STATISTICAL ANALYSIS PLAN

Taisho Pharmaceutical R&D Inc.
TS121-US201

A Multicenter, Randomized, Double-blind, Placebo-controlled Study of the Safety and Efficacy of TS-121 as an Adjunctive Treatment for Patients with Major Depressive Disorder with an Inadequate Response to Current Antidepressant Treatment

Version: Final 2.0
Date: 18 Dec 2018

Revision history

Draft 1:	30 Jun 2017	
Draft 2:	08 Aug 2017	
Draft 3:	16 Oct 2017	
Draft 4:	04 Jan 2018	
Final 1:	21 Feb 2018	
Final 2:	18 Dec 2018	update per amended protocol, address missing values and local hematology results

SIGNATURE PAGE - TAISHO PHARMACEUTICAL R&D INC.

Declaration

Upon review, the undersigned approves the statistical analyses plan. The analysis methods and data presentation are acceptable.

[redacted]	Date (dd mmm yyyy)

SIGNATURE PAGE - [redacted]

Declaration

This Statistical Analysis Plan is authored and issued by:

[redacted]	Date (dd mmm yyyy)

TABLE OF CONTENTS

List of Tables

List of Figures

ABBREVIATIONS AND DEFINITIONS

Term	Definition
ADT	Antidepressant therapy
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATRQ	Antidepressant Treatment Response Questionnaire
BLQ	Below the lower limit of quantification
BMI	Body mass index
bpm	Beats per minute
BUN	Blood urea nitrogen
C	Cortical cataract
CDISC	Clinical Data Interchange Standards Consortium
CFR	Code of federal regulations
CGI-I, S	Clinical global impression- Improvement, -Severity
CI	Confidence interval
Cl	Chloride
CPK	Creatine phosphokinase
CRF	Case report form
CS	Clinically significant
CSP	Clinical study protocol
CSR	Clinical study report
C-SSRS	Columbia suicidality severity rating scale
CV	Coefficient of variation
ECG	Electrocardiogram
eDISH	Evaluation of drug-induced serious hepatoxicity
EOCS	Event of clinical significance
EOT	End of treatment
ET	Early termination
ETDRS	Early treatment diabetic retinopathy study
FDA	Food and drug administration
GGT	Gamma glutamyltransferase
HAM-A	Hamilton anxiety scale
HAM-D	17 items Hamilton depression scale
HDL	High-density lipoprotein
HIV	Human immunodeficiency virus
HPA	Hypothalamus-pituitary-adrenal

HS-CRP	High sensitivity C-reactive protein
IL	Interleukin
IRT	Interactive response technology
IWRS	Interactive web response system
K	Potassium
kg	Kilograms
LDH	Lactate dehydrogenase
LDL	Low-density lipoprotein
LLOQ	Lower limit of quantification
LOCS	Lens opacities classification system
LogMAR	Logarithm of the minimum angle of resolution
MADRS	Montgomery-Asberg depression rating scale
MAR	Missing at random
MCH	Mean corpuscular hemoglobin
MCHC	Mean corpuscular hemoglobin concentration
MCV	Mean corpuscular volume
MDD	Major depressive disorder
MDMA	3,4-methylenedioxy-N-methylamphetamine
MedDRA	Medical dictionary for regulatory activities
MGH	Massachusetts general hospital
MI	Multiple imputation
MINI	MINI-International neuropsychiatric Interview
mg	Milligrams
Mg	Magnesium
mL	Milliliter
mmHg	Millimeter of mercury
MMRM	Mixed model repeated measures analysis
msec	Milliseconds
Na	Sodium
NC	Nuclear color
NCS	Not clinically significant
NK	Not known
NO	Nuclear opalescence
P	Posterior subcapsular cataract
PCP	Phencyclidine
PD	Pharmacodynamics
PGx	Pharmacogenomics
PK	Pharmacokinetics

PT	Preferred term
QTcB	Bazett's corrected QT interval
QTcF	Fridericia's corrected QT interval
RBC	Red blood cell
RDW	Red cell distribution width
SAE	Serious adverse event
SAFER	State versus trait; Assessable; Face validity; Ecological validity; and Rule of three Ps (pervasive, persistent, and pathological)
SAP	Statistical analysis plan
SD	Standard deviation
SDQ	Symptoms of depression questionnaire
SDTM	Study Data Tabulation Model
SE	Standard error
sec	Seconds
SGOT	Serum glutamic-oxaloacetic transaminase
SGPT	Serum glutamic-pyruvic transaminase
SI	Standard international
SNRI	Serotonin norepinephrine reuptake inhibitor
SOC	System organ class
SSRI	Selective serotonin reuptake inhibitor
TBL	Total bilirubin
TEAE	Treatment emergent AE
THC	Tetrahydrocannabinol
[redacted]	[redacted]
TNF-α	Tumor necrosis factor alpha
WBC	White blood cell
WHO-DD	World Health Organization - Drug Dictionary
WHODDE	World Health Organization - Drug Dictionary Enhanced

Statistical Analysis Plan

The Statistical Analysis Plan (SAP) details the statistical methodology to be used in analyzing study data and outlines the statistical programming specifications, tables, figures, and listings. It describes the variables and populations, anticipated data transformations and manipulations, and other details of the analyses not provided in the clinical study protocol (CSP).

The analyses described are based on the CSP Amendment 2 dated Mar 7 2017, Amendment 3 dated Oct 18 2017 and Amendment 4 dated Apr 23 2018. Table and listing shells, in an appended document, are based on the blank case report form (CRF) with date 11-Jul-2017 [Study Version 4.1 Date 28 Aug 2018]. Statistical programming will be performed using Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM) data mapped as defined in the Data Management Plan (Version 1.0, 7 April 2017 date of final signature) and mapping specification document.

The SAP describes the statistical analysis as it is foreseen when the study is being planned and will be finalized prior to database lock. Any revisions (both alternative and additional methods) to the SAP that are used in the final report, and reasons for such revisions, will be described in the final report. This is a pilot study and the results obtained will be used to estimate the effect size and to support future studies.

Amended Statistical Analysis Plan

The CSP Amendment 4 included changes to patient inclusion and exclusion criteria and to prohibited/ restricted concomitant therapy. Data collection and reporting specifications were updated for worsening of depression, medication error or potential overdose and events of clinical significance. Changes between CSP Amendment 3 to 4 with specific impact in data presentation were

- Section 9.2.7.8 abnormal hematology, biochemistry or urinalysis results may be repeated at the investigator's discretion to rule out laboratory error, or to confirm patient eligibility or safety.
- Section 10.5.2.1 on the primary endpoint (change in MADRS score), as assessed at Week 6, a statistically significant GROUP*TIME interaction effect, with greater reductions in MADRS scores in the TS-121 groups, will be interpreted as indicating efficacy (GROUP: Placebo, TS-121 10 mg and TS-121 50 mg, TIME: Week 2, Week 4 and Week 6)

- Section 11.3 Event of Clinical Significance (EOCS): Suicidal ideation and behavior, and lens opacity are considered as EOCS
- Sections 8.5 and 11.4 Patients with a CGI-I score of 6 or 7 at any time during the study will also be evaluated for possible discontinuation from the study. In the case of CGI-I scores of 6 or 7, the investigator will write a note justifying either discontinuation or continuation of the participant in the study. A CGI-I score of 6 or 7 at any time during the study should be considered as an AE of worsening of depression.

The amended SAP addresses missing data of two types: missing item responses for a given rating scale for calculation of the scale's total scores and missing visits data in the mixed model repeated measures analysis (MMRM). The amended SAP also describes presentation of local laboratory hematology data.

1. OVERVIEW OF THE STUDY

This multicenter study is comprised of a screening period of up to 4 weeks, with a 2-week single-blind placebo lead-in period followed by a 6-week randomized, double-blind, placebo-controlled treatment period with a 2-week follow-up period. The objectives are to compare the safety, efficacy, pharmacokinetics, and exploratory pharmacodynamics of TS-121 (drug substance code [redacted]) to placebo as an adjunctive treatment for patients with major depressive disorder (MDD) (Figure 1). Patients must remain on the same single antidepressant, selective serotonin reuptake inhibitor (SSRI), serotonin norepinephrine reuptake inhibitor (SNRI), or bupropion for 6 weeks, with a fixed dose for 4 weeks prior to screening. Patients will remain on this fixed dose of their current antidepressant therapy (ADT) throughout the course of the study.

Visit 1 (Screening): Patients will visit the site for screening assessments and an initial review of study requirements. In addition to confirming the diagnosis of MDD by MINI, a trained site rater will employ the 17-items Hamilton depression scale (HAM-D), which will be subsequently confirmed through an additional HAM-D assessment conducted by a central rater (Visit 1.1) prior to a 2-week placebo lead-in period.

Visit 1.1 (Screening): The central rater will contact the patient as scheduled, and will administer the Massachusetts General Hospital Antidepressant Treatment Response Questionnaire (MGH-ATRQ) for assessment of current ADT, adequacy of duration and dose of prior and current ADT, as well as degree of improvement. The central rater will also complete the HAM-D

assessment and inventory of symptoms: State versus Trait; Assessability; Face validity; Ecological validity; and Rule of three Ps (pervasive, persistent, and pathological) (SAFER) for confirmation of placebo lead-in eligibility. Results will be communicated to the site prior to Visit 2.

Visit 2 (Start of Placebo Lead-In): Following initial eligibility confirmation, qualified patients will return to the site initiating a 2-week, single-blind placebo lead-in period. An ophthalmologist will conduct a thorough eye examination (including slit lamp under mydriasis). Patients will schedule another phone call for central rater assessments (Visit 2.1), and Visit 3 will be scheduled.

Visit 2.1 (End of Placebo Lead-In): Response to placebo will be assessed using HAM-D by a central rater. The current HAM-D score and change in HAM-D score from Visit 1.1 to Visit 2.1 will be used to determine eligibility.

Visit 3 (Baseline): After completing the placebo lead-in period, patients will return to the site for baseline evaluations. Following confirmation of eligibility, patients will be randomly assigned to receive TS-121 10 mg, 50 mg, or placebo in a 1:1:1 ratio. Patients will be instructed to continue taking investigational product once-daily in the morning and to maintain their current ADT.

Visits 4 (Week 2) and 5 (Week 4): Patients will return to the site for safety and efficacy assessments, dosing compliance, drug accountability and dispensing of a new supply of investigational product. Compliance with the dosing schedule will be assessed. Visit 5 and Visit 6 will be scheduled.

Visit 6 (Week 6): Patients will return to the site for end-of-treatment safety (including ophthalmological examination), efficacy, pharmacokinetic, and exploratory pharmacodynamic assessments and final drug accountability.

End of Treatment (EOT): The EOT visit is Visit 6 for patients who complete treatment and is the last assessment closest to the last treatment day for patients who discontinued from treatment early or failed to complete Visit 6.

Visit 7 (Follow-up): Patients will return to the site for a follow-up evaluation of safety (including ophthalmological examination) and efficacy assessments approximately two weeks after Visit 6.

ET Follow-up: The visit completed approximately 14 days after the patient's last dose for patients who discontinued from treatment early.

