Otsuka Pharmaceutical Development & Commercialization, Inc.

Investigational New Drug

REXULTI® (brexpiprazole, OPC-34712, OPC-331, Lu AF41156)

Protocol No. 331-201-00289

A Phase 4, Multicenter, Open-label, Interventional Trial to Assess the Effects on Engagement of Flexible-dose Brexpiprazole (OPC-34712) as Adjunctive Therapy for the Treatment of Adults With Major Depressive Disorder

A Phase 4, Canadian Interventional Trial to Assess Brexpiprazole (OPC-34712) as Adjunctive Therapy in Adults With Major Depressive Disorder

Statistical Analysis Plan

Phase 4
Version: 3.0
Date: 6/29/2022

Protocol Amendment 1: Mar 18, 2021

Table of Contents

List of In-text Tables

List of In-text Figures

List of Appendices

1 Introduction

This statistical analysis plan (SAP) documents the statistical methodology and data analysis algorithms and conventions to be applied for statistical analysis and reporting of efficacy and safety data of study 331-201-00289. All amendments to the protocol are taken into consideration in developing this SAP.

2 Trial Objectives

The primary objective of the trial is:

- To prospectively characterize the effect of brexpiprazole on the concept of engagement, by evaluating Inventory of Depressive Symptomatology Self-Report (IDS-SR)-10-engagement, and the IDS-SR Total score, in patients with MDD with a current depressive episode.

The secondary objectives of the trial are:

- To evaluate patients' impression of improvement of their depression symptoms over time.
- To explore the relationship between engagement and other clinical effects of brexpiprazole, such as improvement in depressive symptoms and functioning.
- To evaluate the safety and tolerability of brexpiprazole (flexible dose; 0.5 to 2 mg once daily [QD]) as adjunctive therapy to antidepressant therapy (ADT) in the proposed patient population with MDD.

3 Trial Design

3.1 Type/Design of Trial

This is a phase 4, multicenter, open-label, flexible dose trial designed to assess the effects of brexpiprazole (flexible dose; 0.5 to 2 mg QD) as adjunctive therapy to ADT in patients with MDD. This trial is being conducted in line with the Canadian Product Monograph[1].

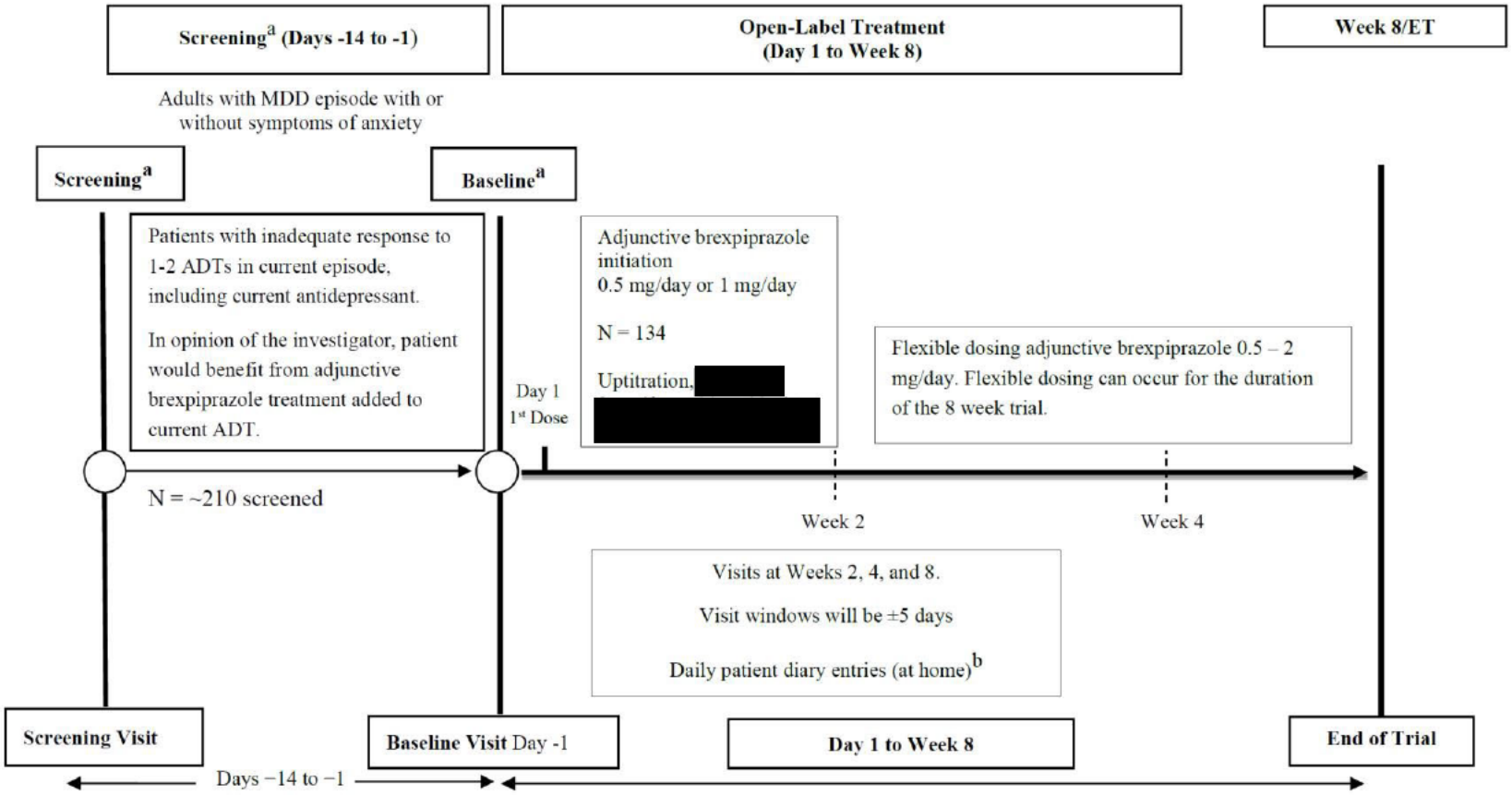
The trial will be organized as follows:

Screening Phase: The screening period of up to 14 days (Days -14 to -1) will begin when the informed consent form (ICF) is signed. The purpose of the screening period is to assess eligibility criteria at 1 or more visits (as necessary to complete screening assessments) and to washout prohibited concomitant pharmacotherapy, if applicable.

