that have a missing value in the PANSS score for any reason (eg, incomplete psychiatric evaluation or missing ET visit) are also considered suspected informative censoring.

Time to event values will be assigned to patients with suspected informative censoring (as defined above) using multiple imputation methodology. These patients may have an increased event hazard starting from time of discontinuation, compared to similar patients on the same treatment who remained in the study at this time-point or discontinued patients considered to be non-informative censoring.

In the suggested framework, and to evaluate the robustness of the results, it is assumed that after suspected informative censoring withdrawn patients have an increased hazard compared to the hazard of 'similar' non-informative censored or on-going patients by the time of withdrawal in the same treatment arm. This can be done for the placebo arm, and the active treatment arms, by defining a shift in hazard, δ, that quantifies the excessive risk after withdrawal compared to a similar patient in *the same arm*. For example, if $\delta=2$ and patient was censored at time t for

reasons listed in Table 1 with a missing ET visit (ie. meets the suspected informative censoring criteria), then this patient has a risk of an event after time t that is *twice* (δ=2) the hazard of event for a similar patient in the same treatment arm that was not censored at time t with the condition described above.

Two-dimensional tipping point analysis will be conducted at the final analysis. The tipping-point analysis strategy will be conducted as follows:

Once the efficacy criteria is met at the final analysis planned at the end of the study (when a total of at least 90 relapse cases are reached), the 2-dimensional tipping point analysis will be performed in 2 steps:

i. utilize 2 shift parameters $\delta_{placebo}$ for the placebo arm and δ_{q1m}, for the q1m active treatment arm

ii. utilize 2 shift parameters $\delta_{placebo}$ for the placebo arm and δ_{q2m}, for the q2m active treatment arm . It should be noted that in case the primary efficacy endpoint of q2m versus placebo does not reach significance (only q1m versus placebo reached significance) this analysis may be performed as well with appropriate changes to the shift parameters in order to better evaluate the reasons for the failure in this comparison.

$\delta_{placebo}$ and δ_{q1m} or δ_{q2m} (as appropriate) are gradually increased until the tipping point is reached, where success criteria no longer holds (ie, the statistical test in favor of a treatment arm is no longer significant) **or** shift parameters exceed 5 for all treatment arms. Additional points may be added at the tipping point neighborhood in order to provide high resolution presentation of the results. The two dimensional tipping point will be presented in a heat-plot, with tipping-point contours (if applicable).

Once the study reaches at least 90 relapse cases, at the final analysis, the criteria follows a sequential procedure: q1m is tested first at 0.05 alpha level and if significant, is followed by a test for q2m at the same significance level. A list of the pairs of p-values for all delta pairs will be presented, in addition to the tipping point heat-plot.

In case the primary analysis results are not significant, the tipping point analysis may be conducted to evaluate the effect of informative censoring on the insignificant results. To do so, the range of δ values would be modified as appropriate.

The statistical test for each δ will be computed using multiple imputations relying on Lipkovich et al, 2016. A detailed description of the implementation of the methodology for this study is provided in Appendix A.

Table 1: Early Dropout Reasons Classified as Potentially Associated with Informative Censoring[a]

Situation/Reason for early treatment termination	Censoring Classification
Death [b]	Non-informative
AE related to mental status change or psychiatric symptoms change	Informative
AE NOT related to mental status change or psychiatric symptoms change	Non-informative
Withdrawal By Subject [c]	Informative
Protocol Deviation [b]	Informative
Pregnancy	Non-informative
Lost To Follow-Up	Informative
Lack Of Efficacy	Informative
Administrative reason or Study termination due to sponsor decision	Non-informative
Other[b]	Informative
COVID-19 reasons[d]	Non-informative

[a] Potential informative censoring requires a combination of both (i) ET for the reason listed in this table and (ii) an increase or missing value in PANSS score at ET visit compared to baseline.
[b] To be discussed during the blinded Statistical Data Review meeting on a case-by-case bases prior to an Interim and Final Analyses.
[c] Excluding clear cases of withdrawal due to confirmed logistical cases (eg, moving to other state). This will be decided and documented during the blinded Statistical Data Review meeting.
[d] In general cases of ET in case a patient cannot return to the clinic for the scheduled visits (eg, due to quarantine, isolation, patient's concern, or closure of the site clinic) are regarded as missing at random (MAR) since the ET is not because of an increased risk of the patient to experience relapse. However, every case will be discussed during the blinded Statistical Data Review meeting prior to the Final Analysis to assign Informative status to relevant cases.
AE = adverse event, ET = early termination; COVID-19 = Coronavirus disease 2019

Death cases that were classified as informative at the Statistical Data Review meeting will be assigned as relapsed patient with date of relapse equal to death date in this sensitivity analysis.

Additional sensitivity analyses utilizing tipping point approach may be conducted as appropriate.

6.2.3.4. Multiple Imputation for COVID-19 Early Termination Patients

If the number of early termination cases will increase substantially as a result of the COVID19 pandemic outbreak and data permit, a supplementary analysis considering dropouts due to the pandemic as if these patients continued the study in spite of the COVID-19 pandemic outbreak (estimand using hypothetical strategy to address this intercurrent event) may be performed. Thus, censoring due to COVID-19 (patient cannot return to the clinic for the scheduled visits (eg, due to quarantine, isolation, patient's concern, or closure of the site clinic)) will be regarded as an intercurrent event that will be multiple imputed by treatment group assuming MAR. This analysis may be performed if data allows. A final decision to perform this analysis will be made prior to unblinding and details will be delineated in the Statistical Analysis Plan addendum as applicable, as applicable.

6.2.4. Sub-Group Analyses

Subgroup analysis will be performed for the primary efficacy endpoint according to the following categories, if applicable:

- Gender;
- Race – in case of at least 15 patients per level-group;
- USA vs outside of USA;
- Time from diagnosis (<3 years, ≥3 years and <5 years, ≥5 years)
- Baseline PANSS (by quartiles);
- COVID-19 status (pre, during and post COVID-19 pandemic outbreak, each patient will be classified into one of the levels), if data permit; details will be described in the SAP addendum as applicable.

The effect and its 2-sided 95% CI will be estimated using Cox proportional hazard model and summarized in a forest plot. [redacted].

6.2.5. COVID-19: Time-Dependent Cox Model

If applicable, a time dependent Cox model will be fitted, incorporating the model in Section 6.2.2 a time-dependent indicator of the COVID-19 pandemic outbreak. This sensitivity analysis (the same estimand as of the primary efficacy estimand) is conducted to evaluate the impact of the COVID-19 pandemic on the treatment effect. An exact definition of this time dependent variable will be defined in a SAP addendum before unblinding, since this definition depends on data recorded and the ability to manipulate and use this data. A SAS code for this model will be provided in a SAP addendum.

All details regarding the definitions of COVID19 time dependent indicator, Start and Stop windows will be found in a SAP addendum, if this analysis is deemed applicable to the data. The final decision based on the applicability will be taken during the blinded Statistical Data Review meeting, or before unblinding.

6.2.6. Supplementary Analysis - Subimpending Relapse

For this supplementary analysis, the impending criteria are modified to define a subimpending event (see Parfionovas et al 2011) is a supplementary analysis to assess the robustness of analysis results from the primary efficacy endpoint; This definition is intended to add a relapse case for some censored patients (as defined in the primary analysis endpoint) who at the time of early termination were close to meeting the impending relapse criteria of the primary endpoint. Subimpeding Relapse is defined as 1 or more of the following items:

1. CGI-I of >5 (minimally worse), **AND**

   an increase of any of the following individual PANSS items: conceptual disorganization, hallucinatory behavior, suspiciousness, and unusual thought content, to a score of >4.

2. CGI-SS of 3 (moderately suicidal) on Part 1

These criteria will be applied to all patients who early terminate the study and do not meet the impending relapse criteria of the primary endpoint. All early terminated patients who meet the subimpending relapse criteria will be considered as having events on the day after the early termination date recorded *in addition* to patients who already met the impending relapse criteria of the primary endpoint. The primary analysis for the primary efficacy estimand will be repeated for this endpoint.

Additional definitions may be used as a post-hoc analyses to better evaluate the effect of patients who early terminate the study while they experience a high relative increase in the total PANSS at early termination.

6.3. Secondary Efficacy Endpoints and Analysis

6.3.1. Key Secondary Efficacy Endpoints and Analyses

The eITT analysis set (Section 3.3) will be used for all summaries in this section, unless otherwise specified. The summaries will be presented by age subset (adults/adolescents) and by total, as applicable.

For the endpoints that are evaluated only in the adult population, the analysis will be conducted on the ITT analysis set.

The key secondary endpoints and their respective methods of analysis are described below.

6.3.1.1. Time to impending relapse in the eITT analysis set

Time to impending relapse (as defined under the primary objective, see Section 1.1) in the total population (adults and adolescents).