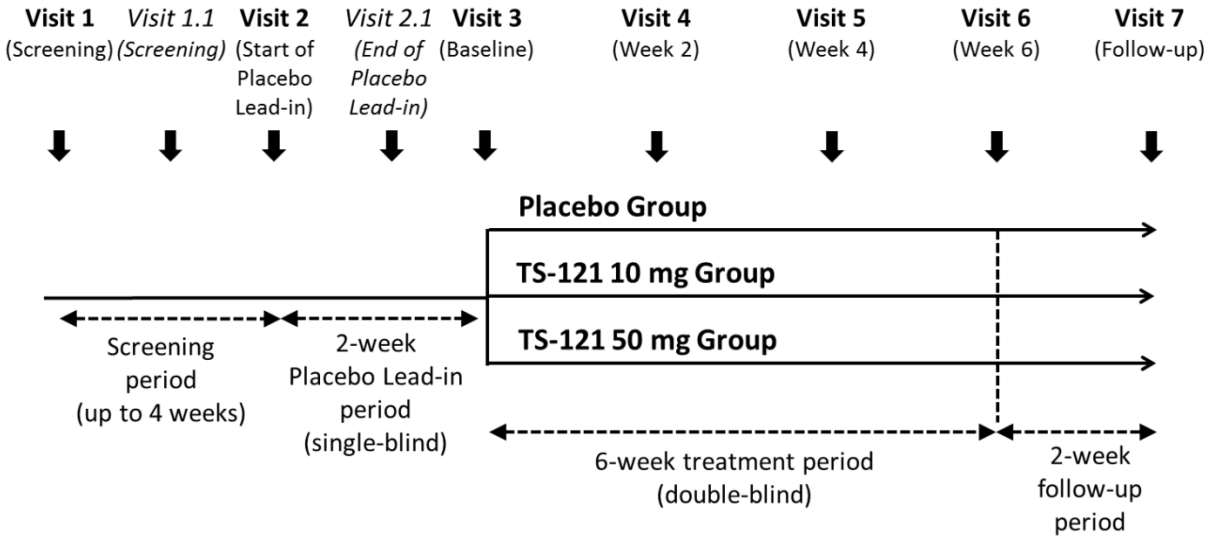

Figure 1. Study Design

2. STUDY OBJECTIVES AND ENDPOINTS

2.1 Primary Objectives

- To evaluate the safety and tolerability of TS-121 compared to placebo in MDD patients
- To evaluate the efficacy of two dose levels of TS-121 compared to placebo on MDD symptoms in MDD patients

2.2 Exploratory Objective

- To evaluate the effectiveness of biomarkers in predicting TS-121 response with retrospective stratification of patients via baseline values of HPA axis and inflammatory biomarkers

2.3 Primary Efficacy Endpoint

- Change in MADRS total score from Visit 3 (Baseline assessment) to Week 6/EOT

2.4 Secondary Efficacy Endpoint

- Change from Visit 3 to Week 6/EOT for CGI-S, SDQ, and HAM-A total scores

- Percentage of CGI-I improvers ("Very much improved" or "Much improved")
- Percentage of MADRS responders (≥ 50% reduction in total score)

2.5 Exploratory Endpoint

- Pharmacokinetic data from Visit 6 may be explored relative to study results (mental health assessments and/or laboratory values)
- Change from Visit 3 to Visit 6 for plasma copeptin and inflammatory markers (TNF-α, IL-1, IL-6, HS-CRP)
- To evaluate the effectiveness of biomarkers in predicting TS-121 response with retrospective stratification of patients via baseline values of HPA axis and inflammatory biomarkers consisting of:
    - Hair cortisol
    - Salivary cortisol
    - Urine cortisol
    - Plasma copeptin, and
    - Inflammatory markers (TNF-α, IL-1, IL-6, HS-CRP)

3. STUDY DESIGN

3.1 Study Procedures and Observations

Refer to Table 1.

Table 1 Study Procedures and Observations

Procedures	Visit 1**** (Screening) Day -28 to Day -7	Visit 1.1**** (Screening) Day -13 to Day -2	Visit 2 (Start of Placebo Lead-in) Day 1	Visit 2.1 (End of Placebo Lead-in) Day 13 (±2)	Screen* failure ONLY End of Study	Visit 3** (Baseline) Day 15 (+3)	Visit 4 (Week 2) Day 29 (±3)	Visit 5 (Week 4) Day 43 (±3)	Visit 6 (Week 6) Day 57 (±3)	Visit 7 (Follow-up) Day 71 (±3)	ET***	
											ET	Follo Up
Site Visit	X		X		X	X	X	X	X	X	X	X
Informed Consent (Study, PGx, HIPAA)[a]	X											
Demography	X											
Inclusion and Exclusion Criteria	X		X			X						
Evaluation of Medical and Psychiatric History	X		X									
Evaluation of Prior Medication History	X											
Serology Screen	X											
Urine Drug Screen	X					X						
Urine Pregnancy (all females)	X		X			X	X	X	X	X	X	X
Height	X											
Body Weight	X		X			X	X	X	X	X	X	X
Body Mass Index (BMI)	X											
Physical and Neurological Examination	X				X	X			X	X	X	X
Ophthalmological Examinations[b]			X						X	X	X	X
Columbia Suicide Severity Rating Scale (C-SSRS)	X				X	X	X	X	X	X	X	X
MINI	X											
HAM-D (SITE Rater)	X											
MGH ATRQ (Central Rater)		X										
HAM-D (Central Rater)		X		X								
SAFER Inventory (Central Rater)		X										
MADRS[c]						X	X	X	X	X		

Table 1 Study Procedures and Observations

Procedures	Visit 1**** (Screening) Day -28 to Day -7	Visit 1.1**** (Screening) Day -13 to Day -2	Visit 2 (Start of Placebo Lead-in) Day 1	Visit 2.1 (End of Placebo Lead-in) Day 13 (±2)	Screen* failure ONLY End of Study	Visit 3** (Baseline) Day 15 (+3)	Visit 4 (Week 2) Day 29 (±3)	Visit 5 (Week 4) Day 43 (±3)	Visit 6 (Week 6) Day 57 (±3)	Visit 7 (Follow-up) Day 71 (±3)	ET***	
											ET	Follow Up
MADRS Central Rater Oversight						X ←	—	—	—	→ X		
CGI-S						X	X	X	X	X		
CGI-I							X	X	X	X		
HAM-A						X	X	X	X	X		
SDQ						X	X	X	X	X		
Vital Signs (seated)	X		X			X	X	X	X	X	X	X
12-Lead Electrocardiogram (ECG)	X					X	X	X	X	X	X	X
Clinical Laboratory Tests	X					X	X	X	X	X	X	X
Exploratory Safety Assessments			X			X	X	X	X	X	X	X
Hair Cortisol Level						X						
Plasma Copeptin Level						X			X			
Salivary Cortisol Level (at home)[d]			Dispense Kit			X						
Urine Cortisol Level (pooled for 8 hr at home)[d]			Dispense Kit			X						
Inflammatory Markers (TNF-α, IL-1, IL-6, HS-CRP)						X			X			
Pharmacokinetic Blood Sampling ([redacted])									X			
Pharmacogenomic Blood Sampling						X						
Placebo administration under single-blind			X ←	—	—	→ X						
Randomization						X						
Investigational product dispensing, accountability and compliance check[e]			X			X	X	X	X			
Dose administration under double-blind						X ←	—	—	→ X			
Drug adherence (AiCure)[f]			X ←	—	—	—	—	—	→ X			
Adverse Event Inquiry & Reporting			X ←	—	—	—	—	—	—	→ X	X	X

Table 1 Study Procedures and Observations

Procedures	Visit 1**** (Screening) Day -28 to Day -7	Visit 1.1**** (Screening) Day -13 to Day -2	Visit 2 (Start of Placebo Lead-in) Day 1	Visit 2.1 (End of Placebo Lead-in) Day 13 (±2)	Screen* failure ONLY End of Study	Visit 3** (Baseline) Day 15 (+3)	Visit 4 (Week 2) Day 29 (±3)	Visit 5 (Week 4) Day 43 (±3)	Visit 6 (Week 6) Day 57 (±3)	Visit 7 (Follow-up) Day 71 (±3)	ET***	
											ET	Follow Up
Concomitant Medication Inquiry & Reporting	X ◄									► X	X	X

* End of Study (Screen failure only): Patients who discontinue between Visits 2 and 3 only will undergo the procedures specified above.

** Visit 3 must be performed after obtaining results for Visit 2.1.

*** The Early Termination (ET) visit should be performed as early as possible after discontinuation. ET follow-up visit will also be necessary approximately 14 days after the patient's last dose.

**** Visit 1 must be performed prior to Visit 1.1.

a Informed consent for PGx must be obtained prior to blood collection at Visit 3.

b Ophthalmological Examination: Must be conducted by an ophthalmologist. Examination can be performed within -5 days of study Day 1 (Day -5 to Day 1) as Visit 2 assessment.

c MADRS should be conducted prior to any other mental health assessments and safety/PK/PD assessments.

d Urinary and Salivary Cortisol: Dispense kits at Visit 2.

e Investigational product dispensing, accountability and compliance checks: Accountability and compliance checks will be performed at Visits 3 to 6. Administration of the investigational product for the placebo lead-in period will start on the morning following Visit 2 (study Day 2). Administration of the investigational product for the treatment period will start on the morning following Visit 3 (study Day 16).

f Drug Adherence: Daily dosing by patients of current ADT and investigational product (TS-121 or placebo) will be monitored using AiCure technology on the patient's smart phone or a device provisioned by the site (See CSP Section 8.1.3, Treatment Compliance).

4. STUDY POPULATION

It is estimated that the study will screen approximately 300 patients in order to randomize approximately 180 MDD patients.

Patients withdrawn for safety or tolerability will not be replaced. Patients discontinued due to other reasons, e.g., dosing compliance, may be replaced at the discretion of the sponsor in order to achieve the study objectives.

5. STATISTICAL BASIS FOR SAMPLE SIZE

The total number of patients enrolled in the study is based on practical consideration and not on statistical power calculations. These sample sizes are considered adequate to meet the study objectives, based on clinical considerations and other MDD trials, and will allow trends to be detected with regards to efficacy and PD effects of TS-121 in this patient population.

6. STATISTICAL ANALYSIS CONVENTIONS

6.1 Analysis Variables, Criteria for Evaluation

6.1.1 Demographic and Background Variables

The following patient history, demographic and anthropometric information will be recorded:

- Date of informed consent
- Age calculated as (date of informed consent – date of birth)/365.25. Date of birth will be collected using interactive response technology (IRT) of the IWRS.
- Gender
- Ethnicity
- Race
- Height and weight
- BMI

- Medical and/or psychiatric history (within the past 3 years)

  New or existing medical or psychiatric conditions reported prior to first administration of the investigational product [from the placebo lead-in period] will be captured as medical history, whereas any new, untoward, or worsening medical or psychiatric conditions reported after first administration of the investigational product during the placebo lead-in period will be captured as adverse events.

- Concomitant medications

  All medications used during the study will be recorded. Prior ADTs other than a current SSRI/SNRI/bupropion and prohibited medications that might influence HPA axis function used within 3 months of Visit 1 will also be recorded.

- Screening laboratory test results will be recorded
  - **Serology:** Hepatitis B surface antigen (HBsAg), HIV screening and serum antibodies against Hepatitis C virus (anti-HCV antibody)
  - **Urine drug screen:** The following controlled substances will be tested at Visit 1 and at Visit 3: amphetamine, barbiturates, benzodiazepines, cocaine, methadone, morphine/opiates, methamphetamine, marijuana (THC), phencyclidine (PCP), and ecstasy (MDMA).
  - **Pregnancy:** For all female participants, urine pregnancy samples will be tested at specified visits indicated in the Schedule of Procedures, Table 1.

All medical history will be coded using Version 19.1 of the Medical Dictionary for Regulatory Activities (MedDRA).

Concomitant medications will be coded using the World Health Organization Drug Dictionary Enhanced (WHODDE September 2016).

6.1.2 Screening Mental Health Assessments

6.1.2.1 MINI-International Neuropsychiatric Interview (MINI)

Results of the MINI (Version 7.0.2) administered at Visit 1 will be recorded in the patient's source records only.

6.1.2.2 17 items Hamilton Rating Scale (HAM-D)

Results of the HAM-D (October 2010 for the DUAG-8 project) conducted by the trained site rater at Visit 1 and of the remote HAM-D conducted by central raters at Visit 1.1 and Visit 2.1 will be captured.

6.1.2.3 SAFER Interview

The interview conducted at Visit 1.1 includes the SAFER inventory of symptoms, the HAM-D and the MGH-ATRQ. Previous and current ADT information collected using the MGH-ATRQ will be recorded with prior and concomitant medications. Start date of MDD current episode will be recorded.

6.1.3 Safety Variables

Safety and tolerability will be assessed from the time the patient is administered the investigational product during the placebo lead-in period through the final study procedure as specified in the Schedule of Procedures, Table 1.

Safety measures include:

- Adverse events
- Clinical laboratory tests (hematology, biochemistry and urinalysis)
- Vital signs
- 12-Lead ECG
- Body weight
- Routine physical examination
- Comprehensive neurological examination
- Ophthalmological examination
- Columbia Suicide Severity Rating Scales (C-SSRS)
- Exploratory Safety Assessments

6.1.3.1 Adverse Events

An adverse event (AE) is defined as any untoward medical occurrence in a patient administered an investigational product, and which does not necessarily have a causal relationship to investigational product. An AE can be any unfavorable or unintended sign (including an abnormal finding), symptom, or disease. AEs include any occurrence that is new in onset or aggravated in severity or frequency from the baseline condition, or abnormal results of diagnostic procedures (including clinically significant laboratory test abnormalities). AEs also include an intentional or accidental overdose, whether or not it is associated with symptoms. An overdose is defined as ingestion of the investigational product (accidental or intentional) more than 3 capsules in the same calendar day.