Open-Label Treatment Phase: Patients meeting all of the inclusion and none of the exclusion criteria will be enrolled at the baseline visit into an 8-week open-label treatment phase. The

screening and baseline visits can occur on the same day if the patient meets all required enrollment criteria. During the open-label treatment phase, all patients will receive brexpiprazole (that is, adjunctive brexpiprazole as a flexible dose; 0.5 to 2 mg/day) and will continue on the stable dose of ADT that they were taking at screening. Patients will attend visits at Weeks 2, 4, and 8 during the open-label treatment phase. If any patient discontinues the trial early, every effort should be made to complete the Week 8/early termination (ET) evaluations.

A schematic of the trial design is presented in Figure 3.1-1.



N = number of patients.

[a]The screening visit can comprise 1 or more visits as needed. The screening and baseline visits can occur the same day if the patient meets all required enrollment criteria.

[b]Patient diary entries will be reduced in frequency from daily to twice weekly from the day after the Week 4 visit to Week 8 (word of the day and PGI-S).

Figure 3.1-1 Trial Design Schematic

3.2 Trial Treatments

Treatment will consist of open-label brexpiprazole plus an ADT. The investigator must be able and willing to prescribe brexpiprazole to the patient per the approved Canadian Product Monograph[1]. Brexpiprazole should be taken at approximately the same time each day and can be taken at the same time as the ADT. Antidepressant therapy is not considered trial drug in this trial. The dosing schedule for brexpiprazole is presented in Table 3.2-1.

Table 3.2-1 Brexpiprazole Dosing Schedule

Trial Days	Week	Brexpiprazole Adjunctive Therapy[a]
1-7	1	Starting dose: Single oral 0.5 or 1 mg tablet QD. Up-titration at weekly intervals [redacted] QD achieved by/at Week 2.
8-14	2	
15-21	3	Single oral [redacted](s) QD. (Dose increases must occur at weekly intervals. The maximum recommended target dose is 2 mg QD).
22-28	4	
29-35	5	
36-42	6	
43-49	7	
50-56	8	

[a]Brexpiprazole tablet strengths are 0.25, 0.5, 1, and 2 mg per the Canadian Product Monograph[1].

Patients will take the first dose of brexpiprazole with or without food on the day after they are enrolled in the trial (Day 1). In accordance with the Canadian Product Monograph[1], the starting dose will be 0.5 mg or 1 mg QD. Up-titration will occur at weekly intervals to reach a dose of at least 1 mg QD by the Week 2 visit. After the Week 2 visit, the dose may be decreased to a minimum of 0.5 mg QD or increased to a maximum of 2 mg QD based on the investigator's clinical judgment. Dose increases may occur at weekly intervals based on the patient's clinical response and tolerability. Dose decreases can occur at unscheduled trial visits. This dosing regimen is consistent with the Canadian Product Monograph[1] and with clinical practice in Canada.

Allowable brexpiprazole doses that may be given starting the day after the Week 2 visit will be [redacted] QD. The dose of 2 mg QD is the recommended target dose and is also the maximum recommended dose.

3.3 Trial Population

The trial population will include male and female outpatients between the regional age of majority (18 or 19 years of age) and 65 years of age, inclusive, with a Diagnostic and

Statistical Manual of Mental Disorders, 5th Edition (DSM-5) diagnosis of a single or recurrent, non-psychotic episode of MDD with or without symptoms of anxiety. Additionally, patients must have a treatment history for the current major depressive episode (MDE) of an inadequate response (assessment of inadequate response is per investigator judgment) to at least 1 and no more than 2 adequate antidepressant treatments (including the antidepressant that the patient is taking at screening).

Patients with a historical diagnosis of any concurrent condition that is exclusionary may be permitted if the investigator determines and documents this prior diagnosis was not appropriate based on current and historical presentation.

Patients with ADHD, obsessive compulsive disorder (OCD), panic disorders, or generalized anxiety disorder can be included under the following conditions:

- These disorders are not the primary focus of treatment at the screening visit
- Changes in any treatment for these disorders (permitted medication and/or psychotherapy) would not likely be required for the duration of the trial
- All other concurrent psychiatric or neurological diagnoses must be discussed with the medical monitor.
  -

4 Sample Size

This is a single-arm trial to assess the effect of brexpiprazole on patient engagement in MDD patients through the following co-primary endpoints: 1) change from baseline to Week 8 in IDS-SR-10-engagement; 2) change from baseline to Week 8 in IDS-SR Total score.