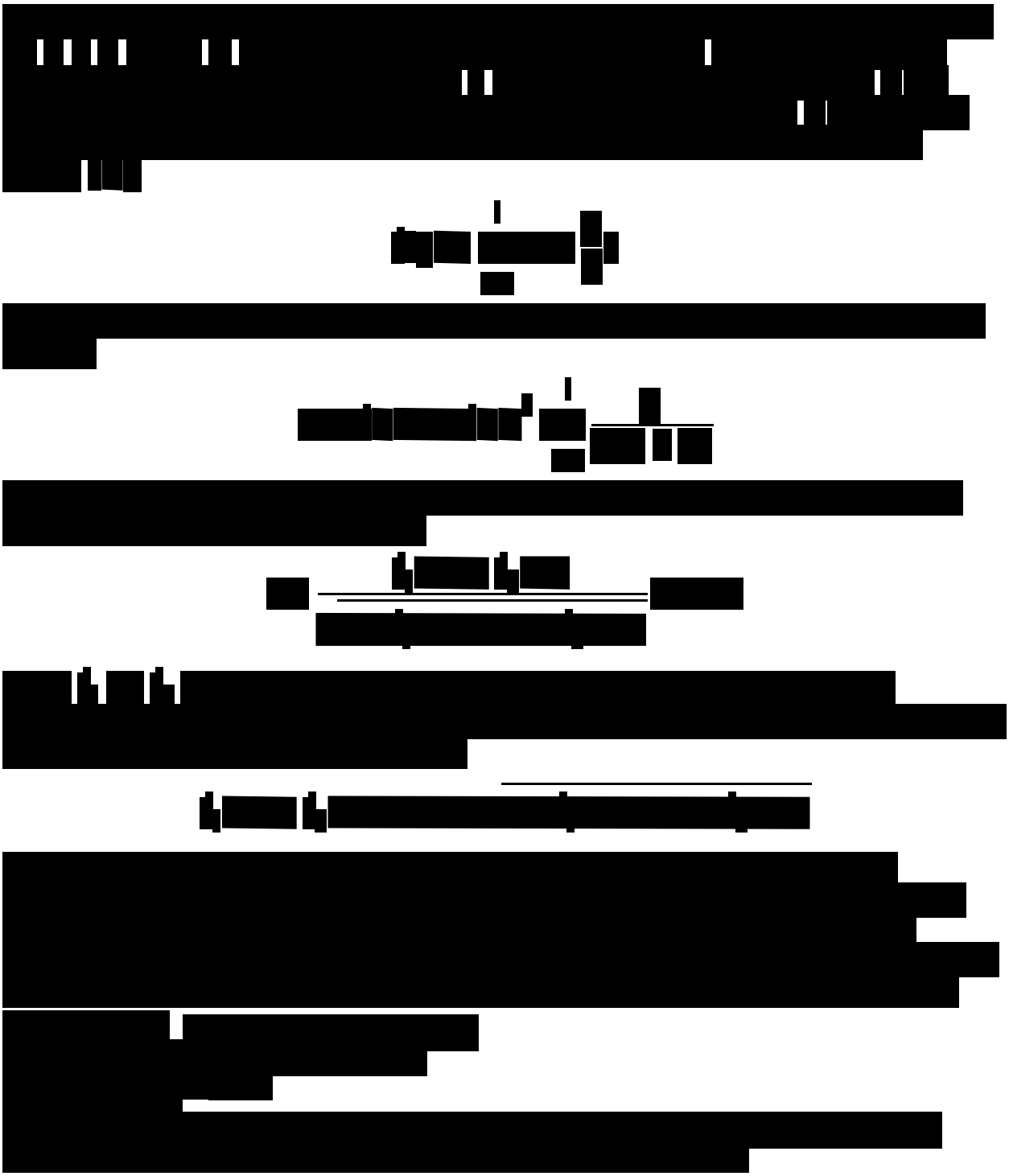
This endpoint will be analyzed similarly to the primary efficacy analysis, as described in Section 6.2.2., using the eITT analysis set including age group (a flag for adolescent) in the Cox model (if applicable) (see Section 3.3).

6.3.1.2. Impending Relapse Rate at Week 24

6.3.1.2.1. Definition

Impending relapse rate at week 24 will be calculated as 1 minus the proportion of patients free of impending relapse events at week 24.

6.3.1.2.2. Analysis



[redacted]

6.3.1.3. Percentage of Patients Who Maintain Stability at Endpoint

6.3.1.3.1. Definition

Stability is defined as meeting all of the following criteria for at least 4 consecutive weeks: outpatient status; PANSS total score ≤80; minimal presence of specific psychotic symptoms on the PANSS, as measured by a score of ≤4 on each of the following items: conceptual disorganization, suspiciousness, hallucinatory behavior, and unusual thought content; CGI-S score ≤4 (moderately ill); and CGI-SS score ≤2 (mildly suicidal) on Part 1 and ≤5 (minimally worsened) on Part 2.

The use of eITT analysis complies with while on treatment estimand and will be comprised of the difference in proportions under the treatment to which the patient was initially randomized of all patients who were successfully stabilized on oral risperidone at a daily dose range of 2 mg to 5 mg. Accordingly, this estimand addresses the clinical question to what extent patients improve their chances to maintain stability once they start the treatment if the patient adheres to the treatment long enough.

The percentage will be calculated as the number of patients who maintained stability at endpoint divided by the number of eITT patients that adhere to treatment for at least 4 weeks (based on the allowed visit time window) or experienced relapse in the given treatment group.

The patients who maintain stability at endpoint might be discussed before unblinding and documented in the study data review document.

6.3.1.3.2. Analysis

[redacted]

6.3.1.4. Percentage of Patients Achieving Remission at Endpoint

6.3.1.4.1. Definition

Positive symptom, negative symptom, and overall symptom remission will be examined and are defined by Andreasen et al (2005), including severity and duration criteria. All remission criteria can be derived from PANSS items.

For overall symptom remission, the patient must not relapse during the study and in addition, over a period of at least 6 months preceding the endpoint, maintain scores of ≤3 on each of the 8 specific PANSS items: P1 (delusions), G9 (unusual thought content), P3 (hallucinatory behavior), P2 (conceptual disorganization), G5 (mannerisms/posturing), N1 (blunted affect), N4 (social withdrawal), and N6 (lack of spontaneity).

The use of eITT analysis complies with the while on treatment estimand, and will be comprised of the difference in proportions under the treatment to which the patient was initially randomized of all patients who were successfully stabilized on oral risperidone at a daily dose range of 2 mg to 5 mg. Accordingly, this estimand addresses the clinical question to what extent patients improve their chances to achieve remission once they start the treatment if the patient adheres to the treatment long enough.

The percentage will be calculated as the number of patients who maintained stability at endpoint divided by the number of eITT patients that adhere to treatment for at least 6 months (based on the allowed visit time window) or experienced relapse in the given treatment group.

The patients achieving remission at endpoint might be discussed before unblinding and documented in the study data review document.

6.3.1.4.2. Analysis

The percentages of patients achieving remission at endpoint will be compared between pooled active treatment group versus placebo using the CMH test adjusting for stratification variables accompanied by Breslow and Day Homogeneity test, similarly to Section 6.3.1.3.2.

6.3.1.5. Observed Rate of Impending Relapse at Endpoint

6.3.1.5.1. Definition

Observed rate of impending relapse will be calculated as the number of patients who relapsed by endpoint divided by the number of patients in each treatment group. The last valid patient assessment will be used as the endpoint.

6.3.1.5.2. Analysis

The observed impending relapse rates at endpoint will be compared between treatment groups versus placebo using the CMH test adjusting for stratification variables, similarly to Section 6.3.1.3.2 [redacted]

A p-value of Breslow and Day Homogeneity test will be presented for each comparison, as well.

6.3.1.6. Drug Attitudes Inventory 10-item Version

6.3.1.6.1. Definition

The Drug Attitudes Inventory (DAI) initially consisted of 30 items which focused on the subjective effects of the neuroleptic medications in patients with schizophrenia. The scale was developed to obtain a more complete understanding of factors influencing medication compliance in these patients.

A brief 10-item scale (DAI-10) that focuses on medication effects was constructed. It demonstrated a high correlation with medication compliance and treatment outcome. The DAI-10 is used in this study.

The response for each item is either “True” (T) or “False” (F) as applies to the patient. An interpretation is required because “true” may indicate either a positive or negative view about medications. Table 2 contains the standard of compliant response for each DAI-10 item/statement.

Table 2: Drug Attitudes Inventory 10-item Version – Standard of a Compliant Response

Item	Statement	Compliant response
1	For me, the good things about medication outweigh the bad.	T
2	I feel weird, like a "zombie", on medication	F
3	I take medications of my own free choice	T
4	Medications make me feel more relaxed	T
5	Medication makes me feel tired and sluggish	F
6	I take medication only when I am sick	F
7	I feel more normal on medication.	T
8	It is unnatural for my mind and body to be controlled by medications.	F

Table 2: Drug Attitudes Inventory 10-item Version – Standard of a Compliant Response (Continued)

Item	Statement	Compliant response
9	My thoughts are clearer on medication	T
10	By staying on medications, I can prevent getting sick.	T

F = false; T = true.

The scoring of the DAI-10 will be done as follows:

1. For each item, a compliant response as appears in Table 2, will be scored as plus 1; an compliant response will be scored as minus 1.
2. The total score is the sum of the item scores from #1.

The total score ranges from -10 to +10. A positive total score indicates a positive attitude toward psychiatric medications and thus corresponds to a compliant response, and vice versa.

The DAI-10 will be assessed in adult patients only at screening, baseline, the ET/EoT visit and at the follow-up visit 2 (8-weeks after dosing visit).