AEs will be recorded after the first administration of the investigational product in the placebo lead-in period through follow-up. AEs with onset date and time (if available) after the first administration of randomized treatment will be considered treatment emergent AE (TEAE).

Any AEs with incomplete start and end dates/times will be treated as follows:

- Adverse events with unknown start and/or end times (but where the date is known) will be imputed as post-dose for events on first treatment day of placebo lead-in or randomized treatment or with a time of 00:00 for other days. The missing time will be shown blank in the listings.
- Adverse events with incomplete start day and known month will be imputed as the later of: day of first placebo lead-in dose or first randomized dose started in that month or first of the month. Missing day will be printed -- in the listings.
- Adverse events with incomplete start dates missing month and day will be considered as the later of: day of first placebo lead-in dose or first randomized dose started in that year or January 1. Missing month and day will be printed ----- in the listings.
- Adverse events with completely unknown start dates will be considered as treatment emergent for the tabulations and the missing dates will be shown as --------- in the listings.

Details of AE classification for severity (mild, moderate or severe), action taken and relationship to test drug (not related, possibly related, probably related and definitely related) are provided in Appendix C. All AEs will be coded using MedDRA Version 19.1.

Serious Adverse Event

AEs that result in death, life-threatening, requiring inpatient hospitalization or prolongation of existing hospitalization, persistent or significant disability/incapacity, a congenital anomaly or birth defect, or any other important medical event, as defined in the CSP Section 11.2.1, will be recorded as serious adverse events (SAE).

Event of Clinical Significance

AEs that are suicidal ideation and behavior or lens opacity, as defined in the CSP Section 11.3, will be recorded as events of clinical significant (EOCS).

6.1.3.2 Clinical Laboratory Tests

Laboratory tests listed in the CSP Section 9.2.7.8 will be performed at times indicated in Schedule of Procedures, Table 1.

6.1.3.3 Vital Signs

Vital signs will include systolic and diastolic blood pressure (mmHg), pulse rate (bpm) and oral temperature (°C) and will be collected at specified time points indicated in the Schedule of Procedures, Table 1.

6.1.3.4 Electrocardiograms

Electrocardiograms will be assessed for

- Ventricular rate (beats per minute [bpm])
- RR interval (msec)
- PR interval (msec)
- QRS duration (msec)
- QT interval (msec)
- QT interval corrected using the Bazett correction formula (QTcB) (msec)
- QT interval corrected using the Fridericia correction formula (QTcF) (msec)

The ECG will be evaluated by the Investigator as ‘Normal’, ‘Abnormal, NCS’ or ‘Abnormal, CS’. Any abnormal CS ECG will be reported as an AE.

6.1.3.5 Weight

Body weight will be recorded in kilograms (kg) at specified time points indicated in the Schedule of Procedures, Table 1.

6.1.3.6 Physical and Neurological Examination

Physical and neurological examinations will be performed and recorded as described in the Schedule of Procedures, Table 1.

The results will be recorded in the patient's study source records. Any abnormalities noted prior to first administration of the investigational product will be listed as medical history. Findings after first administration of the investigational product that meet the criteria of an AE (CSP Section 11.1) will be captured as AEs.

6.1.3.7 Ophthalmological Examination

Ophthalmological examinations will include the following:

- Best corrected visual acuity as measured with the ETDRS chart, and scored in LogMAR
- Evaluation of the crystalline lens through a maximally dilated pupil using lens opacities classification system (LOCS III), including the following features;
    - Posterior subcapsular cataract (P)
    - Cortical cataract (C)
    - Nuclear opalescence (NO)
    - Nuclear color (NC)

Events of clinical significance (EOCS) for ophthalmological findings, as defined in CSP Section 11.3.2, will be recorded as AEs.

6.1.3.8 Columbia Suicidal Severity Rating Scale (C-SSRS)

C-SSRS (Screening, Version 1/14/09) which pertains to the previous 6 months will be administered at Visit 1, and then C-SSRS (Since Last Visit, Version 1/14/09) which pertains to changes since the last visit will be administered at Visits 3 through 7.

At a minimum, responses to questions 1 (Wish to be Dead), 2 (Non-Specific Active Suicidal Thoughts) and suicidal behaviors will be recorded.

The suicidal ideation score is calculated as the maximum score of responses to the first 5 questions.

- Wish to be dead (positive response is scored 1)
- Non-specific active suicidal thoughts (positive response is scored 2)
- Active suicidal ideation with any methods (not plan) without intent to act (positive response is scored 3)
- Active suicidal ideation with some intent to act, without specific plan (positive response is scored 4)
- Active suicidal ideation with specific plan and intent (positive response is scored 5)

The suicidal ideation score will be set Not Applicable (numeric score = 0) when Item 1 and Item 2 responses are negative. An EOCS for suicidal ideation and behavior, as defined in CSP Section 11.3.1, may be recorded for any patient who has a C-SSRS suicidal ideation score of 5 or a "yes" response for any of the C-SSRS suicidal behavior questions.

The suicidal ideation intensity rating will be calculated as the sum of numbers endorsed on 5 intensity items (Frequency, Duration, Controllability, Deterrents and Reasons for Ideation).

6.1.3.9 Exploratory Safety

Blood samples collected at visits indicated in the Schedule of Procedures will be stored at the central lab for possible exploratory safety assessments to be determined at a later stage based on the current understanding of toxicology.

6.1.4 Treatment Compliance

Compliance will be monitored using the AiCure daily record of expected and actual medication taken. Any record reflecting that the daily dose was greater than expected will be noted as an overdose event but will be treated as only the expected daily dose (3 capsules) for calculation of compliance in each visit interval.

Compliance over the treatment interval will be calculated as:

Compliance (%) = 100 * Total number of days kept compliance (took 3 capsules per day)/ number of expected dosing days (The day after V3 to V6/ET visit day).

Treatment compliance will also be calculated for placebo lead-in, baseline to Week 2, Week 2 to Week 4, and Week 4 to Week 6. Compliance will be calculated for any patient with at least one recorded dose in the interval.

6.1.5 Pharmacokinetic Variables and Pharmacogenomic Blood Sampling

A single venous blood sample will be collected for determination of [redacted] pharmacokinetics at Visit 6. Analysis of PGx data, if applicable, will be documented in a separate analysis plan.

6.1.6 Pharmacodynamic Variables

This study will evaluate blood, urine, saliva and hair samples at specified time points indicated in the Schedule of Procedures, Table 1 for exploratory PD analysis, the results of which may guide in predicting TS-121 response in future trials.

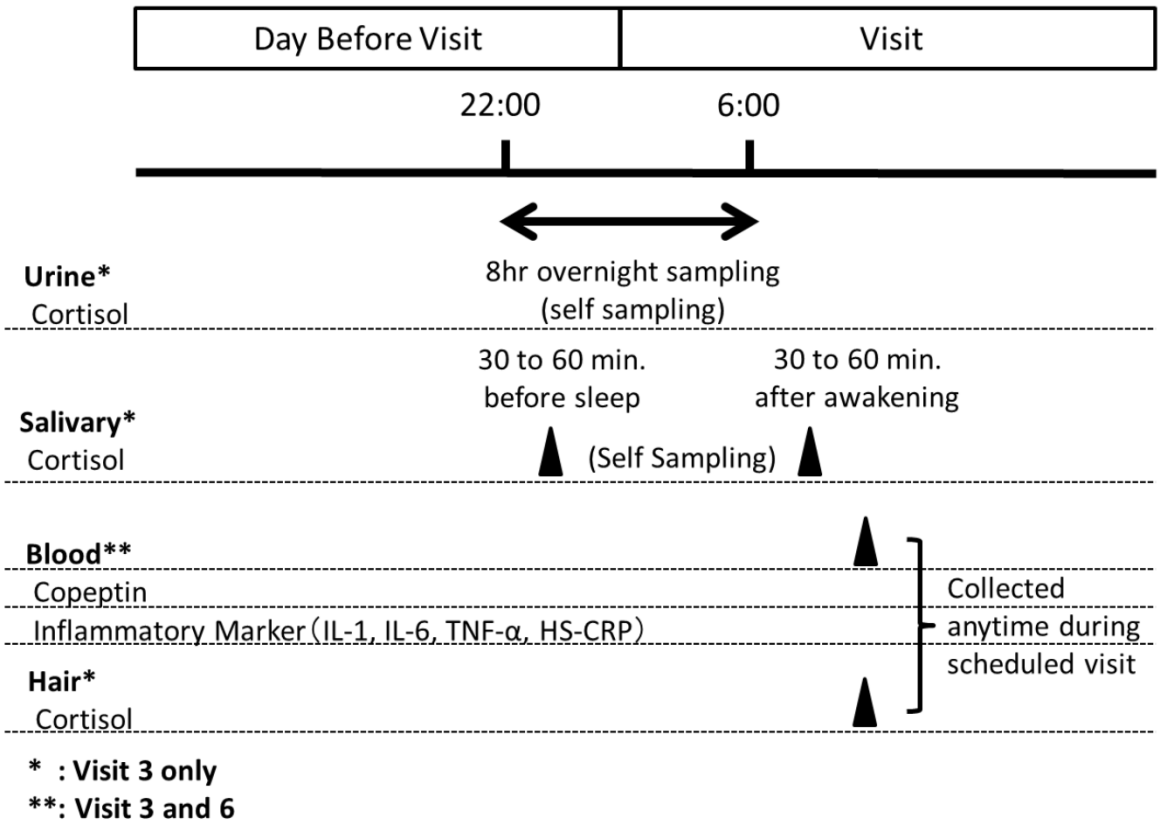

Figure 2. Pharmacodynamic Schedule

Samples for the following PD markers will be collected at visits indicated in Table 1 and Figure 2:

- Plasma Copeptin
- Salivary Cortisol
- Urine Cortisol
- Hair Cortisol
- Inflammatory Markers (TNF-α, IL-1, IL-6, HS-CRP)

6.1.7 Efficacy Variables

Efficacy assessments will be performed at visits indicated in Table 1. The time frame for Montgomery-Asberg Depression Rating Scale (MADRS), Clinical Global Impressions - Severity (CGI-S) and Improvement (CGI-I) Scales, Hamilton Anxiety Scale (HAM-A) and Symptoms of Depression Questionnaire (SDQ) pertains to the previous 7 days. Assessments will be administered by a trained site rater.

Questionnaire interviews will be captured by FDA 21 CFR Part 11 compliant electronic tablets. Results for each question of the HAM-D, MADRS, MGH-ATRQ and SAFER inventory will be captured through this tablet system.

6.1.7.1 Montgomery-Asberg Depression Rating Scale (MADRS)

Scoring information for MADRS is defined in Taisho TS121-US201 MADRS Version 1. Primary and secondary efficacy endpoints are based on MADRS total score change from baseline. The MADRS total score is calculated by summing the 10 item scores. The primary endpoint is change in MADRS total score from Visit 3 (Baseline) to Week 6/EOT and the secondary is percentage of MADRS responders, where a responder is defined as ≥ 50% reduction in total score from Baseline to Week 6/EOT.

6.1.7.2 Clinical Global Impressions - Severity (CGI-S) and Improvement (CGI-I) Scales

Scoring information for CGI-S and CGI-I is defined in Taisho TS121-US201 CGI-S Version 1 and Taisho TS121-US201 CGI-I Version 1, respectively. The secondary endpoint, percentage of CGI-I improvers, whereby a CGI-I improver is defined as patients with CGI-I score of 1 or 2 ("Very much improved" or "Much improved").

6.1.7.3 Hamilton Anxiety Scale (HAM-A)

Instructions for HAM-A administration and scoring are provided in HAM-A (SIGH-A, Version 5/92). The HAM-A total score change from Baseline to Week 6/EOT is a secondary endpoint.

6.1.7.4 Symptoms of Depression Questionnaire (SDQ)

Scoring information for SDQ is defined in Taisho TS121-US201 SDQ Version 1. The SDQ total score is calculated by summing the scores of all 44 items. The SDQ total score change from Baseline to Week 6/EOT is a secondary endpoint.