Based on the results of the appropriate brexpiprazole-plus-ADT arms from four previous Phase 2/3 brexpiprazole MDD trials, it is reasonable to expect a mean change of -2.5 (standard deviation = 6.5) in IDS-SR-10-engagement and a mean change of -5 (standard deviation = 13.5) in IDS-SR Total score from baseline to Week 8. A sample size of 100 completers (ie, patients who have an evaluation on IDS-SR Total score and IDS-SR-10-engagement at Week 8) will yield 95% power to show a difference from baseline in IDS-SR Total score and at least 95% power to show a difference from baseline in IDS-SR-10-engagement at a two-sided significance level of 0.05, respectively. It will ensure the overall power is at least 90%. After considering an ET rate of 25%, a total of 134 patients needs to be enrolled in this trial.

It is anticipated that approximately 210 patients will be screened in order to enroll 134 subjects from approximately 15 sites (to complete 100 patients).

5 Statistical Analysis Datasets

5.1 Datasets for Analysis

The following analysis samples are defined for this trial:

Enrolled Sample: comprises all patients who signed an ICF for the trial and enrolled into the open-label treatment phase.

Safety Sample: comprises all enrolled patients who received at least one dose of open-label brexpiprazole.

Efficacy Sample: the Full Analysis Set comprises all patients in the Safety Sample who have a baseline value and at least one post-baseline efficacy evaluation for IDS-SR Total score and IDS-SR-10-engagement during the open-label treatment phase.

Per-Protocol Sample: Comprises all patients in the Efficacy Sample which excludes major protocol deviation subjects.

5.2 Handling of Missing Data

IDS-SR is utilized as the primary efficacy measurement to assess core diagnostic depressive symptoms as well as the atypical and melancholic symptom features of MDD. The IDS-SR consists of 30 items, all rated on a 0 to 3 scale with 0 being the "best" rating and 3 being the "worse" rating. Besides item 9, two sub-items 9A and 9B exist, with possible scores of 1, 2 or 3 for item 9A, and 0 or 1 for item 9B. The scores for these two sub-items are not included in the calculation of the total score. Item 11 or item 12 should be completed but not both, and similarly, item 13 or item 14 should be completed but not both. Should items 11 and 12 be rated both, then the maximum of the two scores will be used. The same approach will be used for handling items 13 and 14.

The IDS-SR Total Score is the sum of ratings of 28 item scores. The possible IDS-SR Total Score ranges from 0 to 84. The IDS-SR Total Score will be un-evaluable if less than 23 of the 28 items are recorded. If the number of items recorded is at least 23 and at most 27, the IDS-SR Total Score will be the mean of the recorded items multiplied by 28 and then rounded to the first decimal place.

The IDS-SR-10-engagement is the sum of ratings of the selected 10 item scores, including item 8 - Response of your mood to good or desired events, item 15 - Concentration/decision making, item 16 - View of myself, item 17 - View of my future, item 19 - General interest, item 20 - Energy level, item 21 - Capacity for pleasure or enjoyment (excluding sex), item 22 - Interest in sex (please rate interest, not activity),

item 23 - Feeling slowed down, and item 29 - Interpersonal sensitivity. The possible IDS-SR-10-engagement ranges from 0 to 30. The IDS-SR engagement score will be un-evaluable if less than 8 of the 10 items are recorded. If 8 or 9 of the10 items are recorded, the IDS-SR engagement score will be the mean of the recorded items multiplied by 10 and then rounded to the first decimal place.

The observed-cases (OC) data set will consist of actual observations recorded at each scheduled visit during the open-label treatment phase, and no missing data will be imputed. MMRM analysis will be performed on the OC dataset.

The last observation carried forward (LOCF) approach will also be used to impute missing data for continuous and categorical efficacy endpoints. The LOCF data set will include data recorded at a scheduled visit during the open-label treatment phase or, if no observation is recorded at that visit, data carried forward from the previous scheduled visit at the open-label treatment phase. Baseline data will not be carried forward to impute missing values for the LOCF data set.

5.3 Definition of Baseline and Last Visit

Baseline is defined as the last available measurement prior to the first dose of open-label brexpiprazole, while last visit is defined as the last available measurement upon completion or early termination of the open-label treatment phase.

6 Outcome Variables

6.1 Efficacy Variables

The efficacy of open-label brexpiprazole is assessed by IDS-SR, Patient Health Questionnaire 9-item (PHQ-9), Generalized Anxiety Disorder 7-item scale (GAD-7), World Health Organization Disability Assessment Schedule (WHODAS) 2.0 short form questionnaire, Sheehan Disability Scale (SDS), Clinical Global Impression - Severity (CGI-S), Clinical Global Impression - Improvement (CGI-I), Patient Global Impression – Severity (PGI-S), Patient Global Impression – Improvement (PGI-I) scales, and ‘Word of the Day’.

6.2 Safety Variables

The safety of the open-label brexpiprazole is assessed based on reported adverse events (AEs), vital signs, body weight, BMI, and physical examination,

7 Disposition and Demographic Analysis

7.1 Subject Disposition

Subject disposition, subject completion rate and reasons for discontinuation will be summarized for the Enrolled Sample.

7.2 Demographic and Baseline Characteristics

Baseline demographic characteristics, including age, race, ethnicity, gender, weight, height, and BMI for the Enrolled Sample, will be summarized. Summary statistics will consist of mean, median, minimum, maximum, and standard deviation for continuous variables and tabulations of frequency distributions for categorical variables.

Baseline disease characteristics will also be summarized for the Enrolled Sample by summary statistics. They include IDS-SR related assessments, PHQ-9 total score, GAD-7 total score, WHODAS 2.0 short form score, SDS score, CGI-S score, PGI-S score. and 'Word of the Day'.

7.3 Medical and Psychiatric History

A summary of medical and psychiatric history will be presented for the Enrolled Sample.

Listings of medical and psychiatric history will also be provided.