6.3.1.6.2. Analysis

By visit change from baseline to each visit and endpoint in the DAI-10 total score will be calculated. Least squares means and 95% CI of change from baseline to endpoint, if applicable, will be performed using a repeated measures analysis of covariance (repeated measures ANCOVA) model with treatment arm (trt), subject id (subjid) and study visit (avisit), stratification variable (rand_strata) as factors and baseline DAI-10 score (Baseline) as a covariate. If for some visits there are less than 15 observations, they will not by incorporated into the model, as convergence problems might occur. For these cases only mean with 95% CI approximation using t-distribution will be presented.

[redacted]

6.3.1.7. Schizophrenia Quality of Life Scale

6.3.1.7.1. Definition

Schizophrenia Quality of Life Scale (SQLS) comprises 33 items categorized in 2 domains: psychosocial feelings (20 items - Items 3, 4, 5, 6, 8, 10, 11, 13, 15, 16, 17, 18, 19, 21, 22, 24, 25,

27, 29, 30) and cognition and vitality (13 items – Items 1, 2, 7, 9, 12, 14, 20, 23, 26, 28, 31, 32, 33). The items are scored on a five-point scale (1 - never, 2 - rarely, 3 - sometimes, 4 - often, 5 - always). For all items, except the specific 4 items discussed below, higher scores indicate comparatively lower quality of life.

The following four items have reversed scaling and refer to positive aspects of life:

7. I was able to carry out my day to day activities;

12. I felt that I could cope;

14. I slept well;

26. I felt happy.

The above are to be recoded such as 6 minus recorded score (eg. 6-ITEM_7) before the scale total is calculated. Individual domain and total scores are standardized by scoring algorithm to a 0 (best health status) to 100 (worst health status) scale, with higher scores indicating comparatively lower quality of life.

Therefore, the total score, namely TS, is calculated as the total of all 33 items ($R4S_{tot}$) minus the number of items answered, divided by the maximum possible total score, namely $RS_{max} = 33 \times 4 = 132$, and multiplied by 100.

$$TS = \frac{R4S_{tot}}{RS_{max}} \times 100$$

Similarly, each scale score, 'Psychosocial feelings' and 'Cognition and vitality' will be calculated in the same way using the total score and number of items in the scale.

SQLS will be assessed in adult patients only at screening, baseline, week 8, week 12, and every 12 weeks thereafter, including the ET/EoT visit, and at follow-up visit 2 (8 weeks after the last dosing visit).

For handling the missing items per each scale score see Section 4.3.2.

6.3.1.7.2. Analysis

The by-visit change from baseline of the total score will be calculated. Least squares means and 95% CI of change from baseline in the SQLS will be presented using a repeated measures analysis of covariance (repeated measures ANCOVA) method, with treatment arm (trt), stratification variables (rand_strata), subject id (subjid) and study visit (avisit), and baseline (Baseline, SQLS score at the end of Stage 1) as a covariate. [redacted]

SQLS total and domain score and change from baseline will be summarized using descriptive statistics (least-squares mean will be presented as appropriate) at each time point for all treatment groups.

6.3.1.8. Time to Impending Relapse in Adolescents

Time to impending relapse in adolescent patients will be assessed if the number of randomized adolescent patients will be at least 10 with clinically sufficient exposure. This endpoint will be

analyzed similarly to the primary efficacy analysis, as described in Section 6.2.2, but will use the adolescent patient subset of the eITT analysis set (see Section 3.3)

In addition, the treatment effect within the adolescents will be compared to the treatment effect within the adults. Hazard ratio and 95% CI of adolescents compared to adults will be presented (in the same approach presented in Section 6.2.4 for subgroup analysis).

6.4. [redacted]

6.4.1. 5-Level EuroQol Five Dimensions Questionnaire

6.4.1.1. Definition

The 5-Level EuroQol Five Dimensions Questionnaire (EQ-5D-5L) consists of the EQ-5D descriptive system and the EQ visual analogue scale (VAS). The EQ-5D descriptive system includes 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The values for each of the dimensions are: 1=no problems, 2=slight problems, 3=moderate problems, 4=severe problems and 5=unable to/extreme problems.

The EQ VAS records the respondent's self-rated health on a vertical VAS where the endpoints are labeled '100=Best imaginable health state' and '0=Worst imaginable health state'.

EQ-5D-5L will be assessed in adult patients only at screening, baseline, week 8, week 12, and every 12 weeks thereafter, including the ET/EoT visit, and at follow-up visit 2 (8 weeks after the last dosing visit).

6.4.1.2. Analysis

By visit change from baseline and to endpoint in each of the EQ-5D-5L domains will be calculated. The analysis of change from baseline to each visit and endpoint in each of the EQ-5D-5L domains will be performed similarly to Section 6.3.1.6.2, using the appropriate baseline EQ-5D-5L domain as a covariate, respectively.

The change from baseline for each of the EQ-5D-5L domains will also be summarized using descriptive statistics, and least squares means at each time point by treatment group as appropriate.

6.4.2. Healthcare Resource Utilization

6.4.2.1. Definition

The percentage of patients who were hospitalized, number of hospitalizations, and length of hospital stay (number of days); percentage of patients who had emergency room (ER) visits and number of ER visits (ie, not including protocol-mandated outpatient visits); and percentage of patients who had outpatient visits and number of outpatient visits will be calculated.

6.4.2.2. Analysis

The parameters defined above in Section 6.4.2.1 will be assessed in all patients (ie, both adults and adolescents) and will be summarized using descriptive statistics for both study periods.

6.4.3. Change in PANSS Total Score from Baseline to Endpoint

6.4.3.1. Definition

The PANSS comprises a total of 30 items. Each item is scored using 7 rating points, with an increasing level of psychopathology, as follows 1 – absent; 2 – minimal; 3 – mild; 4 – moderate; 5 - moderate severe; 6 – severe; 7 – extreme.

7 of the PANSS items constitute a Positive scale (P1-P7) with a maximum score of 49, 7 a Negative scale (N1-N7) with a maximum score of 49 and 16 a General Psychopathology scale

with a maximum score of 112. The score for these scales are generated by a summation of ratings. The total or overall PANSS score is the sum of all of the 30 items, and it will only be calculated if all the items are present, the maximum score of the total score is 210.

PANSS will be assessed at all in-clinic visits during the study.

Percent reduction of the total score from screening or baseline (=reference) will be conducted by calculating

$$\text{Reduction}(\%) \text{ at visit k} = 100\text{-}\ \frac{100 \cdot (Total\ score\ at\ visit\ k - 30)}{Total\ score\ at\ reference\ visit\ - 30}$$

If the total score at baseline or screening is 30, then the denominator will be set to 1, see Leucht et al. 2010.

6.4.3.2. Analysis

The change in the PANSS total score from screening to each visit will be calculated for Stage 1. The change from baseline in total score will be calculated to each visit and to endpoint for Stage 2. The PANSS total score from baseline to endpoint analysis will be performed similarly to that described in Section 6.3.1.6.2.

PANSS total score and change from screening will be summarized using descriptive statistics for each dose at each time-point during Stage 1. PANSS total score and change from baseline will be summarized using descriptive statistics for each treatment arm at each time point and to endpoint during Stage 2 as well as for the follow up visits using least squares means and 95% CI calculated by a repeated measures analysis of covariance model and nominal p-values for the comparisons versus placebo. However, follow-up visits may not be performed for all patients (ie, patients rolling over into the extension study).

Individual PANSS items and/or scales may be summarized using descriptive statistics during Stage 1, 2 or both.

Percent reduction will be summarized in a frequency table, with using the following percent reduction categories: ≤0%, 1%-24%, 25%-49%, 50%-74%, 75%-100%.

6.4.4. CGI-I Score at Endpoint

6.4.4.1. Definition

The Clinical Global Impression–Improvement (CGI-I) scale permits a global evaluation of the patient's improvement in symptoms overall. The CGI-I scale rates the patient's improvement relative to his or her symptoms on a 7-point scale ranging from 1 (very much improved) to 7 (very much worse).

The CGI-I will be administered by the investigator/trained rather at all visits during the study (except for the screening and baseline visits).

CGI-I during Stage 1 will be relative to screening; CGI-I during Stage 2 will be relative to the baseline visit.

6.4.4.2. Analysis

CGI-I will be analyzed in the same way as described in Section 6.4.1.2, except that the baseline Clinical Global Impression of Severity (CGI-S) will be included as a covariate in the model as the efficacy measure at baseline, and the target variable of the actual value by visit will be analyzed rather than change from baseline for the parameter itself. Subsequently this model is measuring change from a reference point (baseline).

CGI-I will be summarized using descriptive statistics during Stage 1 by oral dose, and during Stage 2 and the follow up visits by treatment group.

6.4.5. Personal and Social Performance Scale

6.4.5.1. Definition

The Personal and Social Performance Scale (PSP) is a clinician-rated instrument that measures personal and social functioning in patients with schizophrenia. The PSP is a 100-point single-item rating scale, divided into10 equal intervals. The score is based on the assessment of patient's functioning in 4 categories in the past: 1) socially useful activities; 2) personal and social relationships; 3) self-care; and 4) disturbing and aggressive behaviors. These 4 patient functioning categories are assessed via the degree of difficulty ranging from 1 (absent) to very severe (6).