6.1.7.5 Handling of Individual Missing Items

In case ≥ 25% of items used to calculate a total score for a given scale are missing for one patient and time point then the total scale score for that patient and time point will be considered missing. If < 25% of items are missing then the total score will be calculated as the mean of non-missing items multiplied by the total expected number of items and rounded to the nearest integer.

6.2 Data Sets Analyzed

6.2.1 Safety Analysis Population

The safety analysis population will include all patients who received at least one dose of investigational product (including placebo during the placebo lead-in period).

6.2.2 Intent to Treat Population

The Intent to Treat Population will include all patients who were randomized and received at least one dose of randomized treatment.

6.2.3 Exploratory PD Analysis Population

The exploratory PD analysis population will include all randomized patients who completed the baseline assessment of HPA axis and inflammatory biomarkers and have sufficient mental health assessments completed. Missing data will not be imputed and are to be excluded from the analysis.

6.2.4 Pharmacokinetic (PK) Analysis Population

The PK analysis population will include all patients who were randomized and received randomized treatment with evaluable PK data and may exclude those that are considered invalid due to relevant missing values, or if any other problems occurring during sampling, laboratory analysis, dosing, or the occurrence of AEs that would invalidate the concentration measurements.

Any data excluded will be described in the clinical study report (CSR).

6.3 Statistical Analysis Methods

6.3.1 Listings and Descriptive Statistics

All original and derived parameters as well as population characteristics will be listed by subject number and described using summary statistics. Listings will be ordered by treatment group, subject number, date and time. Observations that are spurious (extreme relative to the majority of the data or difference from group median >1.5×[75^{th} – 25^{th} percentiles]) will be noted. These will not be altered or removed from any presentation of the data.

Treatment order in summary tables will, in general, be TS-121 10 mg Group, TS-121 50 mg Group and Placebo Group. As appropriate, Total TS-121, Total Placebo Lead-in and Total Screened will also be summarized. Summaries of scheduled events will also be ordered by study period: As grouped in Table 1, screening will include Visit 1 and Visit 1.1 and extend from Day -28 through Day -2, placebo lead-in will include Visit 2 and Visit 2.1 and extend from Day 1 to Day 13 (±2), treatment period will include Visit 3 (Baseline), Visit 4, Visit 5 and Visit 6 or EOT and follow-up will include any visit or study day after Week 6/EOT. Frequency counts (number of patients [n]) and percentages will be reported for each qualitative variable. Descriptive statistics (n, arithmetic mean [mean], standard deviation [SD], standard error [SE], minimum,

median and maximum) will be calculated for each quantitative variable (unless otherwise stated).

All statistical programming will use SDTM data and will be based on SDTM mapping rules for scheduled, unscheduled and repeat visits. No assessment or exclusions other than the mapping specifications will be performed for visits out of the planned visit window.

Listings will include repeat and unscheduled measurements.

The following rules will apply to repeat and unscheduled measurements in summary tables:

- No more than one assessment per patient per time point will be used in the summary of each scheduled safety or efficacy endpoint.
- If the repeated measurement occurs prior to the first dose of study drug for the treatment period, then the last obtained value of any repeated measurement will be used as baseline in the summary tables.
- If the repeated measurement occurs after the first dose of study drug for the treatment period, then the first obtained value of any repeated measurements will be used in the summary tables.

Unless specified otherwise, the baseline used for any change from baseline measurements will be the last measurement prior to the first dose of study drug for the treatment period.

6.3.2 Statistical Significance Level

Any statistical test or confidence interval (CI) will be two-sided and will be performed at the 5% level of significance, unless otherwise stated. No adjustment for multiple comparisons will be performed for calculation of CI or statistical significance.

6.3.3 Software

All statistical analyses will be performed using SAS® Version 9.3 and Enterprise Guide Version 7.15 or using SAS Version 9.4

6.3.4 Missing Data

Rules for incomplete AE start and end dates/times are described in Section 6.1.3.1. Rules for handling missing efficacy items in calculating total scores are described in Section 6.1.7.5. No imputation of missing data will be performed for safety data or for the exploratory PD analysis.

In case an unscheduled clinical laboratory assessment was performed within the visit window for a missing assessment then the unscheduled assessment will be included in the visit summary. In these cases the unscheduled assessment closest to the scheduled time will be used.

Missing data are assumed to be missing at random (MAR).

The primary endpoint for efficacy is Week 6/EOT. Efficacy assessments are not planned to be collected at early termination. The primary analysis will include all available efficacy assessments of the Intent to Treat Population. The primary analysis will use MMRM, in which patients with no post baseline assessment are dropped from the model. A sensitivity analysis will assess the Week 6/EOT of each efficacy endpoint (MADRS total score, CGI, HAM-A total score and SDQ total score) based on a multiple imputation method using all patients with a baseline score, including patients with no post baseline assessment.

6.3.5 Interim Analysis

No interim analysis is planned.

6.3.6 Blinded Efficacy Review

Preliminary blind review of efficacy will be conducted prior to database lock.

6.3.7 Randomization Schema and Codes

Subject number, randomization date and treatment will be listed [Listing 16.1.7].

6.3.8 Documentation of Statistical Methods

In addition to this SAP, the record of statistical model calculations will be included in Section 16.1.9 of the CSR.

6.3.9 Patient Disposition

A record of IRT data, including subject number, placebo lead-in date, randomization date, treatment, first and last dose dates, screen/placebo lead-in fail date and primary reason, the completion status, the date of discontinuation and the reason for discontinuation of all screened patients will be listed. Unblinding date and reason will be listed if any unblinding event occurs [Listing 16.2.1.1]. The number of patients randomized, received the randomized treatment and the frequency and percentage of patients completing the study, patients withdrawing early, and primary reason for withdrawal will be summarized by treatment and overall for all randomized

patients [Table 14.1.1.1]. The number of patients entered placebo lead-in period and randomized or failed will be summarized by treatment and overall [Table 14.1.1.2]. An overall summary table of patients who signed informed consent with numbers of patients who screen failed prior to or after entering placebo lead-in, and reasons for failures will also be presented [Table 14.1.1.3].

Disposition information will be presented in a flow diagram of number of patients screened, screen failures before start of placebo lead-in (Visit 2), placebo lead-in screen failures, patients randomized, patients dosed, reasons for any patient randomized but not dosed, total number of patients receiving double-blind treatment, number per treatment group starting double-blind treatment, number per treatment group of patients completing the study and summary counts, by treatment group, of the reasons that patients discontinued from the study [Figure 14.1.1].

6.3.10 Protocol Deviations and Analysis Populations

All protocol deviations will be recorded by the Investigator, incorporated to the mapped (SDTM) database and will be listed by subject number [Listing 16.2.2.1].

Assignment to analysis populations and data exclusions for individual assessments and time points will be listed by subject number for the safety analysis population [Listing 16.2.3.1].

6.3.11 Demographic Data and Baseline Data

6.3.11.1 Demographics

Demographics and baseline characteristics including date of informed consent, date of birth, age at informed consent date, sex, race, ethnicity, screening height, weight and BMI will be listed for the safety analysis population [Listing 16.2.4.1]. Age, sex, race, ethnicity, and screening height, weight and BMI will be summarized by treatment and overall for the safety analysis population [Table 14.1.2.1].

6.3.11.2 Hamilton Depression Scale (HAM-D)

The HAM-D total score and item scores from site and central raters will be listed for all screened patients [Listing 16.2.4.2.1]. HAM-D total score at Visits 1, 1.1 and 2.1 for screening will be summarized by treatment and overall for the Intent to Treat Population [Table 14.1.2.2]. It will be noted that the qualifying HAM-D total scale was changed in CSP amendment 4 from HAM-D $\geq$ 21 to HAM-D $\geq$ 18 and Visit 1 and 1.1.

6.3.11.3 Baseline Major Depressive Disorder Characteristics

Major depressive episode records from medical history will be listed with the calculated duration (months) of current depressive episode and duration from first MDD Diagnosis (years) [Listing 16.2.4.2.2] and presented in a summary table for the intent to treat population [Table 14.1.3]. The summary table will also present baseline MADRS total score and baseline HAM-A total score.

6.3.11.4 Medical and Psychiatric History

Any findings for medical and psychiatric history will be coded using MedDRA Version 19.1 and will be listed by subject number for the safety analysis population. MedDRA system organ class (SOC) and preferred term (PT) where a condition or abnormality has been reported, will be listed for the safety analysis population [Listing 16.2.4.3].

Medical history and psychiatric history will be presented in summary tables by treatment group and overall by SOC and PT for the safety analysis population [Table 14.1.4].

6.3.11.5 Concomitant Medication

All concomitant medications for patients in the Safety Analysis Population will be listed with the identification number of any associated medical history or adverse event for the safety analysis population [Listing 16.2.4.4]. Prior and concomitant ADTs will be included in this listing and also listed separately [Listing 16.2.4.5]. The record of medications includes reported medication and WHO-DD preferred term. Concomitant medications and indications will be summarized by preferred term by treatment group and overall for the safety analysis population [Table 14.1.5]. Duration of current ADT will be summarized with baseline MDD characteristics [Table 14.1.3].

6.3.11.6 Screening Laboratory Tests

Drug screening [Listing 16.2.4.6.1], screening serology [Listing 16.2.4.6.2] and pregnancy test results [Listing 16.2.4.6.3] will be listed for the safety analysis population.

6.3.12 Study Drug Exposure and Pharmacokinetic Concentration Data

6.3.12.1 Exposure to the Investigational Medicinal Product

The record of dose compliance through the placebo lead-in period and treatment period will be listed [Listing 16.2.5.1] and summarized [Table 14.2.1] for the safety analysis population.

6.3.12.2 Pharmacokinetic Concentrations

The individual plasma concentration of [redacted], the date and time of last preceding dose, actual sampling date and elapsed time will be listed for the PK analysis population [Listing 16.2.5.2].

Plasma [redacted] concentrations will be presented in a scatter plot figure representing concentration versus actual elapsed time post dose and different symbols for TS-121 10 mg or TS-121 50 mg treatment groups [Figure 14.2.1] for the PK analysis population.

Handling of Values Below the Lower Limit of Quantification (BLQ)

All concentrations BLQ or missing data will be labeled as such in the concentration data listings. The lower limit of quantification (LLOQ) will be reported in a listing footnote.

- All BLQ values will be substituted by zeros in the figures.

Significant Digits

Individual drug concentrations values are reported with 3 significant digits.

6.3.13 Efficacy Analysis

6.3.13.1 Primary Endpoint

The Montgomery-Asberg Depression Rating Scale (MADRS) total score with change from baseline will be listed by subject number [Listing 16.2.6.1.1]. A descriptive statistics summary table will be presented for MADRS total score and total score change from baseline by treatment and visit [Table 14.2.2.1]. Baseline for calculating total score change will be the Visit 3 assessment. Each MADRS item will be listed by subject number [Listing 16.2.6.1.2] and MADRS item frequency counts will be presented in a summary table by treatment and visit [Table 14.2.2.2] for the Intent to Treat Population.

Treatment group mean (+/- SE) MADRS total score and total score change from baseline will be plotted versus treatment [Figure 14.2.2.1 and Figure 14.2.2.2].

To evaluate the efficacy of TS-121 compared to placebo on MDD symptoms as measured by the MADRS scale, Mixed Model Repeated Measures analysis (MMRM)/non-parametric methods will be used to compare the difference between the three treatment groups on the primary endpoint (change from baseline in MADRS score) depending on the distribution of the data. The baseline MADRS total score will be included as a covariate. Comparisons of TS-121 10 mg versus placebo and TS-121 50 mg versus placebo will be evaluated at each timepoint.

The following is example SAS programming code that will be used to fit the MMRM model and evaluate treatment mean differences:

```
Proc Mixed Method = REML;
        Class GROUP Subject_ID VISIT;
        Model Change = Baseline GROUP VISIT GROUP*VISIT
                        / solution ddfm=kenwardroger;
        repeated VISIT / sub=Subject_ID type= UN;
        lsmeans GROUP*VISIT / pdiff alpha=0.05 cl;
        ods output lsmeans=lsm covparms=cp tests3=t3 diffs=diffs;
run; quit;
```

In case the repeated measures model with unstructured covariance matrix (type = UN) fails to converge then the compound symmetry model (type = CS) will be used.

