STATISTICAL ANALYSIS PLAN

Product Studied: BIIB104 Protocol

Number(s): 263CS201 / NCT03745820

A Phase 2, Randomized, Double-Blind, Multiple-Dose, Placebo-Controlled Study to Evaluate the Safety and Efficacy of BIIB104 in Subjects with Cognitive Impairment Associated with Schizophrenia (CIAS)

Date of Protocol: 30 October 2019 (version 5)

Date of Statistical Analysis Plan: 30 March 2022 (version 2.0)

Approved By:

[redacted], Biostatistics — Date

[redacted], MSc.

[redacted], Biostatistics — Date

[redacted], MSc.

[redacted], Clinical Development — Date

[redacted], Ph.D.

[redacted], Drug Safety — Date

[redacted], M.D.

VERSION HISTORY

SAP Version	Date	Primary Reasons for Amendment
V1.0	11-JAN-2021	
V2.0	30-MAR-2022	Additional analyses and clarifications made as a result of Interim Analysis, Dry Run review and Advisory Board comments

1. TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	adverse event
ALT	alanine transaminase
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
AMPAR	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor
AST	aspartate transaminase
AUC	area under the plasma drug concentration-time curve
AUC_2	area under the plasma drug concentration-time curve from 0 to 2 hours
AUC_4	area under the plasma drug concentration-time curve from 0 to 4 hours
AUC_{24}	area under the plasma drug concentration-time curve from 0 to 24 hours
AUC_{inf}	area under the plasma drug concentration-time curve from 0 to infinity
AUC_{last}	area under the plasma drug concentration-time curve from 0 to the time of last measurement
AUC_{tau}	area under the concentration-time curve within a dosing interval
BID	twice daily
$C_{b,u}$	unbound brain drug concentration
C_{eff}	efficacious plasma drug concentration
BUN	blood urea nitrogen
[redacted]	[redacted]
CGI-I	Global Clinical Impression – Improvement
CGI-S	Global Clinical Impression – Severity
CIAS	cognitive impairment associated with schizophrenia
C_{max}	maximum observed plasma concentration
C_{min}	minimum plasma drug concentration
CNS	central nervous system
CRF	case report form
eCRF	electronic case report form
CSF	cerebrospinal fluid
$C_{ss,av}$	average steady-state drug plasma concentration
C-SSRS	Columbia-Suicide Severity Rating Scale
CV	coefficient of variation
CYP	cytochrome P450
DNA	deoxyribonucleic acid
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th edition
EC	ethics committee
ECG	electrocardiogram
EDCMS	electronic data capture and management system
[redacted]	[redacted]
EQ-VAS	Euroqol-Visual Analog Scale

ERP	Event-related potential
ET	early termination
EU	European Union
FDG	2-deoxy-2-[^{18}F]fluoro-D-glucose
fMRI	functional magnetic resonance imaging
HIV	human immunodeficiency virus
[redacted]	[redacted]
IA	Interim analysis
ICH	International Council for Harmonisation
ID	identification
IDMC	independent data monitoring committee
IQ	intelligence quotient
IRB	institutional review board
IRT	interactive response technology
ITT	intent to treat
LNS	letter number sequencing
MAD	multiple ascending dose
MAR	missing at random
MATRICS	Measurement and Treatment Research to Improve Cognition in Schizophrenia
MCCB	MATRICS Consensus Cognitive Battery
MGH	Massachusetts General Hospital
MHP	mental health professional
MINI	Mini-International Neuropsychiatric Interview for Psychotic Disorders
MMRM	mixed model repeated measures
MTD	maximum tolerated dose
NMDA	*N*-methyl-D-aspartate
NMDAR	*N*-methyl-D-aspartate receptor
NOAEL	no observable adverse effect level
NOEL	no observable effect level
NPI	Neuropsychiatric Inventory
PANSS	Positive and Negative Syndrome Scale
PD	pharmacodynamics
[redacted]	[redacted]
QD	once daily
QTcF	QT interval corrected for heart rate using Fridericia's correction
RNA	ribonucleic acid
SAE	serious adverse event
SAFER (criteria)	State versus trait, Accessibility, Face validity, Ecological validity, Rule of 3 Ps (pervasive, persistent, and pathological)
SAP	statistical analysis plan

SARA	Scale for the Assessment and Rating of Ataxia
SCI-PANSS	Structured Clinical Interview for Positive and Negative Syndrome Scale
SCID	Structured Clinical Interview for DSM disorders
SCoRS	Schizophrenia Cognition Rating Scale
[redacted]	[redacted]
SD	standard deviation
SM	safety margin
TID	3 times daily
[redacted]	[redacted]
UCSD	University of California, San Diego
ULN	upper limit of normal
US	United States
UPSA-Bi	UCSD Performance-Based Skills Assessment – Brief, international version
WRAT4-R	Wide Range Achievement Test 4 – Reading

1. DESCRIPTION OF OBJECTIVES AND ENDPOINTS

1.1 Primary Objective and Endpoints

The primary objective of the study is to evaluate the efficacy of BIIB104 in subjects with cognitive impairment associated with schizophrenia (CIAS), using the Working Memory Domain of the MATRICS Consensus Cognitive Battery (MCCB).

The primary endpoint that relates to this objective is the change from baseline in MCCB Working Memory Domain score to Week 12.

1.2 Secondary Objectives and Endpoints

The secondary objectives are:

- To evaluate the safety and tolerability of BIIB104 in subjects with CIAS.