7.4 Treatment Compliance

Based on the Study Medication Log of the CRF, compliance in taking study medication is calculated by the dividing total dose taken by the total dose prescribed for subjects during the open-label treatment period. A listing of treatment compliance by subject will be provided on the Safety Sample.

7.5 Concomitant Medications

Number and proportion of subjects taking concomitant medications prior to, during, and post the open-label treatment period will be tabulated by drug classification using the World Health Organization (WHO) drug dictionary on the Safety Sample. In addition, listings of concomitant mediations will be provided.

7.6 Protocol Deviations

Protocol deviation data will be summarized by type of deviations and center. In addition, listing will be provided describing the deviations by subject.

8 Efficacy Analysis

All efficacy analyses pertaining to open-label treatment period will performed on the Efficacy Sample. Statistical comparisons are based on 2-sided, 0.05 significance levels.

8.1 Primary Efficacy Endpoint

The co-primary efficacy endpoints are as follows:

1) Change from baseline to Week 8 in IDS-SR-10-engagement
2) Change from baseline to Week 8 in IDS-SR Total score

8.1.1 Technical Computation Details for Primary Efficacy Analysis

The primary hypothesis is that the mean changes from baseline to Week 8 in both co-primary endpoints will differ from 0, separately. Each of the endpoints will be tested by fitting a MMRM analysis with an unstructured variance covariance structure at a two-sided significance level of 0.05 on OC dataset on Efficacy Sample. The model will include fixed effect terms for visit, baseline value, an interaction term of baseline value by visit, and trial center. The trial will be claimed positive if both p-values are less than 0.05. Therefore, the overall type I error is controlled at 0.05 level.

In case there is a convergence problem with MMRM model with the unstructured variance covariance matrix, the following structures other than unstructured will be used in order of 1) heterogeneous Toeplitz, 2) heterogeneous autoregressive of order 1, and 3) heterogeneous compound symmetry and the first (co)variance structure converging to the

best fit will be used as the primary analysis. If a structured covariance has to be used, the "sandwich" estimator of the standard error of the fixed effects parameters will be used in order to deal with possible model misspecification of the covariance matrix.

The SAS code for the PROC MIXED procedure to carry out the above MMRM analysis with an unstructured variance covariance structure is illustrated as follows:

where baseline is the last IDS-SR assessment total score at baseline.

8.2 Other Efficacy Endpoint Analysis

The other efficacy endpoints are as follows. All of them will be evaluated at a nominal 0.05 level (2-sided) without adjusting for multiplicity on Efficacy Sample:

1) Change from baseline to Weeks 2 and 4 in IDS-SR-10-engagement
2) Change from baseline to Weeks 2 and 4 in IDS-SR Total score
3) Change from baseline to Weeks 2, 4, and 8 in IDS-SR patient-selected engagement items
4) Change from baseline to Weeks 2, 4, and 8 in IDS-SR engagement items selected by both clinicians and patients.
5) Change from baseline to Weeks 2, 4, and 8 in CGI-S score
6) CGI-I score at Weeks 2, 4, and 8
7) Change from baseline to Weeks 4 and 8 in SDS Mean score and SDS individual item scores
8) Change from baseline to Weeks 4 and 8 in WHODAS 2.0 short form score.
9) Change from baseline to Weeks 2, 4, and 8 in PHQ-9 score
10) Change from baseline to Weeks 2, 4, and 8 in GAD-7 score
11) PGI-I score at each timepoint (ie, being measured weekly at home)

12) PGI-I response rate at each timepoint (ie, being measured weekly at home), where response is defined as a PGI-I score of 1 or 2 (very much improved or much improved)

13) Time to first PGI-S response, where response is defined as a reduction of 2 points in PGI-S score from baseline.

Endpoint (1) - (5), (7) - (10) will be analyzed by fitting the same MMRM model described in the primary analysis on OC dataset. Endpoints (6) and (11) will be analyzed using one-sample t-test on LOCF dataset, with the null hypothesis of no change in each of the endpoints. The point estimate with 95% CI will be provided for endpoint (12) using normal approximation on LOCF dataset. Endpoint (13) will be evaluated by time-to-event analysis with Kaplan-Meier curve on OC dataset. The time origin for measuring this event time is the date of the first dose of open-label brexpiprazole. The event date will be the date when a reduction of 2 points in PGI-S score from baseline is achieved.

[redacted]

8.4 Sensitivity analysis

The primary endpoint will be analyzed by Per-Protocol sample. This analyses will be conducted using the same MMRM analysis for the primary efficacy analysis.

9 Safety Analyses

All safety analyses will be conducted based upon the Safety Sample. Safety variables to be analyzed include Adverse Events (AEs), body weights, BMI, vital signs, and physical examinations. In general, baseline of a safety variable is defined as the last observation of the variable before taking the first open-label brexpiprazole unless specified otherwise. Prospectively defined criteria will be used to identify potentially clinically relevant abnormal values for vital signs and body weight.

9.1 Extent of Exposure

Duration of exposure is defined as the last day of open-label brexpiprazole - the first day of open-label brexpiprazole +1. The number and percentage of subject who receive open-

label brexpiprazole will be presented biweekly: i.e., Day 1-14 in Week 1-2, Day 15-28 in Week 3-4, etc. The overall exposure will also be provided by summary statistics.

The mean daily dosage will be summarized biweekly using descriptive statistics. The mean daily dosage per patient biweekly will be determined for every two weeks of the study: ie, Day 1-14 in Week 1-2, Day 15-28 in Week 3-4, etc. This will be calculated by dividing the sum of individual total doses by the number of days in the biweekly interval. The summary will contain the number of patients receiving receive open-label brexpiprazole and the mean and range of the mean daily dose for every two weeks.