Least squares treatment group mean differences with 95% confidence intervals will be reported for TS-121 10 mg versus placebo and for TS-121 50 mg versus placebo evaluated at Visits 4, 5, 6/EOT and 7. The primary endpoints for analysis are treatment mean differences at Week 6/EOT [Table 14.2.2.3]. Statistical significance of Type 3 tests and model variance terms will be reported separately [Table 14.2.2.4].

The primary analysis will assess treatment mean differences at Week 6/EOT. An assessment of time on treatment will be performed as an analysis of covariance on the latest assessment of the primary endpoint, change in MADRS score. The model will include all patients in the specified population and may include patients with the baseline assessment and no post-baseline efficacy data. In this model, where TIME will be included as a continuous predictor in this model, a

statistically significant GROUP*TIME interaction effect, with greater reductions in MADRS scores in the TS-121 groups compared to placebo, will be interpreted as indicating efficacy.

The following is example SAS programming code that will be used to fit the analysis of covariance model:

```
Proc Mixed Method = REML;
        Class GROUP;
        Model Change = Baseline GROUP TIME GROUP*TIME
                        / solution ddfm=kenwardroger;
        lsmeans GROUP / pdiff alpha=0.05 cl;
        ods output lsmeans=lsm2 tests3=t3 solution;
run; quit;
```

Treatment least squares means with 95% confidence intervals will be reported and also slope estimates for the baseline and TIME covariates [Table 14.2.2.5]. Model details will be presented in a statistical appendix.

6.3.13.2 Secondary Endpoints

The MADRS responder status, indicating ≥ 50% reduction in total score from baseline (Visit 3 assessment) will be listed [Listing 16.2.6.1.1] and the number and percentage responders will be summarized by treatment and visit [Table 14.2.2.6]. Percentage responders at each visit will be calculated relative to number of patients evaluated.

The Clinical Global Impression Severity (CGI-S) and CGI Improvement (CGI-I) will be presented in a listing with CGI-S change from baseline and CGI-I-Improver status, indicating ("Very much improved" or "Much improved") [Listing 16.2.6.2]. The CGI-S score and score change from baseline will be presented in a descriptive summary table [Table 14.2.3.1]. The CGI-I summary by treatment group and visit will include score frequencies and the mean and median CGI-I scores [Table 14.2.3.2].

Treatment group mean (+/- SE) CGI-S score and score change from baseline will be plotted versus treatment [Figure 14.2.3.1 and Figure 14.2.3.2].

Hamilton Anxiety Scale (HAM-A) total score, total score change from baseline and item scores will be presented in a listing [Listing 16.2.6.3]. A descriptive summary table will be presented

for HAM-A total score and total score change from baseline (n, mean, SD, SE, CV%, minimum, median and maximum) by treatment and visit [Table 14.2.4.1]. A summary table will be presented for each item response of HAM-A, with score frequencies and the mean and median scores [Table 14.2.4.2]. Treatment group mean (+/- SE) HAM-A observed and change from baseline will be plotted versus treatment [Figure 14.2.4.1 and Figure 14.2.4.2].

Symptoms of Depression Questionnaire (SDQ) total score, total score change from baseline and item scores will be listed [Listing 16.2.6.4]. A descriptive summary table will be presented of total score observed and change from baseline results [Table 14.2.5.1] and a summary table of each item with the mean and median scores [Table 14.2.5.2]. Treatment group mean (+/- SE) SDQ total score and change from baseline will be plotted versus treatment [Figure 14.2.5.1 and Figure 14.2.5.2].

With respect to the secondary qualitative outcome variables, the MMRM/non-parametric methods described above will also be used to compare the difference among the three treatment groups for each secondary endpoint depending on the distribution of the data.

Changes from Baseline (Visit 3) to Week 6/EOT will be assessed for the following:

- CGI-S score [Table 14.2.3.3 LSmeans, Table 14.2.3.4 Type 3 tests]
- HAM-A total score [Table 14.2.4.3 LSmeans, Table 14.2.4.4 Type 3 tests]
- SDQ score [Table 14.2.5.3 LSmeans, Table 14.2.5.4 Type 3 tests]

The Pearson Chi-square test will be used to compare the difference between active treatment groups and the placebo group for the following secondary endpoints evaluated at Week 6/EOT. The 95% confidence interval (asymptotic calculation) for each treatment difference, TS-121 - Placebo, of percentage responders will also be reported.

- Percentage of MADRS responders (≥ 50% reduction in total score) [Table 14.2.2.6]
- Percentage of CGI-I improvers ("Very much improved" or "Much improved") [Table 14.2.3.2]

6.3.13.3 Multiple Imputation for Missing Visits

Multiple imputation method for missing visits will be performed for primary and secondary efficacy endpoints and the summary of treatment differences at Week 6/EOT will be presented as a sensitivity analysis.

The following example SAS programming multiple imputation (MI) code will be used to assess pattern of missing visits:

```
Proc MI Nimpute=0;
        Var GROUP TRT10 TRT50 BASE WEEK2 WEEK4 WEEK6 FOLLOW;
run;
```

The indicator variables TRT10 (1 for TS-121 10 mg and otherwise 0) and TRT50 (1 for TS-121 10 mg and otherwise 0) are equivalent to the GROUP treatment classification variables.

The following MI code will be used to create a set of imputation data sets.

```
Proc MI Seed=xxxxx Nimpute=x out=impute_1 round= . . . .1 .1 .1 .1 ;
        EM INITIAL = AC MAXITER=25 CONVERGE=0.00001;
        Var TRT10 TRT50 BASE WEEK2 WEEK4 WEEK6 FOLLOW;
run;
```

The number of imputed data sets (Nimpute) will be set to have relative efficiency > 95% for each imputed variable. This number is expected to be < 10.

The MMRM of Section 6.3.13.1 is then run by imputation number to generate two data sets: The model parameters (solutionF, solution for model fixed effects) and the treatment by visit mean differences. Only the set of treatment mean differences used to evaluate the Week 6/EOT treatment mean differences (GROUPDIF1-GROUPDIF2: D50-placebo, D10-placebo, D50-D10) will be sorted by GROUPDIF1, GROUPDIF2 and imputation number.

The following example SAS programming code will be used to evaluate the Week 6/EOT treatment mean differences (D50-placebo, D10-placebo, D50-10):

```
Proc Mianalyze Data=Diff(keep=GROUPDIF1 GROUPDIF2 Est_diff SE_diff) Edf=xx;
        By GROUPDIF1 GROUPDIF1;
        Modeleffects Est_diff;
        Stderr SE_diff;
        Ods Output Parameterestimates=Est_diff1 VarianceInfo = SE_diff1;
run;
```

The following example SAS programming code will be used to evaluate the solution for model fixed effects.

```
Proc Mianalyze  Parms(classvar=full)= solutionF;
        Class GROUPDIF TIME;
        Modeleffects BASE GROUP TIME GROUP*TIME;
run;
```

The sensitivity analysis for MADRS treatment mean differences at Week 6/EOT will be presented [Table 14.2.2.3.1] and the MI analysis for MMRM fixed effects will be reported separately [Table 14.2.2.4.1]. Details of the MI analysis will be presented in a statistical appendix.

Supplementary summary tables and statistical appendices will also be presented for the MI analysis for MMRM fit to secondary endpoints.

6.3.14 Pharmacodynamic (PD) Analysis

All results, including baseline PD biomarkers (salivary cortisol, hair cortisol, plasma copeptin, 8h urinary cortisol) and inflammatory markers (IL-1, IL-6, HS-CRP, and TNF-α), will be listed with relevant metadata, as, for example, urine volume and concentration used to calculate 8h urinary cortisol.

6.3.14.1 Plasma Copeptin

Plasma copeptin and change from baseline will be presented in a listing [Listing 16.2.6.5] and summarized [Table 14.2.6.1]. Change from baseline to Visit 6 will be compared between TS-121 treatments and placebo using an analysis of covariance. The following is example SAS programming code that will be used to fit the model and evaluate treatment mean differences:

```
Proc Mixed Method = REML;
        Class GROUP Subject_ID;
        Model Change = Baseline GROUP baseline*GROUP / ddfm=kenwardroger;
        Random subject_ID;
        lsmeans GROUP / pdiff alpha=0.05 cl;
        ods output lsmeans=lsm diffs=diffs;
run; quit;
```

Least squares mean differences with 95% confidence intervals for TS-121 10 mg versus placebo and for TS-121 50 mg versus placebo evaluated at Visit 6 will be presented [Table 14.2.6.2].

Statistical significance of the baseline*Group interaction will be noted. In case this has $p<0.1$ then the distribution of baseline plasma copeptin across treatment groups will be examined and a different model for comparing plasma copeptin change from baseline across treatment groups may be assessed.

The association between baseline plasma copeptin and baseline MADRS total score will be presented in a scatter plot of MADRS baseline total score versus baseline plasma copeptin [Figure 14.2.6.1.1] for the intent to treat population. The association between biomarker response to treatment and treatment efficacy will be presented in a scatter plot of MADRS total score change from baseline versus plasma copeptin change from baseline [Figure 14.2.6.1.2] for the intent to treat population.

6.3.14.2 Inflammatory Markers

Inflammatory markers (IL-1, IL-6, HS-CRP, and TNF-α) and change from baseline will be presented in a listing [Listing 16.2.6.6] and summarized [Table 14.2.7.1]. Change from baseline at Visit 6 will be compared between TS-121 treatments and placebo using an analysis of covariance for change dependent on baseline, treatment and baseline by treatment interaction. Least squares mean differences with 95% confidence intervals for TS-121 10 mg versus placebo and for TS-121 50 mg versus placebo evaluated at Visit 6 [Table 14.2.7.2].

The association between inflammatory markers at baseline and baseline MADRS total score will be presented in separate scatter plots of MADRS baseline total score versus each inflammatory marker for the intent to treat population [Figure 14.2.6.2.1]. The association between biomarker response to treatment and treatment efficacy will be presented in separate scatter plots of MADRS total score change from baseline versus each inflammatory marker change from baseline [Figure 14.2.6.2.2] for the intent to treat population.

6.3.14.3 Cortisol

Pharmacodynamic biomarkers (salivary cortisol, hair cortisol and 8h urinary cortisol) will be presented in listings [Listing16.2.6.7.1, Listing16.2.6.7.2 and Listing 16.2.6.7.3] and summarized [Table 14.2.8.1].

The association between cortisol assessments at baseline and baseline MADRS total score will be presented in a separate scatter plots for the intent to treat population [Figure 14.2.6.3].

6.3.15 Subgroup Analysis for Primary Endpoint

The association between MADRS total score change from baseline to Visit 6 and the following demographic or anthropometric information, baseline severity, baseline PD biomarkers or other background information will be explored using descriptive subgroup analysis of MADRS total score from baseline and scatter plots. Factors and covariates in the analysis, as well as the MADRS total score change will be presented in a listing for the intent to treat population [Listing 16.2.6.8] including any pooled groups and subgroup categories used in the analysis:

- Gender (Male, Female)
- Age
- BMI
- Race (White, Other)
- Baseline MADRS score
- HAM-D score at Visit 1.1
- Concomitant ADT (SSRI, Other)
- Age at MDD diagnosis
- MDD duration
- Baseline PD biomarkers
    - Cortisol (salivary cortisol morning and evening and the ratio of morning to evening, hair cortisol, 8h urinary cortisol)
    - Plasma copeptin
    - Inflammatory markers (IL-1, IL-6, HS-CRP, TNF- α)

Subgroup categories will be determined based on clinical preference or derived from the data based on the sample median or quartiles.

Descriptive summaries by treatment will be presented of MADRS total score change from Visit 3 to Visit 6 in selected subgroup categories for the intent to treat population [Table 14.2.9.1]. The MADRS total score change will be presented in scatter plots versus subgroup categories or the underlying ungrouped covariate corresponding to tabulated subgroup categories [Figure 14.2.7.1].

6.3.16 Safety Analysis

The analysis of the safety variables will be based on the safety analysis population. Safety will be assessed based on reported adverse events, abnormal findings from physical and

ophthalmological examinations, body weight, vital sign measurements, 12-lead ECGs, C- SSRS score and clinical laboratory assessment results.