- To evaluate the efficacy of BIIB104 in subjects with CIAS on measures of cognition, functioning, and psychiatric symptomology, as measured by:
  - University of California, San Diego Performance Based Skills Assessment–Brief, international version (UPSA-Bi)
  - Schizophrenia Cognition Rating Scale (SCoRS)
  - MCCB
  - Positive and Negative Syndrome Scale (PANSS)
  - Clinical Global Impression–Severity (CGI-S) and Clinical Global Impression Improvement (CGI-I).

The secondary endpoints are:

- Incidence of adverse events (AEs) and serious adverse events (SAEs) reported during the study, Scale for the Assessment and Rating of Ataxia (SARA), Columbia–Suicide Severity Rating Scale (C-SSRS)
- Change from baseline in UPSA-Bi assessment to Week 12
- Change from baseline in SCoRS assessment to Week 12
- Change from baseline in MCCB Composite and Individual Domain scores (excluding Working Memory Domain) to Week 12
- Change from baseline in PANSS total score, and subscale scores to Week 12
- Change from baseline in CGI-S to Week 12
- CGI-I at Week 12

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]

[redacted]

- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

2. STUDY DESIGN

2.1 Study Overview

This is a Phase 2, randomized, double-blind, placebo-controlled, parallel-group study to evaluate the safety, tolerability, and efficacy of BIIB104 in subjects with CIAS. Approximately 219 male and female subjects with stable schizophrenia from age 18 to 55 years, inclusive, will be recruited. Dosing with BIIB104 0.15 mg BID and 0.5 mg BID versus placebo BID will be evaluated over a treatment phase of 12 weeks. The study will be conducted at approximately 80 sites globally including non-English speaking countries.

The study includes a screening phase, a placebo lead-in phase, a randomized treatment phase, and a safety follow-up (SFU) phase to begin after the last dose of study treatment.

Screening evaluation will occur up to 35 days prior to randomization. Once eligibility has been established, a 7-day Placebo Lead-In evaluation will be conducted to assess the subject's ability to comply with dosing requirements. After Sponsor confirmation of eligibility and completion of baseline measurements at the Baseline/Day 1 Visit, subjects will be randomized in a 1:1:1 ratio to receive BIIB104 0.5 mg, BIIB104 0.15 mg, or placebo. Subjects will be dosed at approximately 12-hour intervals for 12 weeks. Total duration of subject participation will be approximately 19 weeks from screening to final follow-up visit.

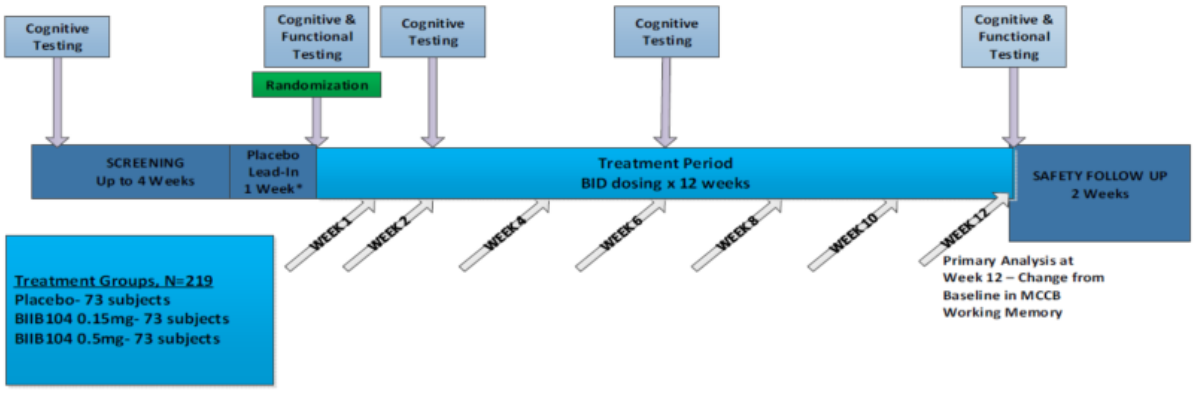
There are 3 treatment groups and subjects are randomized in a 1:1:1 ratio into 1 of 3 treatment groups. The randomization will be stratified by region.

Table1: Treatment groups

Treatment Groups	**Description**
Placebo group	Up to 73 subjects to receive oral dosing BID with placebo for 12 weeks.
BIIB104 0.15 mg group	Up to 73 subjects to receive oral dosing BID with BIIB104 0.15 mg for 12 weeks
BIIB104 0.5 mg group	Up to 73 subjects to receive oral dosing BID with BIIB104 0.5 mg for 12 weeks

2.2 Study Schematic

Figure 1: study design



N = number of subjects; MCCB = MATRICS Consensus Cognitive Battery; BID = twice daily.

*Note: The Placebo Lead-In evaluation may begin any time after eligibility has been established based on screening assessments, provided that the Placebo Lead-In is completed no earlier than 3 days prior to the Baseline/Day 1 Visit.

2.3 Schedule of Events

See protocol section 5.

3. SAMPLE SIZE JUSTIFICATION

The protocol states that:
Fifty-one subjects per treatment group will have approximately 80% power to detect a true mean difference of 3.5 in change from baseline MCCB Working Memory Domain score to Week 12 between the treatment and placebo groups. With the assumed drop-out rate of 30%, 73 subjects are required from each group to maintain the 80% power. This power calculation is based on a 2-

sided *t*-test assuming equal variance with a final significance level of 0.1, and a standard deviation of 7. The statistical software EAST 6.3® is used for sample size calculation. The sample size may be re-estimated based on blinded data review and evolving development of the study.

4. STATISTICAL ANALYSIS METHODS

4.1 General Considerations

This statistical analysis plan (SAP) describes the planned analyses for this phase 2 study evaluating the primary, secondary [redacted] objectives. This SAP also covers the planned interim and final analyses, and therefore there is no separate SAP for the interim analysis.