Higher PSP (also called total) scores represent better personal and social functioning, with ratings from 91 to 100 indicating more than adequate functioning, while scores under 30 indicating functioning so poor that intensive supervision is required.

The PSP will be assessed in adult patients only at screening, baseline, week 8, week 12, and after every 12 weeks thereafter, including ET/EoT visit, and at follow-up visit 2 (8 weeks after the last dosing visit).

6.4.5.2. Analysis

By visit change from baseline in the PSP (also called total) score will be calculated. The PSP total score will be analyzed in the same way as described in Section 6.3.1.6.2 except that the baseline PSP score will be included in the model as the efficacy measure at baseline PSP total score and change from baseline will be summarized using descriptive statistics by treatment arm at each time point and to endpoint.

The areas of functioning assessments will be summarized using descriptive statistics by treatment arm at each timepoint and to endpoint.

7. MULTIPLE COMPARISONS AND MULTIPLICITY

The primary endpoints will be tested using a fixed sequential (hierarchy) methodology used to control the overall type-I error of the study. First, q1m versus placebo will be tested at a 2-sided 0.05 alpha level. If this comparison will be found to be statistically significant, the comparison of q2m versus placebo will be tested at a 2-sided 0.05 alpha level. Thus, the overall type I error in the study will be controlled for both dose comparison tests versus placebo at a 2-sided alpha of 0.05.

If the 2 primary efficacy endpoints are found to be statistically significant, type-I error will be further controlled for the key secondary endpoints by employing a fixed sequential (hierarchical) testing strategy. Key secondary endpoints will be analyzed in a pooled manner for the 2 active treatment groups versus placebo.

The sequence of the secondary endpoint comparisons will be as follows:

1. Time to impending relapse in the eITT analysis set
2. Impending relapse rate at week 24
3. Percentage of patients who maintain stability at endpoint
4. Percentage of patients achieving remission at endpoint
5. Observed rate of impending relapse at endpoint
6. DAI-10 change from baseline to endpoint
7. SQLS change from baseline to endpoint
8. Time to impending relapse in adolescent patients

If the first comparison is found to be significant, then the next comparison of interest will be interpreted inferentially at the same alpha level. This process will continue either until all comparisons are tested inferentially or until the point at which the resulting 2-sided test is insignificant at the same alpha level. At the point where a comparison is found to be insignificant, the rest of the comparisons will be made in an exploratory manner only; any p-value associated with these comparisons will be considered as nominal and will not be used for inference.

Any additional pairwise comparisons between treatment groups (eg, q1m versus placebo) may be presented but will not be interpreted inferentially.

[redacted]

8. SAFETY ANALYSIS

8.1. General

Safety analyses will be performed on the safety analysis set for Stage 2 (Section 3.5) by age group (adolescents [ages 13-17] and adults [18 years of age and above]), before/after COVID-19 pandemic (if data permits), unless otherwise noted. Summaries of treatment emergent adverse events for Stage 1 will be presented on the enrolled analysis set (Section 3.1) by randomized and not-randomized groups and overall, unless otherwise stated. Summaries of treatment emergent adverse events for Stage 2 will be presented by treatment group, unless otherwise stated.

Selected safety data will also be presented by site of injection (abdomen vs arm), as applicable.

Safety assessments and time points are provided in Table 1 and Table 2 of the study protocol.

For continuous variables, descriptive statistics (n, mean, standard deviation, median, minimum, and maximum) will be provided for actual values and changes from baseline to each time point, as appropriate. For categorical variables, patient counts and percentages will be provided as appropriate.

8.2. Duration of Exposure to Risperidone during Study

Duration of exposure to risperidone will be analyzed at the different study stages as the number of days a patient received drug (last day of study drug and first day of study drug). Duration of treatment calculation for the oral and the injectable risperidone will be different in order to account for the difference in frequency of administration. The exposure at each stage will be summarized using descriptive statistics.

8.2.1. Oral Stabilization Period – Stage 1

For the individual patient, duration of exposure to oral risperidone (days) during Stage 1 is the number of days the patient received the oral risperidone during the study (last day of oral risperidone drug – first day of oral risperidone + 1). In patients not randomized to Stage 2 of the study, in case of missing end of study date or a lost-to-follow-up patient, (planned) randomization date will be considered as last date of subject assessments.

Duration of treatment (days) will be summarized using descriptive statistics.

8.2.2. Relapse Prevention - Stage 2

Duration of treatment (days treated) during Stage 2 will be defined based on the first and last days of IMP injection. Duration of treatment calculation for the different treatment regimens will be as follows:

- q1m and placebo - last day– first day + 29;
- q2m - last day– first day + 57.

Weeks of IMP treatment using the categories ≤13 week, >13 to ≤26 weeks, >26 to ≤39 weeks, >39 to ≤52 weeks or more than 52 weeks or as applicable will be summarized using descriptive statistics. Duration of treatment (weeks) will also be summarized using descriptive statistics.

8.2.3. Total of Exposure to Risperidone in the Study

For an individual patient, total exposure to risperidone during the study is the sum of Stage 1 and Stage 2. Total weeks on risperidone in the study using the categories ≤13 weeks, >13 to ≤26 weeks, >26 to ≤39 weeks, >39 to ≤52 weeks or more than 52 weeks will be summarized using descriptive statistics. Duration of study (weeks) will also be summarized using descriptive statistics.

8.3. Study Drug Compliance

8.3.1. Oral Stabilization Period – Stage 1

Per the study protocol, a patient with the total compliance of less than 80% will be considered noncompliant during Stage 1. This is an important criterion in regard to compliance with the study protocol, which is required for the patient to be eligible for randomization.

The compliance is defined as the ratio between the dose taken and the dose required:

$$\%\ Compliance = \frac{Actual\ drug\ taken\ (mg)}{Expected\ drug\ to\ be\ taken\ (mg)} \times 100\%$$

'Actual drug taken' will be calculated as the total amount of drug dispensed less total amount of drug returned, both during the entire oral stage, based on "study drug accountability – oral" field in the CRF. Drug dispensed (in mg) will be the number of dispensed tablets (Total Number Dispensed) multiplied by the corresponding dose (Dispensed Units). The total amount of drug dispensed will be the sum over all the visits where the drug dispensation occurred, both scheduled and unscheduled. The same calculation will be done to obtain the total amount of drug returned, replacing "dispensed" by "returned".

'Expected drug to be taken' will be calculated as $\text{sum}(\text{dose} \times \text{days to be taken})$, where 'dose' is the dose prescribed (with one exception, below) by the investigator from the 'Study drug administration - oral medication' field in the CRF and 'days to be taken' is the number of days between the corresponding 'Start Date of Treatment' dates from the same CRF or the randomization date (for patients with missing randomization data, or lost to follow up the last assessment data will be used). For patients for which the oral dose prescribed is below 2mg/day (eg 1mg or 0.5mg a day), 2mg/day will be used as the expected dose and not the actual dose taken.

If no drug was returned, it will be accounted for as '0' returned amount for this visit in the compliance calculation. It should be noted that if the investigator or the sponsor determines that the patient was not in compliance with the study protocol, the case will be evaluated on a case-by-case basis, and the investigator and the sponsor will determine whether the patient will be randomized to the double-blind period (Stage 2).

Oral drug compliance during Stage 1 will be categorized as <60%, 60% to < 80%, 80% to < 100%, and ≥100% and summarized using descriptive statistics. A numerical example of compliance calculation during Stage 1 is given below.

Numerical Example of Oral Drug Compliance during Stage 1

Assuming that after looking at the Start Dates of Treatment, the patient had a total 84 days of the oral period, of which of 2 days were on 1mg/day, 40 days on 3 mg/day and 42 days on 4 mg/day – the expected drug to be taken (mg) would be $1 \times 2 + 3 \times 40 + 4 \times 42 = 290$ mg.

For of the entire oral stage period, the patient was dispensed a certain amount (2 tablets of 1 mg, 70 tablets of 3 mg/day and 50 tablets of 4 mg/day) and returned a certain amount (30 tablets of 3 mg/day and 10 tablets of 4 mg/day) –

- the total drug dispensed would be $2 \times 1 + 68 \times 3 + 50 \times 4 = 406$ mg;
- the total drug returned would be $30 \times 3 + 10 \times 4 = 130$ mg;
- Actual drug taken would be $406 - 130 = 276$ mg

Drug compliance would be Drug consumed/ Drug prescribed=276/290=95% (>80%)

$$\%\ Compliance = \frac{Actual\ drug\ taken\ (mg)}{Expected\ drug\ to\ be\ taken\ (mg)} \times 100\% = \frac{276}{290} \times 100\% = 95\%$$

8.3.2. Relapse Prevention - Stage 2

Drug compliance during Stage 2 will not be assessed.

8.4. Adverse Events

For recording of adverse events, the study period is defined for each patient as the time period from signature of the ICF to the end of the follow-up period. All adverse events will be coded using the MedDRA. Each patient will be counted only once in each PT or SOC category for the analyses of safety. Summaries will be presented for all adverse events (overall and by severity), including adverse events determined by the investigator to be related to oral risperidone (Stage 1) or TV-46000 (Stage 2) (defined as related or with missing relationship) (overall and by severity), serious adverse events, and adverse events causing withdrawal from the study. Summaries will be presented by randomized, not-randomized and overall using the enrolled patients set (Stage 1) or treatment group using the safety analysis set (Stage 2).