6.3.16.1 Adverse Events

The following AE listings will be produced:

- All AEs [Listing 16.2.7.1]
- SAEs [Listing 16.2.7.2]
- AEs leading to Study Drug Discontinuation [Listing 16.2.7.3]
- EOCSs [Listing 16.2.7.4]
- Overdose [Listing 16.2.7.5]

The following information will be included in AE listings: Subject number, age, sex and race, treatment group, verbatim term, MedDRA preferred term (PT), system organ class (SOC), SAE indicator flag, AE onset date (and time) and study day, AE end date (and time), duration (days), severity, relationship to study drug, action taken for study drug, action taken to treat, AE outcome, and treatment emergent (TEAE) flag. Adverse events with onset date on or later than date of first dose of the randomized treatment, which also meet one of the following two criteria, will be considered TEAE.

- The AE was not present prior to the first dose of study drug for the treatment period
- The AE was present prior to the first dose of study drug for the treatment period and increased in severity

Each summary tabulation of AEs will include the number and percentage of patients who experienced TEAEs grouped to treatment (TS-121 10 mg, TS-121 50 mg and placebo) and the number and percentage of all patients who experienced AEs during placebo lead-in.

The overview of AEs includes the number and percentage of patients with at least one AE, with any SAE or with an AE that led to study discontinuation [Table 14.3.1.1].

Tables of number (and percentage) of patients with AEs summarized by SOC, PT and maximum severity [Table 14.3.1.2] and by SOC, PT and relationship to investigational product [Table 14.3.1.3] will be presented by treatment group for the safety analysis population. Treatment related AEs include those that are marked as “possibly”, “probably” or “definitely related to

study drug". In the case of multiple occurrences of the same AEs (at the same SOC and PT level) in an individual patient, the AE with the greatest severity will be reported in Table 14.3.1.2 and the AE with the greatest relationship to study drug will be reported in Table 14.3.1.3.

Serious AEs and EOCSs, if any occur, will be summarized [Table 14.3.2.1 and Table 14.3.2.2].

6.3.16.2 Clinical Safety Laboratory Tests (Biochemistry, Hematology and Urinalysis)

Clinical laboratory biochemistry and hematology results will be presented, using conventional units (original units, received from the laboratory are consistent with CDISC standards) and standard international (SI) units, will be listed by subject number and study time point [Listing 16.2.8.1.1 through Listing 16.2.8.2.2]. Laboratory analyses will utilize scheduled Visit 3 assessments as baseline values. Only if scheduled Visit 3 result is missing will the latest predose assessment be used as the baseline for that patient and test. Change from baseline will be calculated only for laboratory tests with at least one recorded pre-dose assessment.

Certain hematology test samples were assessed by certain sites' local laboratory and these will be listed together with the corresponding central laboratory results.

Clinical urinalysis and urinalysis microscopic tests will be listed using conventional units [Listing 16.2.8.3].

Laboratory reference ranges, in conventional and SI units will be listed [Listing 16.2.8.4].

All values outside the clinical reference ranges will be flagged in the data listings. The quantitative abnormal values will be flagged with 'L' for values below the lower limit of the clinical reference range and 'H' for values above the upper limit of the clinical reference range and included in the listings. In addition, certain qualitative tests may be flagged abnormal 'A'. The Investigator will assess whether the values outside the clinical reference range are clinically significant and laboratory values outside the reference ranges will be reported as abnormal not clinically significant (NCS) or abnormal clinically significant (CS). Clinically significant laboratory values will be recorded by the Investigator as AEs.

Abnormal laboratory values ordered by test category (biochemistry, hematology and urinalysis), randomized treatment, subject number, laboratory test, date and time will be listed using conventional and SI units [Listing 16.2.8.5.1 and Listing 16.2.8.5.2]. Abnormal CS

laboratory values ordered by test category (biochemistry, hematology and urinalysis), randomized treatment, subject number, laboratory test, date and time will be listed using conventional and SI units [Table 14.3.4.1 and Table 14.3.4.2].

Descriptive statistics (n, mean, SD, SE, CV%, minimum, median and maximum) for non-categorical data including biochemistry, hematology and any quantitative urinalysis will be presented by treatment group and time point for both observed values and changes from baseline in conventional units [Table 14.3.5.1.1] and in SI units [Table 14.3.5.1.2]. Urinalysis categorical results (normal or abnormal) will also be tabulated [Table 14.3.5.2].

Biochemistry, hematology and urinalysis shift tables will be constructed that show the change in quantitative laboratory values (or urinalysis normal/abnormal) from baseline to end of treatment [Table 14.3.5.3.1 through Table 14.3.5.3.3].

Critical ranges for clinically significant laboratory outlier criteria (attached Appendix A) and clinically noteworthy changes (Appendix B) will be used to identify values and trends of potential medical importance. A summary of the number and percentage of patients with clinically significant laboratory outliers, based on Appendix A, for each specified laboratory test will be presented [Table 14.3.5.4.1]. The summary will present across scheduled and unscheduled tests after first dose in Visit 3 and across scheduled and unscheduled test during placebo lead-in period. Clinical laboratory outlier values of individual patients will be presented in a listing [Table 14.3.5.4.2]. Similar analyses will be performed for clinically noteworthy changes from baseline [Table 14.3.5.5.1 and Table 14.3.5.5.2].

Evaluation of Drug-Induced Serious Hepatotoxicity

The evaluation of drug-induced serious hepatoxicity (eDISH) will be based on a tabulation by subject number [Table 14.3.4.3] and scatter plot figure. The table will present all scheduled and unscheduled ALT, AST, ALP and TBL of any patient with AST or ALT > 3x ULN or TBL > 2 x ULN. This table will be assessed to identify any incidence of

1. Three-fold or greater elevations above the ULN of ALT or AST and

2. Among patients showing such AT elevations, any patient with elevation of serum TBL to >2xULN and without initial finding of cholestasis (elevated serum ALP).

Side by side scatter plots will be presented of time-matched ALT versus TBL [Figure 14.3.1.1] and time-matched AST versus TBL [Figure 14.3.1.2] grouped to combined TS-121 treatment

groups and placebo. These plots will be presented using ×ULN units and logarithmic scale. The ALT versus TBL and AST versus TBL plots will also be presented using each patient's post-baseline maximum value [Figure 14.3.2.1 and Figure 14.3.2.2].

6.3.16.3 Vital Signs

Vital signs will be listed [Listing 16.2.9.1.1]. Vital signs change from baseline will also be listed [Listing 16.2.9.1.2]. Baseline is defined as the scheduled predose assessment of Visit 3. Only if scheduled Visit 3 result is missing, will the latest predose assessment be used as the baseline for that subject number and test. Change from baseline will be calculated only for vital signs evaluated post-baseline and with at least one recorded pre-dose assessment.

Descriptive summaries (n, mean, SD, SE, CV%, minimum, median and maximum) of vital signs (by parameter) and oral body temperature measurements observed values and change from baseline will be presented by treatment and time point [Table 14.3.6.1.1 and Table 14.3.6.1.2].

6.3.16.4 12-Lead Electrocardiogram

Electrocardiogram data (ventricular rate, RR, PR, QT, QTcB and QTcF interval data, QRS duration) will be listed [Listing 16.2.9.2.1] with the Investigator interpretation. Any QTcF or QTcB >500 msec will be flagged in the listing. Electrocardiogram data change from baseline will be listed [Listing 16.2.9.2.2] and each flagged to distinguish QTc change from baseline 30 to 60 msec or >60 msec.

Descriptive summaries (n, mean, SD, SE, CV%, minimum, median and maximum) of interval data observed values and change from baseline will be presented by treatment and time point [Table 14.3.6.2.1 and Table 14.3.6.2.2]. Baseline is defined as the scheduled predose assessment of Visit 3. Only if scheduled Visit 3 result is missing, will the latest predose assessment be used as the baseline for that subject number and test.

The categorical analysis of QTcB and QTcF will be presented in a frequency table of patients categorized based on maximum post-baseline QTc (<= 500 msec or >500 msec) and maximum change from baseline (<30 msec, 30 to 60 msec, and >60 msec) [Table 14.3.6.2.3].

The incidence of abnormalities, based on the clinical interpretations from the Investigator, will be enumerated by treatment group and visit [Table 14.3.6.2.4].

6.3.16.5 Body Weight

All body weight (kg) measurements and change from baseline will be listed [Listing 16.2.9.3]. Descriptive summaries (n, mean, SD, SE, CV%, minimum, median and maximum) of body weight and change from baseline will be presented by treatment and time point [Table 14.3.6.3.1 and Table 14.3.6.3.2]. Baseline is defined as the scheduled predose assessment of Visit 3. Only if scheduled Visit 3 result is missing, will the latest predose assessment be used as the baseline for that subject number and test.

6.3.16.6 Ophthalmological Examinations

The observed and change from baseline LOCS III grading scores and best corrected visual acuity scores will be listed [Listing 16.2.9.4]. Any LOCS III grade increase from baseline >=0.5 or >= 0.8 will be flagged. The flagged records will also be interpreted and annotated based on the definition of Class I lens event [increase from baseline in the LOCS III grade of ≥0.5 (Nuclear Opalescence), ≥0.8 (Cortical) or ≥0.5 (Posterior Subcapsular)]. Any record of Class I lens events will be listed [Table 14.3.6.4.1].

Descriptive statistics (n, mean, SD, SE, CV%, minimum, median and maximum) of LOCS III grading items and best corrected visual acuity score and change from baseline will be presented by treatment and time point [Table 14.3.6.4.2 and Table 14.3.6.4.3]. Baseline is defined as the latest predose assessment of Visit 2.

6.3.16.7 Columbia Suicidal Severity Rating Scale (C-SSRS)

Responses to C-SSRS suicidal ideation questions 1 through 5, the suicidal ideation score and score change from baseline will be listed by treatment, subject number and visit [Listing 16.2.9.5.1]. All reported responses to Suicidal Ideation Item 1 (Wish to be dead) and Item 2 (Non-Specific Active Suicidal Thoughts), any positive response to Item 3 Item 4 Item 5 and the suicidal ideation score with change from baseline will be listed.

A shift table of suicidal ideation items 1 and 2 evaluated at baseline (Visit 3) and at Week 2, Week 4, Week 6/EOT, and Follow-up will be presented [Table 14.3.6.5.1]. A descriptive summary (n, mean, SD, SE, min, median, max) by treatment and visit will be presented of observed and change from baseline suicidal ideation score [Table 14.3.6.5.2].

The suicidal ideation intensity rating, including rating 0, any reported intensity of suicidal ideation and any reported suicidal behavior will be listed [Listing 16.2.9.5.2].

7. REFERENCES

1. SAS® Version 9.2 of the SAS System for Unix. Copyright © 2002-2003. SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.
2. U. S. Food and Drug Administration/Center for Drug Evaluation and Research. (2012). Guidance for Industry Suicidal Ideation and Behavior: Prospective Assessment of Occurrence in Clinical Trials.
3. Columbia Suicide Severity Rating Scale Scoring and Data Analysis Guide. http://www.cssrs.columbia.edu.