Data listings will be provided by treatment group, visit, and subject, if applicable. Summary statistics and statistical analyses will only be presented for data where detailed in this SAP. For continuous data, summary statistics will include the arithmetic mean, arithmetic standard deviation (SD), median, minimum (Min), maximum (Max) and n; for log-normal data the geometric mean and geometric coefficient of variation (CV) will also be presented. For categorical data, frequency count and percentages will be presented. Data listings will be provided for all subjects up to the point of withdrawal, with any subjects excluded from the relevant population highlighted. Summary statistics and statistical analyses will generally only be performed for subjects included in the relevant analysis population.

Data will be summarized and presented by treatment group, defined in Table 1. Data will be listed by treatment group. Exposure to study medication will be tabulated by treatment group.

Unless stated otherwise, baseline data are defined as the data collected prior to the time and/or on the day of the first dose, which is usually the same day as the Day1/Baseline visit. If data are not available on Day 1 for safety assessment, the measurements recorded up to 3 days prior to first dose (i.e. Day -3) will be used as baseline. For efficacy assessments, the Day1/Baseline visit may be performed over 2 days, i.e., on Day 1 and Day 2. Therefore, an assessment performed on the second day of visit Day1/Baseline (i.e., Day 2), will be used as baseline if not already performed on the first day of visit Day1/Baseline (i.e., Day 1). If there is more than one value on/before the day of the first dose, the non-missing value closest to and prior to the first dose will be used as the baseline value

Unless stated otherwise, all statistical tests will be 2-sided with statistical significance level of 0.10.

The statistical software, SAS 9.4®, will be used for all summaries and statistical analyses.

4.1.1 Analysis Population

- Intent-to-treat (ITT) population:

The ITT population is defined as all randomized subjects who received at least one dose of study treatment (BIIB104 or placebo). Subjects will be in the treatment group to which they were randomized in statistical analyses based on the ITT population.

- Per-protocol (PP) population:

The PP population is defined as all subjects in the ITT population who do not have major protocol deviation(s) that impact efficacy. Exclusion of subjects from the PP population will be determined by the SMT members including but not limited to medical director, clinical pharmacologist and biostatistician. The main reasons for exclusion will be documented.

- Safety population:

The safety population is defined as all randomized subjects who received at least 1 dose of study treatment (BIIB104 or placebo). Subjects will be in the treatment group they received in statistical analyses based on the safety population.

- [redacted]

4.2 Imputation of Partial Dates for Adverse Events and Concomitant Medication

Partial start dates (of intervention or event):

- Case 1, day is missing (only month and year are present):
    - If year and month are same as treatment period start date, then assign the day of treatment period start date to the partial date. **[worst case scenario (WCS)]**
        - However, if end date of event or intervention is clearly before treatment period start date, assign day '01' to partial start date. **[Max Duration]**
    - Otherwise, assign day of '01' to partial start date. **[Max Duration]**

- Case 2, only year is present:
    - If year is same as treatment period start date, then assign the month and day of the treatment period start date to the partial date. **[WCS]**
        - However, if end date of event or intervention is clearly before treatment period start date, assign 'Jan. 01' to partial start date. **[Max Duration]**
    - Otherwise, assign the month and day of 'Jan. 01' to partial start date. **[Max Duration]**
- Case 3, completely missing date, no imputation is performed.

Partial end dates (of intervention or event):

- Case 1, day is missing (only month and year are present):
    - If year and month are same as study end date, then assign the day of the study end date to the partial date. **[Do not exceed study end]**
    - Otherwise, assign day of last day of the month (28, 29, 30 or 31) to the partial end date **[Max Duration]**

- Case 2, only year is present:
    - If year is same as study end date, then assign the month and day of the study end date to the partial date. **[Do not exceed study end]**
    - Otherwise, assign 'Dec. 31' to the partial end date. **[Max Duration]**
- Case 3, completely missing date, no imputation is performed.

If any partial dates have missing month, with day present, then day is ignored and considered missing. The year-only scenario is considered for imputation purposes.

Also, if the study is ongoing (e.g., interim analysis) and study end date is not available then the cut-off date will be used in the place of study end date. If both a cut-off date and study end date are present for a patient, then the minimum of the two dates will be used as the study (or reference) end date.

4.3 Background Characteristics

The summaries in this section will be based on the ITT population. Unless otherwise specified, summary tables will be presented by treatment group (BIIB104 0.15mg, BIIB104 0.5mg, and placebo). All the listing will be presented by treatment group.

4.3.1 Accounting of Subjects

Accounting of subjects will include number (%) of subjects randomized and dosed, number (%) of subjects who completed the treatment/study, number (%) of subjects who discontinued treatment, and number (%) of subjects who withdrew from study. For subjects who discontinued treatment, the reasons for treatment discontinuation and the number of days on treatment will be summarized. For subjects who withdrew from the study, the reasons for withdrawal and the number of days on the study will be summarized. Time to treatment discontinuation and time to study withdrawal will be displayed by Kaplan-Meier plots.

Accounting of subjects will be also summarized in each analysis population.

4.3.2 Demographics and Baseline Characteristics

The demographic information of age, gender, ethnicity, race, height, weight, smoking status, and body mass index (BMI) will be summarized by treatment group. By-subject listings of such information will be provided.

The categorization of regions is determined from geography, language and culture from each country. The region category will be used as a covariate in statistical models and in subgroup analyses as appropriate.

Baseline clinical assessment variables include baseline MCCB domains, MCCB overall composite score, MCCB neurocognitive composite score, UPSA-Bi, SCoRS, PANSS total score, PANSS Positive Scale, PANSS Negative Scale, PANSS General Psychopathology Scale, SARA, C-SSRS, and CGI-S. Number of years since first episode of schizophrenia, age of first treatment of schizophrenia, number of previous episodes or exacerbations of schizophrenia in the last 24 months, and number of past psychiatric hospitalizations will also be summarized.