9.2 Adverse Events

All AEs will be coded by system organ class and Medical Dictionary for Regulatory Activities preferred term. A treatment-emergent AE (TEAE) is defined as an AE which starts after start of the first open-labelbrexpiprazole, or an AE continues from baseline and is worsening, serious, trial drug-related, or results in death, discontinuation, interruption or reduction of trial medication. The incidence of the following events will be summarized by descriptive statistics:

- TEAEs
- TEAEs by severity
- TEAEs potentially causally related to the trial drug
- TEAEs with an outcome of death
- Serious TEAEs
- TEAEs leading to discontinuation of the trial drug

All TEAEs will be presented in listings. The listings of death due to AEs, serious AEs, AEs leading to discontinuation of the trial drug, will also be provided.

9.3 Clinical Laboratory Data

Listings of pregnancy test results by subject will be provided.

9.4 Vital Sign Data

Vital signs are taken at screening, baseline, Week 4, and Week8/End of Treatment of. Assessments will include blood pressure, heart rate, respiratory rate, and temperature. In

addition, body weight and BMI are measured at screening, baseline, and each visit. The mean changes from baseline in vital signs, body weight, and BMI will be summarized by summary statistics. The incidence of potentially clinically relevant vital sign abnormalities will be tabulated in incidence tables. Criteria for the potentially clinically relevant vital sign abnormalities are provided in Appendix 1. If vital sign assessments are repeated for the same visit, the last repeat values are used for production of summary tables. The potential clinically relevant vital sign abnormalities are also listed by subject.

9.5 Physical Examination Data

Physical examination data will be presented in a listing.

10 Changes in the Planned Analyses

None.

11 Conventions

11.1 Trial Visit Windows

Study visit windows will be used to map visits using study day intervals. The derived study window variable will be named as WEEK and will be footnoted. In listings it will be listed along with the eCRF study visit. The visit window convention shown in Table 11.1-1 is applied to tables and listings for efficacy scales, including IDS-SR, PHQ-9, GAD-7, WHODAS2.0, SDS, CGI-S, and CGI-I. The visit window convention shown in Table 11.1-2 is applied to patient diary entries, including PGI-I and 'word of the day'. In each table, the variable "target day" is defined using the number of days since the 1st use of open-label brexpiprazole. The first day of using open-label brexpiprazole is defined as 'Day 1'.

Table 11.1-1 Study Day and Visit Windows for Efficacy Scales (IDS-SR, PHQ-9, GAD-7, WHODAS2.0, SDS, CGI-S, and CGI-I)

Week	Target Day[a]	Study Day Interval [a]
2	14	2~20
4	28	21~41
8	56	42-70[b]

Protocol 331-201-00289

[a] Relative to the date of 1st use of open-label brexpiprazole; If more than one observation falls within a particular study day interval, then the last observation within that interval is used.

[b] Evaluations occurring more than seven days after the last open-label brexpiprazole date will be excluded.

Table 11.1-2 Study Day and Visit Windows for Patient Diary Entries (PGI-I, PGI-S and 'Word of the Day')

Week	Target Day[a]	Study Day Interval [a]
1	2	2-7
2	9	8-14
3	16	15-21
4	23	22-28
5	30	29-35
6	37	36-42
7	44	43-49
8	51	50-70[b]

[a] Relative to the date of 1st use of open-label brexpiprazole.

[b] Evaluations occurring more than seven days after the last open-label brexpiprazole date will be excluded.

11.2 Scales: Rules for Scoring and Handling of Missing Data

11.2.1 Inventory of Depressive Symptomatology Self-Report (IDS-SR)

IDS-SR is utilized to assess core diagnostic depressive symptoms as well as the atypical and melancholic symptom features of MDD. The IDS-SR consists of 30 items, all rated on a 0 to 3 scale with 0 being the "best" rating and 3 being the "worse" rating. Besides item 9, two sub-items 9A and 9B exist, with possible scores of 1, 2 or 3 for item 9A, and 0 or 1 for item 9B. The scores for these two sub-items are not included in the calculation of the total score. Item 11 or item 12 should be completed but not both, and similarly, item 13 or item 14 should be completed but not both. Should items 11 and 12 be rated both, then the maximum of the two scores will be used. The same approach will be used for handling items 13 and 14.

The IDS-SR Total Score is the sum of ratings of 28 item scores. The possible IDS-SR Total Score ranges from 0 to 84. The IDS-SR Total Score will be un-evaluable if less than 23 of the 28 items are recorded. If the number of items recorded is at least 23 and at most 27, the IDS-SR Total Score will be the mean of the recorded items multiplied by 28 and then rounded to the first decimal place.

The IDS-SR symptom severity is rate using the following categories based on IDS-SR total Score: 0-13: None, 14-25: Mild, 26-38: Moderate, 39-48: Severe, and 49-84: Very severe.

The IDS-SR related scales are assessed at each scheduled visit.

11.2.2 IDS-SR-10-Engagement

The IDS-SR-10-engagement is the sum of ratings of the selected 10 item scores by clinician. It includes item 8 - Response of your mood to good or desired events, item 15 - Concentration/decision making, item 16 - View of myself, item 17 - View of my future, item 19 - General interest, item 20 - Energy level, item 21 - Capacity for pleasure or enjoyment (excluding sex), item 22 - Interest in sex (please rate interest, not activity), item 23 - Feeling slowed down, and item 29 - Interpersonal sensitivity. The possible IDS-SR-10-engagement ranges from 0 to 30. The IDS-SR-10 engagement will be un-evaluable if less than 8 of the 10 items are recorded. If 8 or 9 of the10 items are recorded, the IDS-SR-10 engagement will be the mean of the recorded items multiplied by 10 and then rounded to the first decimal place.