Summaries will include treatment-emergent adverse events which are defined as adverse events occurring at or after the first day of oral risperidone until the last day of oral risperidone in Stage 1 and for Stage 2 as adverse events occurring at or after the first IMP injection.

In case of the missing dates and severity, the imputation will be performed as described in Section 4.3.

Multiple records with the same PT and adverse event onset date for the same patient, or with overlapping or consecutive dates, are counted only once selecting the adverse event with the highest severity and seriousness. For injection site reactions, multiple records with overlapping or consecutive dates are counted only once only if the injection site is the same, selecting the adverse event with the highest severity and seriousness. If the onset date of adverse events with the same PT is partially unknown or duration is < 24 hours, then these adverse events will be counted as separate adverse events, except for the cases where an adverse event with duration of < 24 hours has the same onset date as another adverse event with a longer duration, then the

adverse events with the longer duration adverse events will be counted. The event rate per 100 years (PY) of patient treatment exposure is calculated as 100*(Number of cases/PY).

Safety data collected in Stage 1 will also be summarized using descriptive statistics in the enrolled patients set. Adverse events that started in Stage 1 and ended in Stage 2 will be reported only in the oral phase (ie, under Stage 1).

Selected safety data will also be presented by site of injection (abdomen vs arm).

Separate summaries for adolescent patients may be presented separately for some analyses, as applicable.

Patient listings of all adverse events, serious adverse events and adverse events leading to withdrawal and adverse events leading to death will be presented.

Summary of the number of adverse events per week following each sc injection will be presented.

8.5. Deaths

If any patient dies during the study, a listing of deaths will be provided, and all relevant information will be discussed in the patient narrative included in the CSR.

8.6. Clinical Laboratory Tests

The clinical laboratory tests (serum chemistry, hematology and urinalysis) are detailed in Table 7 of the study protocol.

Laboratory test results will be presented in standard international (SI) units.

Summary statistics for chemistry, hematology, and urinalysis laboratory tests will be presented at time points detailed in the schedule of assessment Tables 1 and 2 of the study protocol including last assessment (as a separate visit). Laboratory values and changes from screening or baseline (as appropriate) to each visit will be summarized using descriptive statistics.

Boxplots for selected laboratory tests, including neutrophils, alanine aminotransferase, aspartate aminotransferase and bilirubin, will be presented by visit and treatment group.

Shifts (below, within, and above the normal range) from screening or baseline (as appropriate) to each visit and endpoint assessment will be summarized using patient counts.

Summaries of potentially clinically significant abnormal values will include all post-baseline values (including scheduled, unscheduled, and withdrawal visits). The incidence of potentially clinically significant abnormal values will be summarized for laboratory variables using descriptive statistics with the criteria specified in Table 3.

Table 3: Criteria for Potentially Clinically Significant Laboratory Values

Test		Criterion value
Serum chemistry		
Alanine aminotransferase (ALT)		≥3x ULN
Aspartate aminotransferase (AST)		≥3x ULN
Alkaline phosphatase		≥3x ULN
Lactate dehydrogenase (LDH)		≥3x ULN
Blood urea nitrogen (BUN)		≥10.71 mmol/L
Creatinine		≥177 µmol/L
Uric acid	Men	≥625 µmol/L
	Women	≥506 µmol/L
Bilirubin (total)		≥2x ULN
Potassium		≤3 mmol/L ≥6 mmol/L
Calcium		≤1.5 mmol/L ≥3.5 mmol/L
Hematology		
Hemoglobin	Men	≤115 g/L
	Women	≤95 g/L
White blood cell (WBC) counts		≤3 x 10^9/L ≥20 x 10^9/L
Absolute neutrophil counts (ANC)		≤1 x 10^9/L
Platelet counts		≤75 x 10^9/L ≥700 x 10^9/L

ULN=upper limit of normal range.

8.6.1. Laboratory Values Meeting Hy's Law Criteria

All occurrences of possible drug-induced liver injury that meet Hy's law criteria, defined as **all** of the below, must be reported by the investigator to the sponsor as a serious adverse event:

- ALT or AST increase of >3× the upper limit of the normal range (ULN)
- total bilirubin increase of >2× ULN
- absence of initial findings of cholestasis (ie, no substantial increase of alkaline phosphatase)

An adverse event that does not meet any of the criteria for seriousness listed above will be regarded as a non-serious adverse event.

All incidences will be listed.

8.6.2. Other Clinical Laboratory Tests

8.6.2.1. Virology and Thyroid Screening Tests

At screening, patients will be tested for HIV, hepatitis B surface antigen, hepatitis C antibody, thyroid stimulating hormone, triiodothyronine, and thyroxine. Due to potential effects on clinical psychiatric symptoms, levels of thyroid hormones (thyroid stimulating hormone, triiodothyronine, and thyroxine) will also be assessed.

The data will be presented as listings.

8.6.2.2. Human Chorionic Gonadotropin Tests

A FSH test will be performed at the screening visit for any female who has been without menses for at least 12 months to confirm post-menopausal status. Women may be allowed not to use contraceptives during the study if they had no menses for at least 12 months (without an alternative medical cause) and the FSH concentration is above 35 U/L.

A serum beta human chorionic gonadotropin (β-HCG) test will be performed for all women of child bearing potential at screening, baseline, and the follow-up/exit visits (see Table 1 and Table 2 of the study protocol). At baseline, both a serum and a urine β-HCG test will be performed. Urine β-HCG (dipstick) test will be performed for women of child bearing potential at all visits where study drug is administered (prior to study drug administration). A negative result must be obtained prior to study drug administration.

The FSH, serum and urine β-HCG data will be listed.

8.6.2.3. Urine Drug Screen

A urine drug screen will be performed at the time points specified in Table 1 and Table 2 of the study protocol. The urine drug screen detects the presence of drugs of abuse, including amphetamine, barbiturates, benzodiazepines, cocaine, opiates, and tetrahydrocannabinol.

Any findings will be listed.

8.6.2.4. Prolactin

Blood samples will be obtained for prolactin test at the time points specified in Table 1 and Table 2 of the study protocol. The test results will remain blinded to the sponsor until the end of the study.

The test results will be presented as described for all the clinical laboratory test results outlined in Section 8.6 above. In addition, shift tables from screening - from normal to 2x ULN and 5x ULN - will be presented.

8.6.2.5. Glomerular Filtration Rate

Glomerular Filtration Rate (GFR) will be calculated based on plasma creatinine levels, weight, gender and age using the Cockcroft-Gault equation.

The estimated GFR values and changes from baseline will be presented by visit using descriptive statistics. The individual GFR measurements will also be presented in the listing.

8.7. Physical Examinations

Physical examinations, including height (to be measured at the screening visit only) and weight, will be performed at the time points detailed in Table 1 and Table 2 of the study protocol. The full physical examination will consist of examining the following body systems: cardiovascular, respiratory, abdominal, skin, neurological, and musculoskeletal systems. The physical examination will also include examination of general appearance, including head, eyes, ears, nose, and throat; chest; abdomen; skin; lymph nodes; and extremities. Body weight and tympanic temperature will be measured at each visit. Systolic and diastolic blood pressure and pulse rate will be measured with the patient in a seated position. Any physical examination finding that is judged by the investigator as a clinically significant change (worsening) compared with a baseline value will be considered an adverse event, recorded on the CRF, and monitored as described in Section 7.1.2 of the study protocol.

Descriptive statistics for weight and height will be provided. All the data will be listed. Boxplots for weight will be presented by visit and treatment group.

In addition, BMI results and percent changes from screening to each visit in Stage 1 and percent changes from baseline to each visit in Stage 2 will be presented for both the enrolled patients set and safety analysis set. BMI shifts from reference to each visit by treatment group will be presented, where the reference is the screening measurement for Stage 1 and the baseline measurement for Stage 2, using the following limits: '<0', '≥0-5%', '≥5-10%' and '≥10%'.

8.8. Vital Signs

Summary statistics for vital signs (blood pressure [systolic/diastolic], respiratory rate, tympanic temperature, and pulse) will be presented at the time points including last assessment (as a separate visit) detailed in Table 1 and Table 2 of the study protocol.

Vital signs values and changes from screening or baseline to each visit will be summarized by stage of study using descriptive statistics. All the data will be listed.

Boxplots for selected vital signs, including pulse, systolic blood pressure and diastolic blood pressure, will be presented by visit and treatment group.

Summaries of potentially clinically significant abnormal values will include all post baseline values (including scheduled, unscheduled, and withdrawal visits). The incidence of potentially clinically significant abnormal vital sign values, specified by the criteria in Table 4, will be summarized using descriptive statistics. Systolic and diastolic blood pressure will be included in all summaries.

Table 4: Criteria for Potentially Clinically Significant Vital Signs

Vital Sign	Criterion value
Pulse	≥120 bpm
	≤50 bpm
Systolic blood pressure	≥180 mm Hg
	≤90 mm Hg

Table 4: Criteria for Potentially Clinically Significant Vital Signs (Continued)

Vital Sign	Criterion value
Diastolic blood pressure	≥105 mm Hg
	≤50 mm Hg
Body temperature	≥38.3°C

bpm = beats per minute; mm Hg = millimeters of mercury.