Medical history will be coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA). The number (%) of subjects with history (including both ongoing and not ongoing medical conditions) will be summarized by system organ class and preferred term. A listing of medical history will be generated.

4.3.3 Concomitant Medications and Antipsychotic Treatment

Concomitant medications will be coded using the latest version of the World Health Organization (WHO) medication dictionary. Concomitant non-drug treatments will be coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA). A concomitant medication could be taken on or after the day of the first dose of study drug. This includes concomitant medications that start prior to the initiation of the first dose if their use continues or after the date of first dose. An antipsychotic treatment is any treatment for CIAS that is ongoing at the time of trial entry.

To define concomitant use for therapies or antipsychotic treatments with missing start or stop dates, the following additional criteria will be used:

- If both the start and stop dates of a therapy are missing, then the therapy will be considered concomitant;
- If the start date of a therapy is missing and the stop date of that therapy falls on/after the first dose date, the therapy will be considered concomitant;

- If the start date of a therapy is prior to the date of the first dose and the stop date of the therapy is missing and the therapy is listed as ongoing, that therapy will be considered concomitant;

- If the start date of a therapy is prior to the date of first dose and the stop date of that therapy is missing and the therapy is not listed as ongoing, the therapy will be considered non-concomitant.

- If start date of a therapy is after the last dose date but before the end of study date, the therapy will be considered concomitant.

The number and percent of subjects taking concomitant medication, non-drug treatments and antipsychotic treatment will be summarized by treatment group.

4.3.4 Protocol Deviations

Protocol deviations identified during site monitoring will be captured in a protocol deviation log and categorized as major or minor deviations. The major protocol deviations will be summarized by treatment groups. Listings will be generated for the major and minor protocol deviations.

4.3.5 Study Drug Exposure and Study Drug Compliance

A summary table of study drug exposure and compliance will be provided. Number of weeks on study treatment (BIIB104 or placebo), will be calculated as (date of last dose – date of first dose +1) /7. The cumulative duration of weeks on study treatment will be displayed (≥ 1, ≥ 2, ≥ 4, ≥ 6, ≥ 8, ≥ 12 weeks). Descriptive statistics will be presented for days on study treatment.

Compliance, which is the percentage of study treatment taken up to the last dose, calculated as (the actual number of dose taken / by the number of doses a subject is expected to take until the date of last dose) *100, will be summarized both as a continuous variable and a categorical variable (i.e., < 80, 80 to < 90, 90 to <=100, and >100%).

Time on study (in weeks) will be summarized both as a continuous variable and a categorical variable. The categorical variable will be grouped as ≤ 2, 2 - ≤ 4, 4 - ≤ 6, 6 - ≤ 8, 8 - ≤ 10, 10 - ≤ 12, and >12 weeks.
A listing of study drug administration records for placebo subjects who received any dose of active treatment will be provided.

4.4 Efficacy Analysis

4.4.1 General Considerations

For efficacy endpoints, the following treatment groups of BIIB104 (per randomization) will be evaluated and compared with placebo:

- High dose (0.5 mg BID)
- Low dose (0.15 mg BID)

All efficacy analyses will be performed on the ITT population. In addition, the per-protocol population will also be used in the analyses of primary and key secondary efficacy endpoints including MCCB Working Memory, MCCB Neurocognitive Composite, UPSA-Bi, SCoRS.

Primary analysis for continuous endpoints

The primary analysis is the mean difference of the change from baseline to Week 12 between treatment and control group in the ITT population. Data will be censored when a subject prematurely discontinues the study treatment. The estimate of this analysis reflects the treatment effect of BIIB104 if the drug were taken as directed. This is considered [de jure] efficacy estimand [ICH E9 (R1) Addendum 2014, 2017].

The change from baseline will be summarized by treatment group at each post-baseline visit. A mixed model repeated measures (MMRM) model will be used as the primary analysis to analyze change from baseline of the endpoint of interest, using fixed effects of treatment group, study visit, study visit-by-treatment interaction, baseline value of the endpoint, and baseline-by-visit interaction. Region (stratification factor for randomization) will be included as a fixed effect but may be omitted in the case that the model fails to converge. An unstructured covariance matrix will be used to model the within-patient variance-covariance errors. If the unstructured covariance matrix results in a lack of convergence, the following structures will be used in order: heterogeneous Toeplitz (TOEPH), heterogeneous compound symmetry (CSH), antedependence (ANTE1), heterogeneous autoregressive (ARH1), autoregressive (AR1) and compound symmetry (CS). The first covariance structure yielding convergence will be used. If a structured covariance is used, the empirical “sandwich” estimator of the standard error (SE) of the fixed effects parameters will be used to deal with possible model misspecification of the covariance matrix. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. The primary treatment comparison is BIIB104 treatment group versus placebo at the end of Week 12. In the primary analysis, missing data are assumed to be missing at random (MAR) [Rubin 1976]. To account for the interim analysis at the time of the final analysis, the median unbiased estimate of the primary treatment effect will be calculated as a supporting analysis.

However, for UPSA-Bi and SCoRS, which only have assessments at baseline and at Week 12, the primary analysis will be based on subjects who have both baseline and Week 12 assessments. An analysis of covariance (ANCOVA) model will be applied adjusting for treatment group and baseline value of the endpoint (UPSA-Bi or SCoRS, respectively).

Sensitivity analysis: Reference-based multiple imputation approach
The following sensitivity analysis will be performed to assess the robustness of the primary analysis with respect to deviation from the MAR assumption. All sensitivity analyses will be conducted for the ITT population.