The weighted IDS-SR-10 engagement is a weighted average of the above selected 10 items by clinician. According to the mean relevance rating, the weights are shown in Table 11.2.2-1. If 1 or 2 items are missing, the average of the remaining items will be used to replace those missing items, and the weights below will then be applied for calculation.

Table 11.2.2-1 Weights for Calculating the weighted IDS-SR-10 Engagement

IDS-SR Item	Weight
Items 20, 21	0.109 each
Items 8, 15	0.106 each
Item 19	0.103
Item 17	0.1
Item 16, 23	0.097 each
Item 22	0.088
Item 29	0.085

11.2.3 IDS-SR Patient-Selected Engagement Items

The IDS-SR patient-selected engagement items is the sum of rating of the selected 10 items score by patients. It includes item 5 - Feeling sad, item 8 - Response of your mood to good or desired events, item 10- The quality of your mind, item 15 - Concentration/decision making, item 16 - View of myself, item 17 - View of my future, item 19 - General interest, item 20 - Energy level, item 21 - Capacity for pleasure or enjoyment (excluding sex), and item 23 - Feeling slowed down. The possible IDS-SR patient-selected engagement items range from 0 to 30. The IDS-SR patient-selected engagement items will be un-evaluable if less than 8 of the 10 items are recorded. If 8 or

9 of the10 items are recorded, the IDS-SR patient-selected engagement items will be the mean of the recorded items multiplied by 10 and then rounded to the first decimal place.

The sum of the common 8 IDS-SR item scores selected by both clinician and patients will be also calculated. If 1 of 8 items is missing, then the total score will be un-evaluable.

11.2.4 Patient Health Questionnaire-9 (PHQ-9)

The PHQ-9 is a standardized, self-administered rating scale that assesses the severity of depressive symptoms. The scale consists of 9 items, representing the 9 criteria upon which the diagnosis of DSM-5 depressive disorders is based. Each item is rated from 0 (not at all) to 3 (nearly every day). The total score ranges from 0 to 27. A higher score on the PHQ-9 represents a higher severity of depressive symptoms. The PHQ-9 is assessed at screening and each scheduled visit.

11.2.5 Generalized Anxiety Disorder-7 (GAD-7)

The GAD-7 is a self-reported questionnaire designed to assess anxiety in subjects. The scale contains 7 items and each item is rated from 0 (not at all) to 3 (nearly every day). The total score ranges from 0 to 21. A higher score on the GAD-7 represents greater anxiety symptomatology. The GAD-7 is assessed at each scheduled visit.

11.2.6 World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) Short Form Questionnaire

The WHODAS 2.0 short form is a 12-item self-assessment questionnaire to assess a patient's activity limitations and participation restrictions (ie, disability) across 6 domains of life: cognition (understanding and communicating), mobility (moving and getting around), self-care (hygiene, dressing, eating, staying alone), getting along (interacting with others), life activities (domestic responsibilities, leisure, work and school), and participation (community and society). Each item is rated from 1 (none) to 5 (extreme or cannot do). To calculate WHODAS 2.0 short form score, a single sum method of scoring all 12 items identified in the manual of WHODAS 2.0[2] is used. Therefore, the short form score ranges from 5 (no disability) to 60 (complete disability). When only one item is missing, the simple approach mentioned in the manual of WHODAS 2.0[2] is used to handle missing data. This method should not be used if more than one item is missing. In situation where one or two items are missing, it will be handled as described in the manual of WHODAS 2.0. The WHODAS 2.0 short form is assessed at baseline, Weeks 4 and 8.

11.2.7 Sheehan Disability Scale (SDS)

The Sheehan Disability Scale (SDS) is a self-rated instrument used to measure the effect of the patient's symptoms on work/school, social life, and family/home responsibilities. For each of the three items, scores range from 0 through 10. The number most representative of how much each area was disrupted by symptoms is marked along the line from 0 = not at all, to 10 = extremely. For the work/school item, no response was to be entered if the patient did not work or go to school for reasons unrelated to the disorder and a response therefore not being applicable. The Mean SDS Score will be calculated over the three item scores. All three item scores need to be available with the exception of the work/school item score when this item is not applicable. The SDS is assessed at baseline, Weeks 4 and 8.

11.2.8 Clinical Global Impression - Severity of Illness Scale (CGI-S)

The severity of agitation for each subject will be rated using the Clinical Global Impression- Severity of Illness Scale (CGI-S). To perform this assessment, the investigator (or designee) will answer the following question: "Considering your total clinical experience with this particular population, how mentally ill (as related to agitation) is the subject at this time?" Response choices are 0 = not assessed; 1 = normal, not at all ill; 2 = borderline mentally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill; and 7 = among the most extremely ill subjects. The score 0 (not assessed) will be set to missing. The CGI-S is therefore a 7-point scale from 1 through 7. The CGI-S is assessed at each scheduled visit.

11.2.9 Clinical Global Impression - Improvement (CGI-I)

The efficacy of trial treatment will be rated for each patient using the Clinical Global Impression - Improvement (CGI-I) scale. The rater or investigator will rate the patient's total improvement whether or not it is due entirely to drug treatment. Response choices include: 0 = not assessed, 1 = very much improved, 2 = much improved, 3 = minimally improved, 4 = no change, 5 = minimally worse, 6 = much worse, and 7 = very much worse. The score 0 (not assessed) will be set to missing. The CGI-I is therefore a 7-point scale from 1 through 7. The CGI-I is assessed at Weeks 2, 4, and 6.