8.9. Electrocardiography

A standard 12-lead ECG will be recorded at the time points detailed in Tables 1 and 2 of the study protocol. Triplicate measurements will be performed at screening and baseline and single measurements at all other in-clinic visits.

Electrocardiogram (ECG) findings (normal and abnormal) at screening and baseline will be summarized using descriptive statistics; the data presentation will include both eResearch Technology (the central ECG vendor) and investigator interpretation. The average of the 3 ECG screening and baseline assessments will be calculated; if more measurements are taken, the 3 last measurements taken prior to first dose administration will be used both for the average calculation and for choosing the worst result for the interpretation.

If the investigator interpretation is missing, the following derivation rule will be used:

1. If ERT interpretation is 'normal' then investigator interpretation on the same date will be derived as 'normal';
2. Otherwise investigator interpretation will not be derived.

Shifts (normal and abnormal) from screening to baseline (for randomized patients) and from baseline to overall result interpretation at each visit and last assessment will be summarized using patient counts. For overall result interpretation the worst post baseline finding for the patient (the abnormal finding if there are both normal and abnormal findings) will be used in the summaries. Summary statistics for ECG variables values will be presented. Actual values and changes from screening to baseline and from baseline to each visit and last assessment will be summarized using descriptive statistics.

The incidence of potentially clinically significant abnormal values for ECG variables will be summarized using descriptive statistics with the criteria specified below:

- QT interval corrected using Fridericia's formula (QTcF) values >450 msec or >480 msec or >500 msec.
- QTcF change from baseline values >30 or >60.
- PR change from baseline ≥25% and a value >200.
- QRS change from baseline ≥25% and a value >110.
- Heart rate value <60 bpm or >100 bpm.

8.10. Concomitant Medications or Therapies

Concomitant therapies and medications, including medications that are taken on an as needed basis and occasional therapies, will be monitored during the study. Details of prohibited medications may be found in Section 5.7 of the study protocol. All concomitant medications will be coded using the WHO Drug.

The incidence of concomitant therapies and medications will be summarized using descriptive statistics by therapeutic class category and PT. Patients are counted only once in each therapeutic class, and only once in each PT category.

A distinction will be made between the study stages. For Stage 1, concomitant therapies and medications will be presented using the enrolled patients analysis set (Section 3.1) and will include all medications up to randomization day (Day 1), excluding the randomization day. For Stage 2, the concomitant therapies and medications will be presented using the safety analysis set and will include all medications from the day of randomization (baseline, Day 1) and up to the end of study as defined in the study protocol.

Descriptive statistics for allowed rescue medications (see Section 5.7 in the study protocol) will be presented by treatment group and by study stage.

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

Any positive answer to the behavior subcomponents at screening or baseline identifies a patient with “Suicidal Behavior at Screening” or “Suicidal Behavior at Baseline” respectively.

For the derivation of Screening Behavior, all assessments until the beginning of Stage 1 should be used, while for the derivation of Baseline Behavior, all assessments after the beginning of Stage 1 and till the beginning of Stage 2 should be used. Same derivation will be done also for "Suicidal Ideation at Screening” and "Suicidal Ideation at Baseline”. A patient identified with either “Suicidal Behavior at Screening/Baseline” or with “Suicidal Ideation at Screening/Baseline” is also classified with “Suicidal Behavior or Ideation at Screening/Baseline”.

For post-screening/ post-baseline periods, any positive answer to the behavior subcomponents in any of the post screening or baseline assessments, according to the study period (Oral or sc), identifies a patient with “Suicidal Behavior Post Dosing” for that period. Same derivation will be done also for Suicidal Ideation post Dosing” for both the oral and sc periods. A patient identified with either “Suicidal Behavior Post Dosing” or with “Suicidal Ideation Post Dosing” is also classified with “Suicidal Behavior or Ideation Post Dosing”, the derivation should be done for both periods.

Frequency counts and percentages of the C-SSRS outcomes: Suicidal Behavior at Screening/Baseline, Suicidal Ideation at Screening/Baseline, Suicidal Behavior or Ideation at Screening/Baseline, Suicidal Behavior Post Dosing, Suicidal Ideation Post Dosing, Suicidal Behavior or Ideation Post Dosing, and shifts from screening/baseline for Stage 1/Stage 2 will be summarized.

8.12. Abnormal Involuntary Movement Scale (AIMS)

The AIMS will be performed at the time points specified in Table 1 and Table 2 of the study protocol. The AIMS scores the occurrence of tardive dyskinesia in patients receiving neuroleptic medications. The AIMS is a 14-item scale that includes assessments of orofacial movements, extremity and truncal dyskinesia, examiner's judgment of global severity, subjective measures of awareness of movements and distress, and a yes/no assessment of problems concerning teeth and/or dentures. Higher scores indicate greater severity of the condition.

AIMS total score will be calculated as a sum of items 1 through 7.

The AIMS total score and the individual score for each of the AIMS items 8-10 and changes from screening during Stage 1 will be presented using descriptive statistics. Shift from screening analysis of AIMS items 11-14 during Stage 1 will be presented.

The AIMS total score and the individual score for each of the AIMS items 8-10 and changes from baseline during Stage 2 will be presented using descriptive statistics. Shift from baseline analysis of AIMS items 11-14 during Stage 2 will be presented.

Enrolled analysis set and eITT analysis set will be used as appropriate.

8.13. Simpson-Angus Scale (SAS)

The SAS will be performed at the time points specified in Table 1 and Table 2 of the study protocol. The SAS is a 10-item instrument for the assessment of neuroleptic-induced parkinsonism. The items on the scale include measurements of hypokinesia, rigidity, glabellar reflex, tremor, and salivation. Each item is rated on a 5-point scale (0 to 4), with a higher score indicating greater severity of symptoms. The mean score is calculated by adding the individual item scores and dividing by 10.

The SAS mean score and changes from screening during Stage 1 will be presented using descriptive statistics by enrolled analysis set and all patients in safety analysis set, as appropriate.

The SAS mean score and changes from baseline during Stage 2 will be presented using descriptive statistics by treatment group.

8.14. Barnes Akathisia Rating Scale (BARS)

The BARS will be performed at the time points specified in Table 1 and Table 2 of the study protocol. The BARS is an instrument that assesses the severity of drug-induced akathisia. The BARS includes 3 items for rating objective restless movements, subjective restlessness, and any subjective distress associated with akathisia that are scored on a 4-point scale of 0 to 3, and summed yielding a total score ranging from 0 to 9. The BARS also includes a global clinical assessment of severity scored on a scale of 0 to 5. Higher scores are indicative of greater severity of akathisia.

BARS total score, global clinical assessment of severity values, and changes from baseline to each visit will be summarized using descriptive statistics - during Stage 1 by enrolled analysis set and all patients in safety analysis set, as appropriate, and during Stage 2 by treatment group.

8.15. Calgary Depression Scale for Schizophrenia (CDSS)

The CDSS will be performed at the time points specified in Table 1 and Table 2 of the study protocol. The CDSS is specifically designed to assess the level of depression separate from the positive, negative, and EPS in schizophrenia. This clinician-administered instrument consists of 9 items, each rated on a 4-point scale from 0 (absent) to 3 (severe) that are added together to form the patient's CDSS depression score.

Descriptive statistics of CDSS depression score values and changes will be presented - during Stage 1 by enrolled analysis set and all patients in safety analysis set, as appropriate, and during Stage 2 by treatment group.

8.16. Clinical Global Impression-Severity of Suicidality (CGI-SS)

The CGI-SS scale provides an overall clinician-rated assessment of the risk of suicidality. The CGI-SS consists of a 5-point scale in Part 1 ranging from 1 (not at all suicidal) to 5 (attempted suicide) and a 7-point scale in Part 2 ranging from 1 (very much improved) to 7 (very much worse).

The CGI-SS will be assessed at screening and all visits during the study.

Descriptive statistics of CGI-SS Part 1 and 2 will be presented - during Stage 1 by enrolled analysis set and all patients in safety analysis set, as appropriate, and during Stage 2 by treatment group.

8.17. Clinical Global Impression- of Severity (CGI-S)

The CGI-S will be administered by the investigator/trained rater at all in-clinic visits. The CGI-S scale was developed to rate the severity of a patient’s condition on a 7-point scale ranging from 1 (normal, not at all ill) to 7 (among the most extremely ill patients).

Descriptive statistics of CGI-S Part 1 and 2 will be presented - during Stage 1 by enrolled analysis set and all patients in safety analysis set, as appropriate, and during Stage 2 by treatment group.

8.18. Structured Clinical Interview for DSM-5 (SCID-5)

The Structured Clinical Interview for DSM-5 (SCID-5) is a semi-structured interview guide for making DSM-5 diagnoses. It will be administered at screening by a clinician or trained mental health professional that is familiar with the DSM-5 classification and diagnostic criteria.

The SCID-5 will be presented in listings.

9. TOLERABILITY VARIABLES AND ANALYSIS

9.1. Assessment of Local Tolerability and Pain

In case an adverse event related to an injection site reaction is reported, an assessment of the sc injection site (ie, local tolerability [skin at injection site]) will be made. The presence and severity of erythema, swelling, induration, and pain at the injection site may be assessed using the scales described in Section 7.8. of the protocol.