Intermittent (non-monotonic) missing data will be imputed based on the MAR assumption and with a multiple imputation model using Markov chain Monte Carlo (MCMC) method within each treatment group. The imputation models will include baseline value and the values at each post-baseline week for the endpoint of interest. The MCMC method in the PROC MI procedure in SAS will be used with a single chain, with a burn-in of 1000, and a thinning of 100 and non-informative priors for all parameters.

Reference-based multiple imputation (MI) approach will be used to impute missingness post treatment discontinuation and consider a missing not at random (MNAR) mechanism for monotonic missing data. Imputed values in the reference group (placebo) will assume MAR. Imputed values in the BIIB104 treatment groups will be based on data from the placebo group using sequential regression MI model. Each sequential regression model (i.e., for imputation of values at a given week) will include visits of endpoint. This approach assumes no sustained benefit of BIIB104 after discontinuation of study treatment and limits post-discontinuation effects to that of placebo.

Two hundred imputations will be performed and from each of the imputed datasets, least square (LS) mean differences from placebo and their standard errors will be combined using Rubin's method [Rubin 1987]. PROC MIANALYZE procedure is used to produce a pooled LS mean estimate of treatment difference at Week 12, its 95% CI, and a pooled p-value.

Complementary Analysis: include data after intercurrent events

This is a complementary analysis to the primary and sensitivity analyses (efficacy estimand) for the ITT population. All observed data including those collected after intercurrent events [ICH E9 (R1) Addendum 2017], i.e. premature discontinuation of the study treatment, will be included. The estimate of this analysis reflects the treatment effect of BIIB104 seen as actually taken. This is considered [*de facto*] effectiveness estimand. When all data after intercurrent events are included, the same MMRM model as the primary analysis will be applied. In this complementary analysis, missing data are assumed to be missing at random [Rubin 1976].

The per-protocol analysis will be done using the same model as the primary analysis and applying in the per-protocol population.

Responder Analysis: threshold determined using distribution-based method

This is a supplementary analysis to assess the effect of individual treatment response in the ITT population only. The threshold, defined as the minimal clinically important difference (MCID), will be estimated based on the distribution-based approach and will use 1/2 pooled standard deviation at baseline calculated from subjects in the ITT population. A subject will be classified as a responder if their change from baseline (in terms of improvement) is greater or equal to the threshold, otherwise they will be classified as a non-responder. All subjects with missing data at Week 12 will be classified as non-responders.

The number of responders and non-responders will be modelled using a logistic regression with treatment group and baseline included as covariates. Region will also be included but may be omitted in the case the model fails to converge.

Visit windows for mapping efficacy endpoints

For efficacy data that are summarized by visit, data collected on all scheduled visits and all unscheduled visits will be mapped to an appropriate analysis visit using the windowing scheme shown in the following few tables. If there are 2 or more assessments available in the same analysis window for a subject, the assessment closest to the target visit day will be used for analysis. If there are 2 or more assessments in the same analysis window with the same distance from the target visit day, the later assessment will be used.

Table 2: Visit Windows for Efficacy Endpoints: MCCB, [redacted], and PANSS

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Prior to the time of, or within one day after, the first treatment dose
Week 2	15	[2, 28]
Week 6	43	[29, 70]
Week 12	85	[71, 91]

* If subject terminate the treatment before Day 71, their visits will be mapped to corresponding analysis visit based on the analysis day (ADY).

Table 3: Visit Windows for efficacy endpoints: UPSA-Bi, SCoRS, [redacted]

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Prior to the time of, or within one day after, the first treatment dose
Week 12	85	[71, 91]
Early Termination	N/A	[2, 70]

* If subject terminate the treatment early before Day 71, their visits will be mapped to nominal visit of "early termination".

Table 4: Visit Windows for Efficacy Endpoints: CGI-S/CGI-I

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Prior to the time of, or within one day after, the first treatment dose
Week 6	43	[2, 70]
Week 12	85	[71, 91]

* CGI-I does not have baseline collected and therefore its visits are applied to Week 6 and 12.
* If subject terminate the treatment early before Day 71, their visits will be mapped to corresponding analysis visit based on the analysis day (ADY).

Considerations for multiple comparison adjustments

There will be no multiple comparison adjustments for primary, secondary efficacy [redacted] endpoints, nor the subgroup analyses.

4.4.2 Primary Efficacy Endpoint

4.4.2.1 Primary Analysis

The primary endpoint is the change from baseline in MCCB Working Memory Domain score to Week 12. The MMRM model will be applied to the primary endpoint and it is described in Section 4.4.1. The same model will also be applied to per-protocol population.

4.4.2.2 Sensitivity Analysis

Reference-based multiple imputation approach will be performed on the primary endpoints using the methods as described in Section 4.4.1.

4.4.2.3 Complementary Analysis

All observed data including those collected after intercurrent events will be applied by MMRM model to the primary endpoint.

4.4.2.4 Responder Analysis

The number of responders and response rate in MCCB Working Memory at Week 12 will be summarised and modelled using the methods as described in Section 4.4.1.

4.4.3 Secondary Efficacy Endpoints

Three key secondary endpoints including changes from baseline to Week 12 in MCCB neurocognitive and overall composite scores, UPSA-Bi, and SCoRS will be applied by the primary analysis, sensitivity analysis, and complementary analysis described in Section 4.4.1. A responder analysis for MCCB neurocognitive, UPSA-Bi Total score, and SCoRS Interviewer Rating Subtotal score only will also be applied as described in Section 4.4.1.

Other secondary efficacy endpoints including MCCB individual domain (except working memory domain), PANSS total score, PANSS positive syndrome, PANSS negative syndrome, CGI-S will be applied by MMRM model of the primary analysis only, which is described in Section 4.4.1.