11.2.10 Word of the Day

The "Word of the Day" is a word contained in a previously-generated list of meaningful words that has been developed through patient interviews. Words in this list have been assigned a valence (ie, positive, negative, or neutral) and will be selected by patients to best describe how the patient has felt that day. The 'Word of the Day' will be reduced in frequency from daily to twice weekly from Week 5 to Week 8.

11.2.11 Patient Global Impression - Severity (PGI-S)

The Patient Global Impression – Severity (PGI-S) scale is a 7-point single-item self-report scale for the patient to rate the severity of symptoms of MDD. Patients answer the following question: "Taking into account all of your symptoms, how severe is your Major Depressive Disorder at this time?" Scores range from 1 "no symptoms" to 7 "very severe." The PGI-S will be reduced in frequency from daily to twice weekly from Week 5 to Week 8.

11.2.12 Patient Global Impression - Improvement (PGI-I)

The Patient Global Impression – Improvement (PGI-I) scale is a 7-point single-item self report scale depicting a patient's rating of overall improvement in their condition since starting trial medication. Patients answer the following question: "Since starting trial medication, how much have your symptoms of Major Depressive Disorder changed?" Scores range from 1 "very much improved" to 7 "very much worse." The PGI-I will be administered weekly throughout the trial.

12 Reference

[1] REXULTI (brexpiprazole tablets) Canadian Product Monograph. Otsuka Canada Pharmaceutical Inc., Saint-Laurent, Quebec, Canada; Otsuka Pharmaceutical Co., Ltd., Tokyo, Japan; Feb 2017.

[2] Üstün TB, Kostanjsek N, Chatterji S, Rehm J, editors. Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0) Malta: World Health Organization; 2010.

13 Appendix

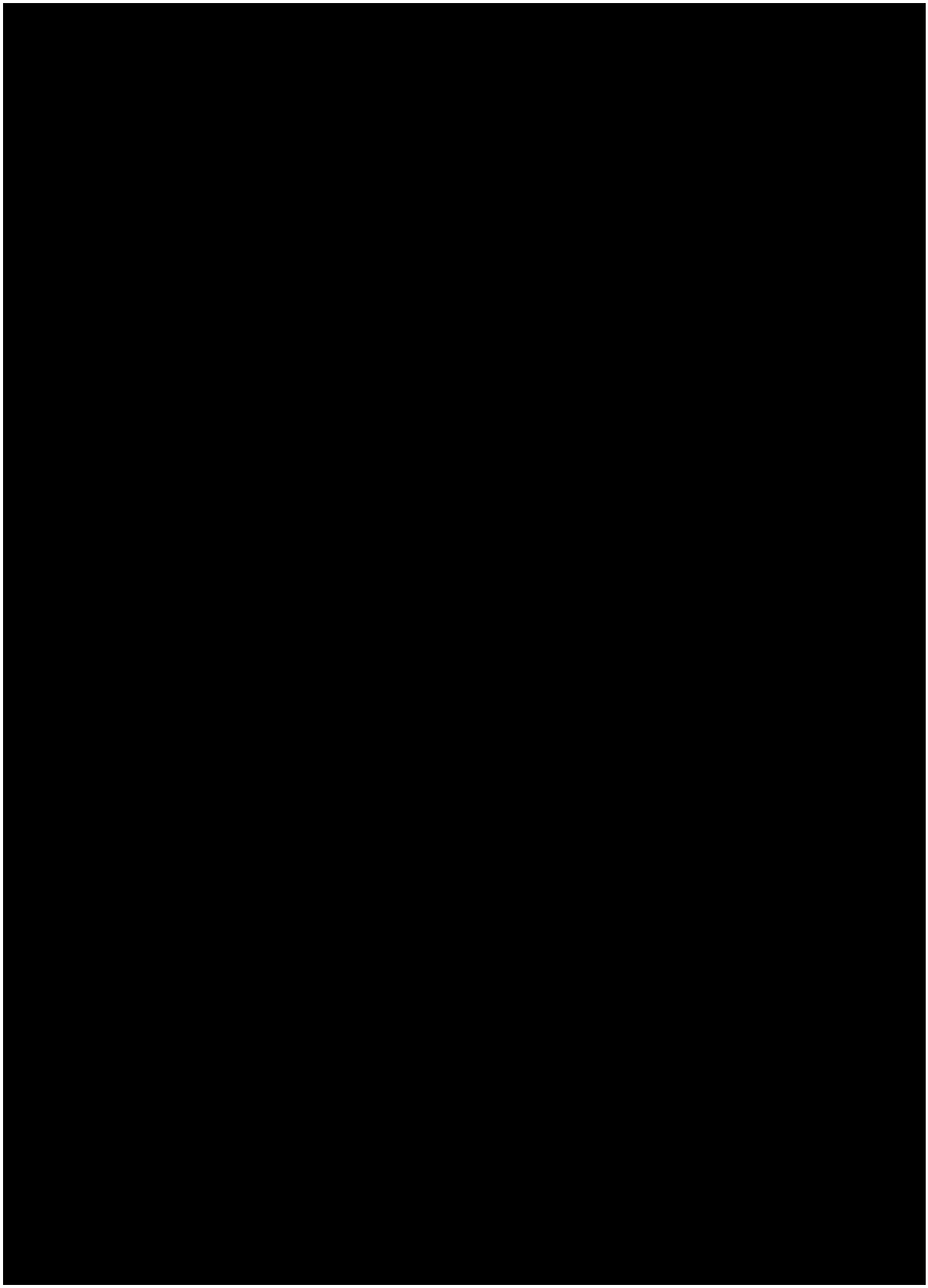
Appendix 1 Criteria for Potentially Clinically Relevant Vital Sign Abnormalities