Data will be summarized as follows:

- Descriptive analysis of frequency counts and percentages of the initial patient's *maximal* (per injection type) scoring of erythema, edema, induration and nodule assessment, and of injection pain intensity assessment using the Numeric Pain Rating Scale (NPRS);
- Descriptive analysis of patient's *mean* time to stabilization, namely until the score of 0, of erythema, edema, and of induration and nodule assessment.

This analysis will be presented by both

- treatment group and the injection type (placebo or TV-46000), thus dividing the patients into non-intersecting groups, and distinguishing between the 2 types of injections for patients randomized to the TV-46000 q2m group, and by low (≤0.3ml) and high (>0.3ml) volume (the cut off is equivalent to the 100 mg dosage, which is the only overlapping dose between the 2 regimens, see Table 5);
- injection type (placebo or TV-46000 dose) – injection type will be identical to the treatment arm for the placebo/TV-46000 q1m patients, while the TV-46000 q2m patients will appear under 2 categories.

Table 5: Oral Risperidone Doses and Corresponding TV-46000 Doses

	Oral Risperidone Doses and Corresponding TV-46000 Doses			
	2 mg/day (mL)	**3 mg/day (mL)**	**4 mg/day (mL)**	**5 mg/day (mL)**
TV-46000 q1m	[redacted]	[redacted]	[redacted]	[redacted])
TV-46000 q2m	[redacted]	[redacted]	[redacted]	[redacted]

NOTE: The administered volume corresponding to each dose is shown in parentheses next to it.

9.2. All-cause Discontinuation Rate Assessment

All-cause discontinuation rates and discontinuation rates due to adverse events (dropout rates) will be calculated as the number of patients who withdrew early for all reasons, and the number of patients who withdrew early due to adverse events, respectively, divided by number of patients in each treatment group, and will be analyzed using descriptive statistics.

Time to all-cause discontinuation will be calculated as the discontinuation date minus the randomization date plus 1. The date of last contact will be used for the lost to follow-up patients.

Kaplan Meier curves for the time to discontinuation as a result of all causes will be plotted by treatment group.

The ITT analysis set will be used for the analysis. Separate summaries for adolescent patients may be presented separately for some analyses, according to the total number of the adolescents in the study.

[redacted]

11. PHARMACOKINETIC/PHARMACODYNAMIC ANALYSIS

[redacted]

12. BIOMARKER AND PHARMACOGENETICS ANALYSIS

[redacted]

13. PLANNED INTERIM ANALYSIS

There will be no interim analysis in this study.

14. STATISTICAL SOFTWARE

All data listings, summaries, and statistical analyses will be generated using SAS® version 9.4 or later.

15. CHANGES TO ANALYSES SPECIFIED IN THE STUDY PROTOCOL

15.1. [redacted]

[redacted]

15.2. Analysis due to the COVID-19 Pandemic

In Appendix N of the protocol, a general statement was added that supplementary and sensitivity analyses will be conducted to evaluate the impact of the change to remote monitoring (VC visits) and the impact of the COVID-19 pandemic on the impending relapse and rating scales, and that details of these analyses will be presented in the statistical analysis plan

Further details on the supplementary analysis, sensitivity analyses were added to the SAP throughout the document, as applicable.

16. REFERENCES

Andreasen NC, Carpenter WT, Kane JM, Lasser RA, Marder SR, Weinberger DR. Remission in schizophrenia: proposed criteria and rationale for consensus. Am J Psychiatry 2005;62:441-9.

Bartlett Jonathan. London School of Hygiene & Tropical Medicine [Internet]. DIA working group, Imputation based approaches. Multiple imputation for time to event data under Kaplan-Meier, Cox or piecewise-exponential frameworks – SAS® macros [updated: 02 April 2019; cited 01 November 2019]. Available from: http://www.missingdata.org.uk/. The macros can be downloaded from: http://missingdata.lshtm.ac.uk/files/2019/04/Package_Release_V3-final.zip

EAST 6 (Version 6.3) manual, Cytel Statistical Software & Services, Cytel Inc. Volume 10, Page 1931, 11Jul2014.

Kane JM, Sanchez R, Perry PP, Jin N, Johnson BR, Forbes RA, et al. Aripiprazole intramuscular depot as maintenance treatment in patients with schizophrenia: a 52-week, multicenter, randomized, double-blind, placebo-controlled study. J Clin Psychiatry 2012;73(5):617-24.

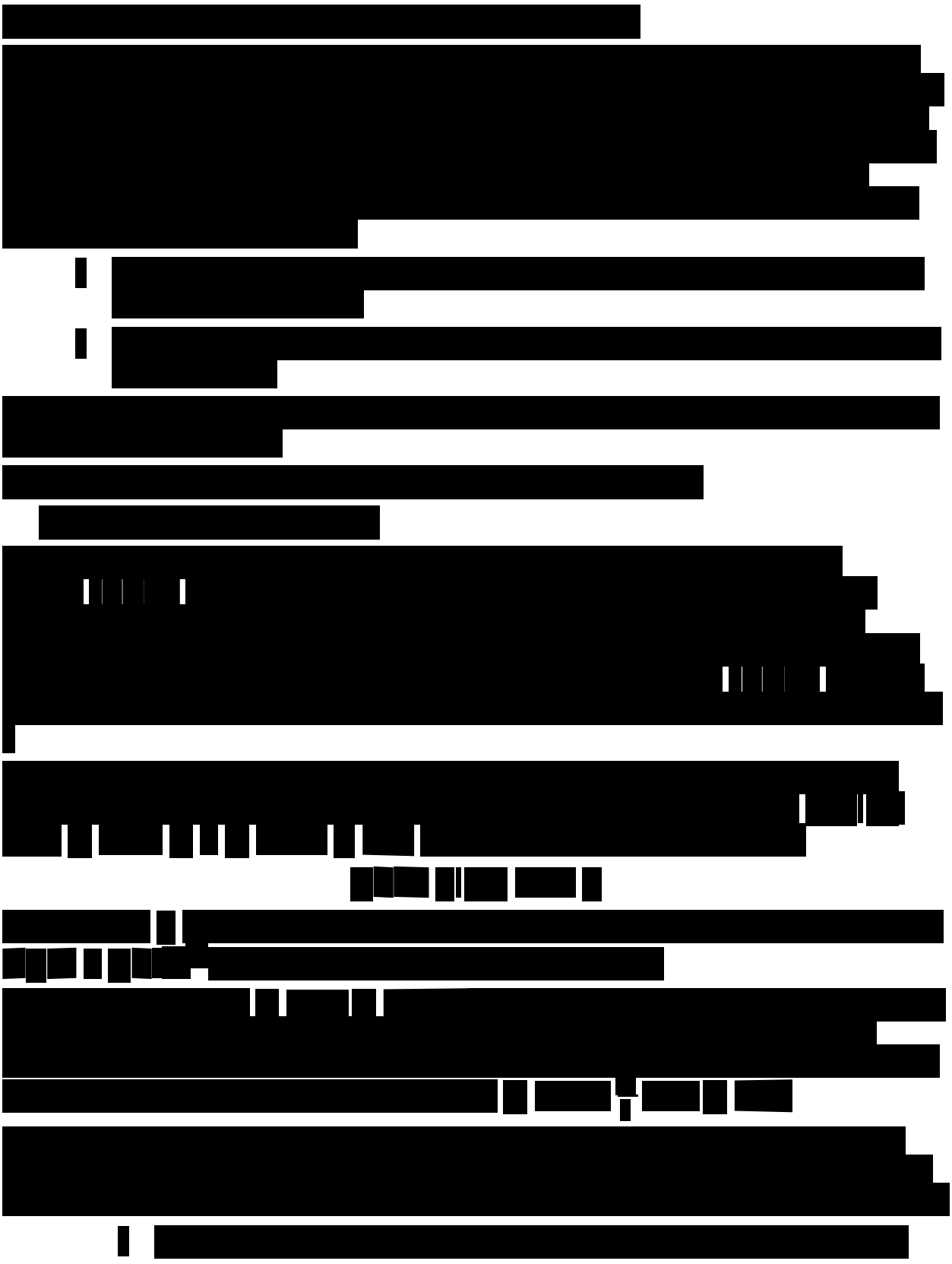
Leucht S, Kissling W, Davis JM. The PANSS should be rescaled. Schizophr Bull. 2010;36(3):461–2.