The summaries of CGI-I at Week 12 will be displayed.

[redacted]

The analysis methods used for primary, secondary, [redacted] endpoints are listed in table 5:

Table 5: Analysis for Efficacy Endpoints:

Endpoints	Analysis	Analysis Population	Assumption for Missing Data
Primary endpoint and key secondary endpoints: MCCB Working Memory, MCCB Neurocognitive Composite, UPSA-Bi, SCoRS	Primary (for MCCB Working Memory, and Neurocognitive Composite): analysis of change from baseline to Week 12 (MMRM)	ITT and per-protocol population	Missing at Random (MAR)
	Primary (for UPSA-Bi and SCoRS): analysis of change from baseline to Week 12 (ANCOVA)	ITT and per-protocol population	N/A
	Sensitivity: multiple imputation using MCMC (MAR) for intermittent missing data; reference-based multiple imputation missing data post treatment discontinuation	ITT	Missing at Random (MAR) for intermittent missing data and Missing Not at Random (MNAR) for missing data post treatment discontinuation
	Complementary: include data after intercurrent events	ITT	Missing at Random (MAR)

	Responder (for MCCB Working Memory, Neurocognitive Composite, UPSA-Bi Total score, and SCoRS Interviewer Rating Subtotal score only): threshold determined using distribution-based method	ITT	N/A
Other secondary endpoints: MCCB individual domain (except working memory), MCCB overall composite score, PANSS total score, PANSS positive scale, PANSS negative scale, and CGI-S	Primary: analysis of change from baseline to Week 12 (MMRM)	ITT	Missing at Random (MAR)
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

4.5 Safety Analysis

4.5.1 General considerations

Analysis population

Safety population will be used for safety analyses of adverse events, clinical laboratory data, Columbia Suicide Severity Rating Scale (C-SSRS) data, ECG data and vital signs data.

Safety treatment groups

Different from efficacy analysis, for safety treatment group if a subject who was randomized to a placebo group accidentally received one or more doses of the active treatment during the study, he or she will be classed to low or high dose group based on the dose group of the active treatment received. A listing of these subjects will be provided. Safety treatment groups will be the same as the randomization treatment groups for other subjects (subjects randomized to active treatment groups, and subjects randomized to placebo without any accidental active dose). Safety treatment groups will be used for all the safety analyses unless otherwise specified.

Incidence and incidence proportion

Incidence and incidence proportion will be provided in incidence tables. Incidence is defined as the number of subjects who experienced an event. Incidence proportion is defined as the number of subjects who experienced an event divided by total number of subjects in the analysis population, i.e., percentage. Each subject will be counted only once within each category.

4.5.2 Adverse Events

All AEs will be analysed based on the principle of treatment emergence. A TEAE is defined as an AE that started or worsened between the time of first dose of study treatment and the end of study date (EOS/ET Visit) recorded on the eCRF.

To define treatment emergence for AEs with missing start or stop date or time the following additional criteria will be used:

- if both the start and stop dates for an AE are missing, then this AE is considered treatment emergent;
- if the start date for an AE is missing and the stop date falls after the first dose date, then this AE is considered treatment emergent;
- if the start date for an AE was the same as the first dose date, then that event is considered treatment emergent.

4.5.2.1 Summary of adverse events

Overall TEAE summary table will include the number of subjects with any TEAE, with any TEAE by maximum severity, the number of subjects with any related TEAE, the number of subjects with SAE, the number of subjects with related SAE, the number of subjects with TEAE leading to study drug discontinuation, the number of subjects with TEAE leading to study withdrawal, and the number of deaths.

A subject is counted only once within each category in each table. A subject with more than one TEAE for a preferred term will be counted only once in preferred term. Similarly, a subject who experienced more than one TEAE for a system organ class (SOC) will be counted only once in that SOC category.

The sorting order for AE incidence tables, unless otherwise specified, will be by decreasing frequency order of BIIB104 total column within each category. A subject is counted only once within each category in each table. For the table of TEAEs by system organ class and preferred terms sorted by decreasing frequency presented by treatment group, system organ class will be presented in decreasing frequency order of BIIB104 total column, and within each system organ class, preferred terms will be presented in decreasing frequency order of total column. A subject is counted only once within each system organ class and preferred term.

The following TEAE incidence tables will be presented by treatment group:

- Overall summary of adverse events
- TEAEs by system organ class and preferred term sorted by decreasing frequency
- TEAEs by system organ class and preferred term sorted by alphabetical order
- TEAEs with at least 5% higher in incidence for either low or high dose compared to placebo, by SOC and preferred term
- TEAEs by preferred term with an incidence rate of 5% or more in any treatment group
- SAEs by system organ class and preferred term by decreasing frequency
- SAEs by preferred term

- TEAE by maximum severity by system organ class and preferred term
- TEAEs by maximum severity by preferred term
- Related TEAEs by system organ class and preferred term
- TEAEs that led to discontinuation of study treatment by system organ class and preferred term
- TEAEs that led to withdrawal from study by system organ class and preferred term

The following listings will be provided:

- A listing of subjects experiencing seizure event
- A listing of subjects experiencing TEAEs
- A listing of subjects experiencing SAEs
- A listing of subjects experiencing TEAEs leading to discontinuation of study treatment
- A listing of subjects experiencing TEAEs leading to withdrawal from study
- A listing of deaths

4.5.3 Clinical Laboratory Data

The following laboratory safety assessments will be conducted:

- Hematology: complete blood count (CBC) including red blood cell count (RBCs), white blood cell counts (WBCs) with differentials (neutrophils, lymphocytes, monocytes, eosinophils, and basophils), hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), and mean corpuscular hemoglobin concentration (MCHC).