Table 2 Tables to be Included in Section 14 of the Clinical Study Report

*Note: Top line tables are indicated **

SECTION 14.1:	**Demographic Data**
*Table 14.1.1.1	Patient Disposition by Treatment Group and Overall (All Randomized Patients)
Table 14.1.1.2	Patient Disposition by Treatment Group and Overall (Safety Analysis Population)
Table 14.1.1.3	Patient Disposition Through End of Placebo Lead-in (All Consented Patients)
Table 14.1.2.1	Demographics and Baseline Characteristics including Screening Height, Weight and Body Mass Index (Safety Analysis Population)
*Table 14.1.2.2	Hamilton Depression Scale HAM-D Total Score at Screening and Baseline (Safety Analysis Population)
*Table 14.1.3	Baseline Major Depressive Disorder Characteristics (Safety Analysis Population)
Table 14.1.4	Medical History and Psychiatric History (Safety Analysis Population)
Table 14.1.5	Concomitant Medications (Safety Analysis Population)
SECTION 14.2:	**Treatment Compliance, PK, Efficacy and PD Data**
Table 14.2.1	Treatment Dose Compliance (Safety Analysis Population)
*Table 14.2.2.1	Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Observed and Change from Baseline by Treatment and Visit (Intent to Treat Population)
Table 14.2.2.2	Montgomery-Asberg Depression Rating Scale (MADRS) Item Response Frequency by Treatment and Visit (Intent to Treat Population)
*Table 14.2.2.3	Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline Least Squares Treatment Mean Differences by Visit (Intent to Treat Population)
Table 14.2.2.3.1	Multiple Imputation Method for Missing Visits Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline Least Squares Treatment Mean Differences by Visit (Intent to Treat Population)
*Table 14.2.2.4	Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline Model Summary (Intent to Treat Population)
Table 14.2.2.4.1	Multiple Imputation Method for Missing Visits Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline Model Summary (Intent to Treat Population)
*Table 14.2.2.5	Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline at Week 6/EOT (Intent to Treat Population)
Table 14.2.2.6	Montgomery-Asberg Depression Rating Scale (MADRS) Responders Comparison of TS-121 Treatment versus Placebo at Week 6/EOT (Intent to Treat Population)
Table 14.2.3.1	Clinical Global Impression Severity (CGI-S) Score Observed and Change from Baseline by Treatment and Visit (Intent to Treat Population)

Table 3 Figures to be Included in Section 14 of the Clinical Study Report

*Note: Top line figures are indicated **

Figure 14.1.1	Patient Disposition (All Screened Patients) ***Refer to shells document.***
Figure 14.2.1	Plasma [redacted] Concentrations versus Actual Time Post Dose (Pharmacokinetic Population) ***Refer to shells document.*** *Scatter plot using different symbols for TS-121 10 mg and TS-121 50 mg treatment groups. X-axis is actual time post dose.*
*Figure 14.2.2.1	Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Treatment Mean (+/- SE) versus Visit (Intent to Treat Population) ***Refer to shells document.*** *Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks.*
*Figure 14.2.2.2	Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Treatment Mean (+/- SE) Change from Baseline versus Visit (Intent to Treat Population) ***Refer to shells document.*** *Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks. Reference line at zero change from baseline.*
Figure 14.2.3.1	Clinical Global Impression Severity (CGI-S) Score Treatment Mean (+/- SE) versus Visit (Intent to Treat Population) *Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks.*
Figure 14.2.3.2	Clinical Global Impression Severity (CGI-S) Score Treatment Mean (+/- SE) Change from Baseline versus Visit (Intent to Treat Population) *Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks. Reference line at zero change from baseline.*
Figure 14.2.4.1	Hamilton Anxiety Scale (HAM-A) Treatment Mean (+/- SE) versus Visit (Intent to Treat Population) *Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks.*

Figure 14.2.4.2 Hamilton Anxiety Scale (HAM-A) Treatment Mean (+/- SE) Change from Baseline versus Visit (Intent to Treat Population)

Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks. Reference line at zero change from baseline.

Figure 14.2.5.1 Symptoms of Depression Questionnaire (SDQ) Total Score Treatment Mean (+/- SE) versus Visit (Intent to Treat Population)

Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks.

Figure 14.2.5.2 Symptoms of Depression Questionnaire (SDQ) Total Score Treatment Mean (+/- SE) Change from Baseline versus Visit (Intent to Treat Population)

Profile plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis in weeks from Baseline through follow-up. Plot Week 6 or ET at 6 weeks and Follow-up at 10 weeks. Reference line at zero change from baseline.

Figure 14.2.6.1.1 Baseline Montgomery-Asberg Depression Rating Scale (MADRS) Total Score versus Baseline Plasma Copeptin (Intent to Treat Population)

Refer to shells document.

Scatter plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis is baseline plasma copeptin. Plot baseline MADRS total score.

Figure 14.2.6.1.2 Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline versus Plasma Copeptin Change from Baseline (Intent to Treat Population)

Scatter plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis is plasma copeptin change. Plot MADRS total score change.

Figure 14.2.6.2.1 Baseline Montgomery-Asberg Depression Rating Scale (MADRS) Total Score versus Baseline Inflammatory Markers (IL-1, IL-6, HS-CRP, and TNF-α) (Intent to Treat Population)

Scatter plot, one pager per marker using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis is baseline Marker. Plot baseline MADRS total score.

Figure 14.2.6.2.2 Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline versus Inflammatory Markers (IL-1, IL-6, HS-CRP, and TNF-α) Change from Baseline (Intent to Treat Population)

Scatter plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis is inflammatory marker. Plot MADRS total score change.

Figure 14.2.6.3 Baseline Montgomery-Asberg Depression Rating Scale (MADRS) Total Score versus Baseline Pharmacodynamic Biomarkers (Intent to Treat Population)

One page each for *salivary cortisol, hair cortisol and 8h urinary cortisol. Scatter plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis is baseline Pharmacodynamic Biomarker (Cortisol level). Plot baseline MADRS total score.*

Figure 14.2.7.1 Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline versus Baseline Patient Characteristics (Intent to Treat Population)

Refer to shells document.

One page per 'patient characteristic' in Listing 16.2.6.8. Scatter plot using different symbols for TS-121 10 mg, TS-121 50 mg and Placebo treatment groups. X-axis is baseline Patient Characteristic (may be group score). Plot MADRS total score change from baseline. In case of grouped scores in only 2 or 3 groups then present overlay box plots (median with 25th to 75th percentiles) of the MADRS Total score change per group and treatment.

Figure 14.3.1.1 Time-matched Post-Baseline Alanine Aminotransferase versus Total Bilirubin (Safety Analysis Population)

Refer to shells document.

One page, two panels. Left is Placebo Group and right is Overall TS-121 Groups. Include scheduled and unscheduled post-baseline assessments. Units are ×ULN for horizontal and vertical axes and scale is logarithmic.

Figure 14.3.1.2 Time-matched Post-Baseline Aspartate Aminotransferase versus Total Bilirubin (Safety Analysis Population)

As Figure 14.3.1.1.

Figure 14.3.2.1 Peak Post-Baseline Alanine Aminotransferase versus Peak Total Bilirubin (Safety Analysis Population)

As Figure 14.3.1.1, one point per patient and unmatched unscheduled results are included.

Figure 14.3.2.2 Peak Post-Baseline Aspartate Aminotransferase versus Peak Total Bilirubin (Safety Analysis Population)

As Figure 14.3.1.1, one point per patient and unmatched unscheduled results are included.

Table 4 Listings to be Included in Section 16 of the Clinical Study Report

*Note: Listings that support top line tables or figures are indicated **

Study Information

Listing 16.1.7	Randomization and Interactive Response Technology (IRT) Record

Documentation Of Statistical Methods

Appendix 16.1.9.2.1	Statistical Analysis Details of Table 14.2.2.3 and Table 14.2.2.4 Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline
Appendix 16.1.9.2.2	Statistical Analysis Details of Table 14.2.2.5 Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline at Week 6/EOT
Appendix 16.1.9.2.3	Statistical Analysis Details of Table 14.2.2.6 Montgomery-Asberg Depression Rating Scale (MADRS) Responders Comparison of TS-121 Treatment versus Placebo at Visit 6 (Intent to Treat Population)
Appendix 16.1.9.2.4	Statistical Analysis Details of Table 14.2.3.2 Summary and Analysis of Clinical Global Impression Improvement (CGI-I) Score Frequency, CGI-I Summary Scores and CGI-I Improver Frequency by Treatment and Visit (Intent to Treat Population)
Appendix 16.1.9.2.5	Statistical Analysis Details of Table 14.2.3.3 and Table 14.2.3.4 Analysis of Clinical Global Impression Severity (CGI-S) Score Change from Baseline (Intent to Treat Population)
Appendix 16.1.9.2.6	Statistical Analysis Details of Table 14.2.4.2 and Table 14.2.4.3 Analysis of Hamilton Anxiety Scale (HAM-A) Change from Baseline (Intent to Treat Population)
Appendix 16.1.9.2.7	Statistical Analysis Details of Table 14.2.5.3 and Table 14.2.5.4 Analysis of Symptoms of Depression Questionnaire (SDQ) Total Score Change from Baseline (Intent to Treat Population)
Appendix 16.1.9.2.8	Statistical Analysis Details of Table 14.2.7.2 Analysis of Inflammatory Markers Change from Baseline at Visit 6 (Intent to Treat Population)
Appendix 16.1.9.2.9	Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.2.3.1 and Table 14.2.2.4.1 Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline
Appendix 16.1.9.2.10	Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.2.5.1 Analysis of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Change from Baseline at Week 6/EOT
Appendix 16.1.9.2.11	Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.2.6.1 Montgomery-Asberg Depression Rating Scale (MADRS) Responders Comparison of TS-121 Treatment versus Placebo at Visit 6 (Intent to Treat Population)
Appendix 16.1.9.2.12	Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.3.2.1 Summary and Analysis of Clinical Global Impression Improvement

(CGI-I) Score Frequency, CGI-I Summary Scores and CGI-I Improver Frequency by Treatment and Visit (Intent to Treat Population)

Appendix 16.1.9.2.13 Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.3.3.1 and Table 14.2.3.4.1 Analysis of Clinical Global Impression Severity (CGI-S) Score Change from Baseline (Intent to Treat Population)

Appendix 16.1.9.2.14 Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.4.2.1 and Table 14.2.4.3.1 Analysis of Hamilton Anxiety Scale (HAM-A) Change from Baseline (Intent to Treat Population)

Appendix 16.1.9.2.15 Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.5.3.1 and Table 14.2.5.4.1 Analysis of Symptoms of Depression Questionnaire (SDQ) Total Score Change from Baseline (Intent to Treat Population)

Appendix 16.1.9.2.16 Multiple Imputation Method for Missing Visits Statistical Analysis Details of Table 14.2.7.2.1 Analysis of Inflammatory Markers Change from Baseline at Visit 6 (Intent to Treat Population)

SECTION 16.2.1: Disposition

*Listing 16.2.1.1 Study Disposition (All Screened Patients)

SECTION 16.2.2: Protocol Deviations

Listing 16.2.2.1 Protocol Deviations

SECTION 16.2.3: Analysis Sets

Listing 16.2.3.1 Assignment to Analysis Populations (Safety Analysis Population)

SECTION 16.2.4: Demographics And Baseline Characteristics

*Listing 16.2.4.1 Demographics (Safety Analysis Population)

Listing 16.2.4.2.1 Screening and Baseline 17 Items Hamilton Rating Scale (HAM-D) (All Screened Patients)

Listing 16.2.4.2.2 Baseline Major Depressive Disorder Characteristics (Intent to Treat Population)

Listing 16.2.4.3 Medical and Psychiatric History (Safety Analysis Population)

Listing 16.2.4.4 Concomitant Medications (Safety Analysis Population)

Listing 16.2.4.5 Prior Antidepressant Therapy (Safety Analysis Population)

Listing 16.2.4.6.1 Drug Screening Results (Safety Analysis Population)

Listing 16.2.4.6.2 Serology Screen Results (Safety Analysis Population)

Listing 16.2.4.6.3 Pregnancy Test Results (Safety Analysis Population)

SECTION 16.2.5: Exposure/Compliance/Drug Concentration

Listing 16.2.5.1 Investigational Product Dose Compliance (Safety Analysis Population)

Listing 16.2.5.2 Pharmacokinetic Sample Times and Plasma [redacted] Concentrations (Pharmacokinetic Analysis Population)

SECTION 16.2.6: Efficacy and Pharmacodynamic Data

*Listing 16.2.6.1.1	Montgomery-Asberg Depression Rating Scale (MADRS) Total Score, Total Score Change from Baseline and Responder Status (Intent to Treat Population)
Listing 16.2.6.1.2	Montgomery-Asberg Depression Rating Scale (MADRS) Item Responses (Intent to Treat Population)
Listing 16.2.6.2	Clinical Global Impression Severity (CGI-S) and CGI Improvement (CGI-I) (Intent to Treat Population)
Listing 16.2.6.3	Hamilton Anxiety Scale (HAM-A) (Intent to Treat Population)
Listing 16.2.6.4	Symptoms of Depression Questionnaire (SDQ) (Intent to Treat Population)
Listing 16.2.6.5	Plasma Copeptin Observed and Change from Baseline (Intent to Treat Population)
Listing 16.2.6.6	Inflammatory Markers, IL-1, IL-6, HS-CRP, and TNF-α (Intent to Treat Population)
Listing 16.2.6.7.1	Salivary Cortisol (Intent to Treat Population)
Listing 16.2.6.7.2	Hair Cortisol (Intent to Treat Population)
Listing 16.2.6.7.3	Urinary Cortisol (Intent to Treat Population)
Listing 16.2.6.8	Patient Characteristics, Demographic and Baseline Pharmacodynamic Biomarkers, Listed with Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Baseline and Change from Baseline (Intent to Treat Population)