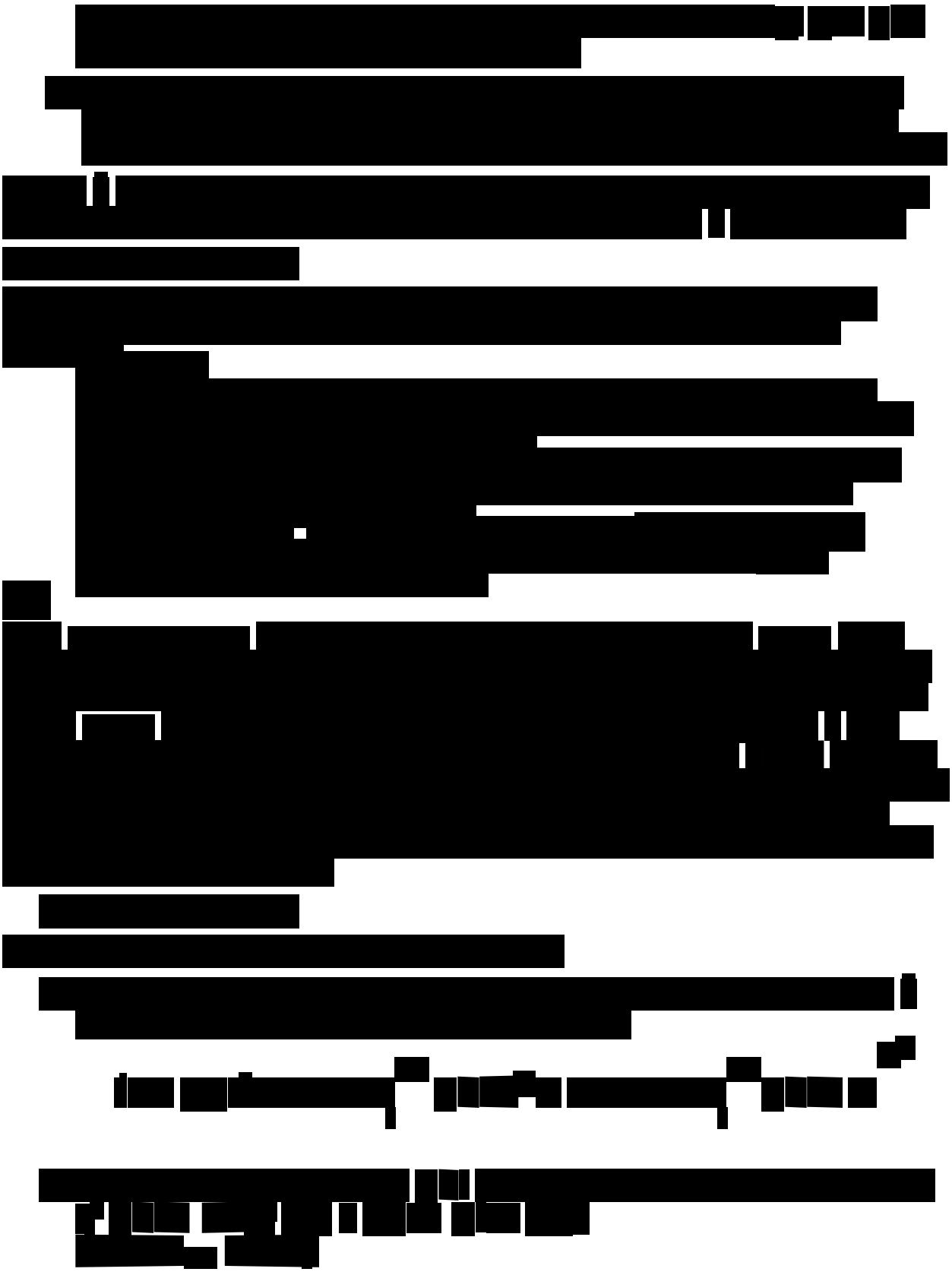
Lipkovich I, Ratitch B, O'Kelly M. Sensitivity to censored-at-random assumption in the analysis of time-to-event endpoints. Pharm Stat. 2016;15:216-29.Peto R, Pike MC, Armitage P, et al. "Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design". Br J Cancer. 1976;34(6):585–612.

Parfionovas A, Yang P, Hung J. Food and Drug Administration (FDA): Center for Drug Evaluation and Research; 2011. Aripiprazole (Abilify) Application 202971Orig1s000. Statistical Review and Evaluation. 27 pages.

APPENDIX A. [redacted]

[redacted]





[redacted]

3. [redacted]

[redacted]

iv. Combining Results using Rubin's Rule

[redacted]

```
modeleffects T;
```

[redacted]

d. [redacted]

APPENDIX B. LIST OF CONCOMITANT MEDICATIONS THAT MAY BE RELATED TO INCREASED RISK OF RELAPSE

Antipsychotics, sleep aids and antidepressants that may potentially indicate a need for symptoms reduction are listed below. Patients that withdrew from the study and used these medication might be at increased risk for relapse, and therefore might be used, if possible, in the imputation model. This list is not exhaustive and may be updated at the Statistical Data Review meeting prior to data closure.

Medication Class	Standardized Medication Name
ANTIEPILEPTICS	LORAZEPAM
PSYCHOANALEPTICS	AMITRIPTYLINE
PSYCHOANALEPTICS	AMITRIPTYLINE HYDROCHLORIDE
PSYCHOANALEPTICS	BUPROPION
PSYCHOANALEPTICS	BUPROPION HYDROCHLORIDE
PSYCHOANALEPTICS	CITALOPRAM
PSYCHOANALEPTICS	CITALOPRAM HYDROBROMIDE
PSYCHOANALEPTICS	DOXEPIN
PSYCHOANALEPTICS	DULOXETINE
PSYCHOANALEPTICS	DULOXETINE HYDROCHLORIDE
PSYCHOANALEPTICS	ESCITALOPRAM
PSYCHOANALEPTICS	ESCITALOPRAM OXALATE
PSYCHOANALEPTICS	FLUOXETINE
PSYCHOANALEPTICS	FLUOXETINE HYDROCHLORIDE
PSYCHOANALEPTICS	MIRTAZAPINE
PSYCHOANALEPTICS	PAROXETINE
PSYCHOANALEPTICS	PAROXETINE HYDROCHLORIDE
PSYCHOANALEPTICS	SERTRALINE
PSYCHOANALEPTICS	SERTRALINE HYDROCHLORIDE
PSYCHOANALEPTICS	TRAZODONE
PSYCHOANALEPTICS	VENLAFAXINE
PSYCHOANALEPTICS	VENLAFAXINE HYDROCHLORIDE
PSYCHOANALEPTICS	VILAZODONE
PSYCHOANALEPTICS	VILAZODONE HYDROCHLORIDE
PSYCHOANALEPTICS	VORTIOXETINE
PSYCHOLEPTICS	ALPRAZOLAM
PSYCHOLEPTICS	AMISULPRIDE

Medication Class	Standardized Medication Name
PSYCHOLEPTICS	ARIPIPRAZOLE
PSYCHOLEPTICS	ASENAPINE MALEATE
PSYCHOLEPTICS	BREXPIPRAZOLE
PSYCHOLEPTICS	BUSPIRONE
PSYCHOLEPTICS	BUSPIRONE HYDROCHLORIDE
PSYCHOLEPTICS	CARBAMAZEPINE
PSYCHOLEPTICS	CARIPRAZINE
PSYCHOLEPTICS	CARIPRAZINE HYDROCHLORIDE
PSYCHOLEPTICS	CHLORPROTHIXENE
PSYCHOLEPTICS	CHLORPROTHIXENE HYDROCHLORIDE
PSYCHOLEPTICS	CLONAZEPAM
PSYCHOLEPTICS	CLOZAPINE
PSYCHOLEPTICS	DIAZEPAM
PSYCHOLEPTICS	DULOXETINE HYDROCHLORIDE
PSYCHOLEPTICS	ESZOPICLONE
PSYCHOLEPTICS	FLUOXETINE
PSYCHOLEPTICS	FLUPENTIXOL DECANOATE
PSYCHOLEPTICS	FLUPHENAZINE
PSYCHOLEPTICS	FLUPHENAZINE DECANOATE
PSYCHOLEPTICS	HALOPERIDOL
PSYCHOLEPTICS	HALOPERIDOL DECANOATE
PSYCHOLEPTICS	HYDROXYZINE
PSYCHOLEPTICS	HYDROXYZINE EMBONATE
PSYCHOLEPTICS	HYDROXYZINE HYDROCHLORIDE
PSYCHOLEPTICS	ILOPERIDONE
PSYCHOLEPTICS	LEVOSULPIRIDE
PSYCHOLEPTICS	LITHIUM
PSYCHOLEPTICS	LITHIUM CARBONATE
PSYCHOLEPTICS	LORAZEPAM
PSYCHOLEPTICS	LURASIDONE
PSYCHOLEPTICS	LURASIDONE HYDROCHLORIDE
PSYCHOLEPTICS	MELATONIN
PSYCHOLEPTICS	OLANZAPINE
PSYCHOLEPTICS	PALIPERIDONE

Medication Class	Standardized Medication Name
PSYCHOLEPTICS	PALIPERIDONE PALMITATE
PSYCHOLEPTICS	PAROXETINE
PSYCHOLEPTICS	PERPHENAZINE
PSYCHOLEPTICS	QUETIAPINE
PSYCHOLEPTICS	QUETIAPINE FUMARATE
PSYCHOLEPTICS	RISPERIDONE
PSYCHOLEPTICS	SERTRALINE HYDROCHLORIDE
PSYCHOLEPTICS	TASIMELTEON
PSYCHOLEPTICS	TEMAZEPAM
PSYCHOLEPTICS	THIORIDAZINE
PSYCHOLEPTICS	TIOTIXENE
PSYCHOLEPTICS	VALPROATE SEMISODIUM
PSYCHOLEPTICS	VENLAFAXINE
PSYCHOLEPTICS	ZIPRASIDONE
PSYCHOLEPTICS	ZIPRASIDONE HYDROCHLORIDE
PSYCHOLEPTICS	ZOLPIDEM
PSYCHOLEPTICS	ZOLPIDEM TARTRATE
PSYCHOLEPTICS	ZUCLOPENTHIXOL
PSYCHOLEPTICS	ZUCLOPENTHIXOL HYDROCHLORIDE