- Serum chemistry: albumin, alkaline phosphatase, alanine aminotransferase (ALT), aspartate aminotransferase (AST), blood urea nitrogen (BUN), calcium, chloride, carbon dioxide, creatinine, creatine kinase, direct bilirubin, gamma-glutamyl transpeptidase (GGT), glucose, HbA1c, lactate dehydrogenase (LDH), magnesium, phosphorus, potassium, sodium, total cholesterol, total bilirubin, high-density lipoproteins, low-density lipoproteins, triglycerides, total protein, and uric acid

- Urinalysis: including determination of the presence of urine protein, glucose, ketones, occult blood, and WBCs by dipstick, with microscopic examination if indicated

4.5.3.1 Quantitative Analyses

Hematology, serum chemistry, and urinalysis will be summarized by treatment group and visit. Descriptive statistics will be presented at each visit by treatment group, including number of subjects, mean, standard deviation, median, 25%, 75% quartiles, minimum, and maximum values.

Changes from baseline in quantitative laboratory values will be summarized using descriptive statistics by treatment group and visit. Plots of mean values by visits will presented for numeric laboratory parameters, except for mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), and mean corpuscular hemoglobin concentration (MCHC).

Individual subject listings of clinical laboratory parameters at each visit will be provided. Values for hematology, serum chemistry, and urinalysis values outside the central laboratory reference ranges will be flagged on the individual subject data listings.

Visit windows for by visit summaries

For laboratory by visit summaries, the analysis visit is defined by visit windows in the table below. If there is more than 1 value in the same analysis visit window for a certain parameter for a subject, then the closest record to the target visit day will be used for the by visit analysis. If there are 2 values in the same analysis visit window with the same distance from the target visit day for a certain parameter for a subject, then the record with the later date will be used for the by visit analysis. If there are 2 values on the same day for a certain parameter for a subject, then the record with the later time will be used for the by visit analysis.

Table 6: Visit Windows for Laboratory

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Most recent non-missing pre-dose value
Week 1	8	[2, 10]
Week 2	15	[11, 21]
Week 4	29	[22, 42]
Week 8	57	[43, 70]
Week 12	85	[71, one day before start of safety follow-up]
Safety follow-up	Safety follow-up	[start of safety follow up, the end of study]

4.5.3.2 Qualitative analysis

Clinically significant laboratories

Clinically significant laboratory data will be summarized using shift tables where appropriate. For each subject, the lab values in hematology, blood chemistry and urinalysis will be flagged as “low”, “normal”, or “high” relative to the normal ranges of the central laboratory, or as “unknown” if no result is available. The categorical value in urinalysis will be flagged as “positive” (or “high”), “negative” (or “normal”), or “unknown”.

For blood chemistry and serum chemistry, shifts tables from baseline to low/high post-baseline will be presented. Shift to low includes normal to low, high to low, and unknown to low. Shift to high includes normal to high, low to high, and unknown to high. For a subject if a lab is collected multiple times at post baseline, as long as one lab value is low this subject is regarded to have "low" post-baseline for this lab parameter. Similarly, if at least one post-baseline lab value is "high" this subject is regarded to have "high" post-baseline for this lab parameter.

For urinalysis, shifts tables from baseline to high/positive post-baseline will be presented. Shift to high/positive includes low to high/positive, normal/negative to high/positive, or unknown to high/positive.

Markedly abnormal laboratory

Markedly abnormal indicate clinically significant laboratory abnormalities determined by Biogen. For hematology and blood chemistry the number and percentage of subjects with markedly abnormal laboratory post-baseline will be summarized by treatment group from the criteria in Table 7. Shifts tables from baseline to markedly abnormal post-baseline will also be summarized. For a subject if a lab is collected multiple times at post baseline, as long as one lab value is markedly abnormal this subject is counted to be markedly abnormal for this lab parameter.

Subject's listings will also be presented for all subjects with any markedly abnormal laboratory. In the listings, each subject's complete history from screening to last study visit for that specific laboratory test meeting the criteria will be listed.

Table 7: Criteria for markedly abnormal laboratory

Parameter name	Unit	Low	High
Hematology			
White blood cells	$x10^9$ cells/L	<3.0	>16
Lymphocytes	$x10^9$ cells/L	<0.8	>12
Neutrophils	$x10^9$ cells/L	<1.5	>13.5
Monocytes	$x10^9$ cells/L	N/A	>2.5
Eosinophils	$x10^9$ cells/L	N/A	>1.6
Basophils	$x10^9$ cells/L	N/A	>1.6
Red blood cells (RBC)	$x10^{12}$ cells/L	≤ 3.5	≥ 6.4
Hemoglobin - Females	g/L	≤ 95	≥ 175
- Males		≤ 115	≥ 190
Hematocrit - Females	%	≤ 32	≥ 54
- Males		≤ 37	≥ 60
Platelet count	$x10^9$ cells/L	≤ 75	≥ 700
Blood Chemistry			
Alanine aminotransferase (ALT)		N/A	> 3 x ULN

Parameter name	Unit	Low	High
Aspartate aminotransferase (AST)		N/A	> 3 x ULN
Alkaline phosphatase (ALP)		N/A	>3 x ULN
Total bilirubin		N/A	>2 x ULN
Blood urea nitrogen (BUN)		N/A	≥10.7 mmol/L
Creatinine		N/A	≥176.8 umol/L
Sodium	mmol/L	≤ 126	≥ 156
Potassium	mmol/L	≤ 3	≥ 6
Chloride	mmol/L	≤ 90	≥ 118
Bicarbonate	mmol/L	≤ 16	≥ 35
Glucose (non-fasting)	mmol/L	≤ 2.2	≥ 9.7
HbA1C	%	N/A	≥ 6.5%
Calcium	mmol/L	≤ 2	≥ 3
Phosphorus	mmol/L	≤ 0.5491	≥ 1.7119
Albumin	g/L	≤ 25	≥ 100
Total protein	g/L	≤ 45	≥ 100

Liver Function Abnormalities

The number and percentage of subjects with the following potentially clinically significant liver function laboratory abnormalities:

- post-baseline ALT ≥3x ULN or ≥5x ULN
- post-baseline AST ≥3x ULN or ≥5x ULN
- post-baseline total bilirubin >1.5x or ≥2x ULN
- post-baseline AST or ALT ≥3x ULN accompanied by concurrently elevated total bilirubin ≥2x ULN, where concurrent is defined as on the same day.