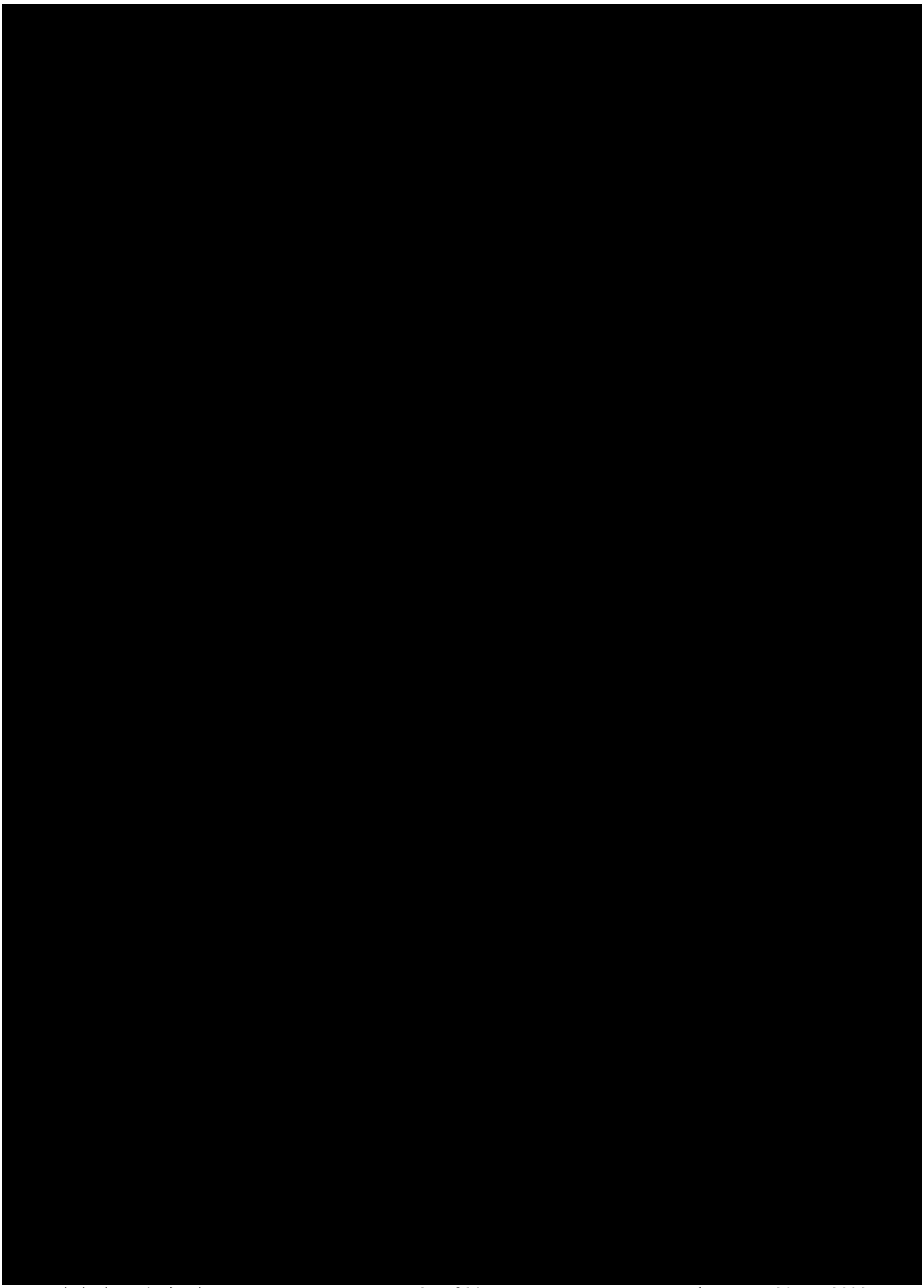
Variable	Criterion Value[a]	Change Relative to Baseline[a]
Heart rate[b]	**> 100 bpm**	**≥ 10 bpm increase**
	< 50 bpm	**≥ 10 bpm decrease**
Systolic blood pressure[b]	**≥ 140 mmHg**	**≥ 20 mmHg increase**
	< 90 mmHg	**≥ 20 mmHg decrease**
Diastolic blood pressure[b]	**≥ 90 mmHg**	**≥ 10 mmHg increase**
	< 60 mmHg	**≥ 10 mmHg decrease**
Weight	–	**≥ 7% increase**
		≥ 7% decrease

bpm = beats per minute; mmHg = millimeter of mercury

[a]In order to be identified as potentially clinically relevant, an on-treatment value must meet the "Criterion Value" and also represent a change from the subject's baseline value of at least the magnitude shown in the "Change Relative to Baseline" column.

[b]As defined in "Supplementary Suggestions for Preparing an Integrated Summary of Safety Information in an Original NDA Submission and for Organizing Information in Periodic Safety Updates," FDA Division of Neuropharmacological Drug Products draft (2/27/87).





This page is a manifestation of an electronically captured signature

SIGNATURE PAGE

Document Name: 331-201-00289_Statistical_Analysis_Plan_Version 3.0 Final

Document Number: [redacted]

Document Version: 2.0

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:min) - UTC timezone
[redacted]	Medical Affairs Approval	11-Jul-2022 12:45:00
[redacted]	Biostatistics Approval	07-Jul-2022 23:26:31
[redacted]	Clinical Approval	11-Jul-2022 14:12:48