SECTION 16.2.7: Adverse Events

*Listing 16.2.7.1	Adverse Events (Safety Analysis Population)
*Listing 16.2.7.2	Serious Adverse Events (SAEs) (Safety Analysis Population)
Listing 16.2.7.3	Adverse Events Associated with Study Drug Discontinued (Safety Analysis Population)
*Listing 16.2.7.4	Events of Clinical Significance (Safety Analysis Population)
Listing 16.2.7.5	Overdose Events (Safety Analysis Population)

SECTION 16.2.8: Listing of Individual Laboratory Measurements by Subject Number

Listing 16.2.8.1.1	Clinical Laboratory Results (Conventional Units): Biochemistry (Safety Analysis Population)
Listing 16.2.8.1.2	Clinical Laboratory Results (Standard International Units): Biochemistry (Safety Analysis Population)
Listing 16.2.8.2.1	Clinical Laboratory Results (Conventional Units): Hematology (Safety Analysis Population)
Listing 16.2.8.2.2	Clinical Laboratory Results (Standard International Units): Hematology (Safety Analysis Population)
Listing 16.2.8.2.3	Local Clinical Laboratory Results and Corresponding Central Laboratory Results (Conventional Units): Hematology (Safety Analysis Population)
Listing 16.2.8.3	Clinical Laboratory Results: Urinalysis (Safety Analysis Population)

Listing 16.2.8.4 Clinical Laboratory Reference Ranges

Listing 16.2.8.5.1 Abnormal Clinical Laboratory Values (in Conventional Units) (Safety Analysis Population)

Listing 16.2.8.5.2 Abnormal Clinical Laboratory Values (Standard International Units) (Safety Analysis Population)

SECTION 16.2.9: Vital Signs, Physical Findings And Other Observations Related To Safety

Listing 16.2.9.1.1 Vital Signs - Observed Values (Safety Analysis Population)

Listing 16.2.9.1.2 Vital Signs – Change from Baseline (Safety Analysis Population)

Listing 16.2.9.2.1 12-Lead Electrocardiogram – Observed Values and Investigator Interpretation (Safety Analysis Population)

Listing 16.2.9.2.2 12-Lead Electrocardiogram – Change from Baseline Values (Safety Analysis Population)

Listing 16.2.9.3 Height, Observed and Change from Baseline Weight and BMI (Safety Analysis Population)

Listing 16.2.9.4 Ophthalmological LOCS III Grading Scores and Best Corrected Visual Acuity Scores Observed and Change from Baseline (Safety Analysis Population)

Listing 16.2.9.5.1 Columbia-Suicide Severity Rating Scale, Suicidal Ideation (Safety Analysis Population)

Listing 16.2.9.5.2 Columbia-Suicide Severity Rating Scale, Ideation Intensity and Behavior Detail (Safety Analysis Population)

APPENDIX A: CLINICALLY SIGNIFICANT LABORATORY OUTLIER CRITERIA

	Outlier Criteria [a]	
Analyte	**SI Units**	**Conventional Units**
HEMATOLOGY		
Hemoglobin (increase)	Value >2 mmol/L above the ULN	Value >2 g/dL above the ULN
Hemoglobin (decrease)	Value <6.2 mmol/L	Value <10 g/dL
Leukocytes (decrease)	Value <3.0 $x10^9$/L	Value <3.0 $x10^3$/µL
Lymphocytes (increase)	Value >4.0 $x10^9$/L	Value >4.0 $x10^3$/µL
Lymphocytes (decrease)	Value <0.8 $x10^9$/L	Value <0.8 $x10^3$/µL
Neutrophils (decrease)	Value <1.5 $x10^9$/L	Value <1.5 $x10^3$/µL
Platelets (decrease)	Value <75 $x10^9$/L	Value <75 $x10^3$/µL
BIOCHEMISTRY		
Albumin (decrease)	Value <30 g/L	Value <3 g/dL
Alkaline phosphatase (increase)	Value >2.5 xULN	Value >2.5 xULN
ALT (increase)	Value >2.5 xULN	Value >2.5 xULN
AST (increase)	Value >2.5 xULN	Value >2.5 xULN
Bicarbonate (decrease)	Value <16 mmol/L	Value <16 mEq/L
Bilirubin (increase)	Value >1.5 xULN	Value >1.5 xULN
Calcium (decrease)	Value <2.0 mmol/L	Value <8.0 mg/dL
Calcium (increase)	Value >2.9 mmol/L	Value >11.5 mg/dL
Cholesterol (increase)	Value >7.75 mmol/L	Value >300 mg/dL
CPK (increase)	Value >2.5 xULN	Value >2.5 xULN
Creatinine (increase)	Value >1.5 xULN	Value >1.5 xULN
GGT (increase)	Value >2.5 xULN	Value >2.5 xULN
Glucose (decrease)	Value <3.0 mmol/L	Value <55 mg/dL
Glucose (increase)	Value >8.9 mmol/L	Value >160 mg/dL
Magnesium (decrease)	Value <0.5 mmol/L	Value <1.2 mg/dL
Magnesium (increase)	Value >1.23 mmol/L	Value >3.0 mg/dL
Phosphate (decrease)	Value <0.8 mmol/L	Value <2.5 mg/dL
Potassium (decrease)	Value < 3.0 mmol/L	Value < 3.0 mEq/L
Potassium (increase)	Value >5.5 mmol/L	Value >5.5 Meq/L
Sodium (decrease)	Value <130 mmol/L	Value <130 mEq/L
Sodium (increase)	Value >150 mmol/L	Value >150 mEq/L
Triglycerides (increase)	Value >2.5 xULN	Value >2.5 xULN
Uric acid (increase)	Value >0.59 mmol/L	Value >10 mg/dL
URINALYSIS		
Hemoglobin, urine (increase)	Present	Present
Protein, urine (increase)	Result of 2+ or 3+	Result of 2+ or 3+

Notes: LLN = lower limit of normal range; ULN = upper limit of normal range; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CPK = creatine phosphokinase; GGT = gamma-glutamyl transpeptidase.

[a] Outlier values based on Grade 2 (moderate) or higher from the Common Toxicity Criteria for Adverse Events v3.0 (CTCAE) Publish Date: August 9, 2006.

APPENDIX B: CRITERIA FOR CLINICALLY NOTEWORTHY LABORATORY CHANGES

Analyte	Criteria
Hemoglobin	Male: < 0.85 x Baseline value Female: < 0.85 x Baseline value
Hematocrit	Male: < 0.85 x Baseline value Female: < 0.85 x Baseline value
White Blood Cell Count (WBC)	if Baseline value is ≤ LLN, < 0.85 x Baseline value or if Baseline value is ≥ ULN, > 1.15 x Baseline value or if Baseline value is > ULN & a post-treatment value is < LLN or if Baseline value is < LLN & a post-treatment value is > ULN
Neutrophils	< 0.85 x Baseline value
Eosinophils	> 1.5 x Baseline value
Platelet Count	if Baseline value is ≤ LLN, < 0.85 x Baseline value
Alkaline Phosphatase	> 1.5 x Baseline value
ALT	> 1.5 x Baseline value
AST	> 1.5 x Baseline value
CPK	> 1.5 x Baseline value
Total Bilirubin	> 1.25 x Baseline value
Albumin	< 0.9 x Baseline value
Glucose	if Baseline value is < LLN, < 0.8 x Baseline value or if Baseline value is > ULN, > 2 x Baseline value or if Baseline value is > ULN & a post-treatment value is < LLN or if Baseline value is < LLN & a post-treatment value is > ULN
Sodium	if Baseline value is < LLN, < 0.95 x Baseline value or if Baseline value is > ULN, > 1.05 x Baseline value or if Baseline value is > ULN & a post-treatment value is < LLN or if Baseline value is < LLN & a post-treatment value is > ULN
Potassium	if Baseline value is < LLN, < 0.9 x Baseline value or if Baseline value is > ULN, > 1.1 x Baseline value or if Baseline value is > ULN & a post-treatment value is < LLN or if Baseline value is < LLN & a post-treatment value is > ULN
Chloride	if Baseline value is < LLN, < 0.9 x Baseline value or if Baseline value is > ULN, > 1.1 x Baseline value or if Baseline value is > ULN & a post-treatment value is < LLN or if Baseline value is < LLN & a post-treatment value is > ULN
Calcium	if Baseline value is < LLN, < 0.9 x Baseline value or if Baseline value is > ULN, > 1.1 x Baseline value or if Baseline value is > ULN & a post-treatment value is < LLN or if Baseline value is < LLN & a post-treatment value is > ULN
Blood Urea Nitrogen (BUN)	> 1.2 x Baseline value
Creatinine	> 1.33 x Baseline value
LDH	> 1.5 x Baseline value

Notes: LLN = lower limit of normal range; ULN = upper limit of normal range; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CPK = creatine phosphokinase; LDH = lactate dehydrogenase.

APPENDIX C: CLASSIFICATION OF ADVERSE EVENTS

Severity	**Mild**	awareness of sign or symptom, but easily tolerated
	Moderate	discomfort enough to cause interference with usual activity
	Severe	incapacitating with inability to work or do usual activity
Duration	Record the start and stop dates of the adverse experience. If less than 1 day, indicate the appropriate length of time and units	
Action taken	Did the adverse experience cause the test drug to be discontinued? Interrupted? Dosage increased or decreased? Other action?	
Relationship to test drug	Did the test drug cause the adverse experience? The determination of the likelihood that the test drug caused the adverse experience will be provided by an investigator. The investigator's signed/dated initials on the source document supporting the causality noted on the AE form ensures that a medically qualified assessment was done. This initialed document must be retained for the required regulatory time frame. The criteria below are intended as reference guidelines to assist the investigator in assessing the likelihood of a relationship between the test drug and the adverse experience based upon the available information. The greater the correlation with the components and their respective elements (in number and/or intensity), the more likely the test drug caused the adverse experience (AE):	
	Exposure	Is there evidence that the subject/subject was actually exposed to the test drug such as: reliable history, acceptable compliance assessment (pill count, diary, etc.), expected pharmacologic effect, or measurement of drug/metabolite in bodily specimen?
	Time Course	Did the AE follow in a reasonable temporal sequence from administration of the test drug? Is the time of onset of the AE compatible with a drug-induced effect?
	Likely Cause	Is the AE not reasonably explained by another etiology such as underlying disease, other drug(s)/vaccine(s), or other host or environmental factors?
	Dechallenge	Was the dose of test drug discontinued or reduced? If yes, did the AE resolve or improve? If yes, this is a positive dechallenge. If no, this is a negative dechallenge. **(Note: This criterion is not applicable if: (1) the AE resulted in death or permanent disability; (2) the AE resolved/improved despite continuation of the test drug; or (3) the study is a single-dose drug study.)**

	Rechallenge	Was the subject/subject reexposed to the test drug in this study? If yes, did the AE recur or worsen? If yes, this is a positive rechallenge. If no, this is a negative rechallenge. **(Note: This criterion is not applicable if:** **(1) the initial AE resulted in death or permanent disability, or (2) the study is a single-dose drug study.)**
	Consistency with Study Drug Profile	Is the clinical/pathological presentation of the AE consistent with previous knowledge regarding the test drug or drug class pharmacology or toxicology?
	The assessment of relationship will be reported on the case report forms/worksheets by an investigator according to his/her best clinical judgment, including consideration of the above elements. **Use the following scale of criteria as guidance (not all criteria must be present to be indicative of a drug relationship).**	
	Definitely related	A clinical event, including laboratory test abnormality, occurring in a plausible time relationship to drug administration, and which cannot be explained by concurrent disease or other drugs or chemicals. The response to withdrawal of the drug (dechallenge) should be clinically plausible. The event must be definitive pharmacologically or phenomenologically, using a satisfactory rechallenge procedure if necessary.
	Probably related	A clinical event, including laboratory test abnormality, with a reasonable time sequence to administration of the drug, unlikely to be attributed to concurrent disease or other drugs or chemicals, and which follows a clinically reasonable response on withdrawal (dechallenge). Rechallenge information is not required to fulfil this definition.
	Possibly related	A clinical event, including laboratory test abnormality, with a reasonable time sequence to administration of the drug, but which could also be explained by concurrent disease or other drugs or chemicals. Information on drug withdrawal may be lacking or unclear.
	Not related	A clinical event, including laboratory test abnormality, with a temporal relationship to drug administration which makes a causal relationship improbable, and in which other drugs, chemicals or underlying disease provide plausible explanations.