4.5.4 Vital Signs

Vital sign parameters include temperature, systolic blood pressure, diastolic blood pressure, pulse rate, respiration rate and weight. The descriptive statistics for actual values and change from baseline will be summarized at each visit.

The analysis of vital signs will also focus on the incidence of clinically relevant outliers based on the following criteria. The incidence and percentage of clinically relevant outliers determined by each criterion will be summarized. A listing of subjects with clinically relevant vital signs will be provided.

Table 8: Criteria to determine clinically relevant abnormalities in vital signs

Variable	Low	High
Systolic Blood Pressure	<90 mmHg or > 15 mmHg decrease from baseline	>180 mmHg or > 20 mmHg increase from baseline
Diastolic Blood Pressure	< 50 mmHg or > 15 mmHg decrease from baseline	>105 mmHg or > 15 mmHg increase from baseline
Pulse*	< 50 bpm post-baseline or > 20 bpm decrease from baseline	>120 bpm post-baseline or > 30 bpm increase from baseline
Temperature	>38˚C or > 1˚C increase from baseline	
Respiration Rate	≤10 breaths/min post-baseline and >10 breaths/min at baseline	≥20 breaths/min post-baseline and <20 breaths/min at baseline

* bpm = beats per minute

Visit windows for by visit summaries

For vital sign by visit summaries, the analysis visit should be defined using window in the table below. If there is more than 1 value in the same analysis visit window for a certain parameter for a subject, the closest record to the target visit day will be used for the by visit analysis. If there are 2 values in the same analysis visit window with the same distance from the target visit day for a certain parameter for a subject, then the record with the later date will be used for the by visit analysis.

Table 9: Visit windows for vital sign

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Most recent non-missing pre-dose value
Week 1	8	[2, 10]
Week 2	15	[11, 21]
Week 4	29	[22, 35]
Week 6	43	[36, 49]
Week 8	57	[50, 63]
Week 10	71	[64, 77]
Week 12	85	[78, EOS]

4.5.5 C-SSRS data

The Columbia Suicide Severity Rating Scale (C-SSRS) is an assessment that evaluates suicidal ideation and behavior.

There are 5 “Yes/No” questions from suicidal ideation section with the following increasing severity order: (1) Wish to be dead, (2) Non-specific active suicidal thoughts, (3) Active suicidal ideation with any methods (not plan) without intent to act, (4) Active suicidal ideation with some intent to act, without specific plan, (5) Active suicidal ideation with specific plan and intent.

From suicidal behavior section, there are 6 “Yes/No” questions with the following order: (1) Actual attempt, (2) Has subject engaged in non-suicidal self-injurious behavior? (3) Interrupted attempt (4) Aborted Attempt (5) Preparatory Acts or Behavior (6) Suicidal behavior was present during the assessment period?

The summary table for C-SSRS will include the number of subjects who answered “Yes” to one of the 5 questions in suicidal ideation section, or answer “Yes” to one of the 6 questions in suicidal behaviour section. A listing of subjects with any suicidal ideation or suicidal behaviour at baseline or post-baseline will be provided.

4.5.6 Physical and Neurological Exams

The full physical examination will include head, ears, eyes, nose, mouth, skin, heart and lung, lymph nodes, and gastrointestinal, musculoskeletal, and neurological systems. The brief physical examination will be focused on general appearance and respiratory and cardiovascular systems, as well as on subject-reported symptoms.
The full neurological examination includes observation for cerebellar (intention) tremor and noncerebellar (e.g., resting or positional) tremor, nose-finger, heel-shin, Romberg, tandem walk, normal walk/gait, positional nystagmus, gaze-evoked nystagmus, reflexes, muscle strength, cranial nerves, and sensory function of the upper and lower extremities. The neurological examination will be performed by a physician or the equivalent.
The brief neurological examination will include an assessment of motor and sensory function, cranial nerves, reflexes, noncerebellar tremors (e.g., resting or positional), and cerebellar function. The assessment of cerebellar function will be complemented by the SARA. The brief neurological examinations will be conducted by a physician or the equivalent.

The summary table for Physical and Neurological Exams will include number of subjects who are assessed as Normal, Abnormal (Not clinically significant), Abnormal (clinically significant), and Not done/Missing for each function.

Table 10: Visit windows for Brief Physical and Neurological Examination

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Most recent non-missing pre-dose value
Week 1	8	[2, 10]
Week 2	15	[11, 28]
Week 6	43	[29, 63]
Safety follow-up	Safety follow-up	[64, the end of study]

4.5.7 SARA

Descriptive statistics will be used to summarize SARA measurements at each visit for each treatment group. Individual subject listings of SARA at each timepoint will also be provided by treatment group.

Table 11: Visit windows for SARA

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Most recent non-missing pre-dose value
Week 1	8	[2, 10]
Week 2	15	[11, 28]
Week 6	43	[29, 63]
Week 12	85	[64, one day before start day of safety follow-up]
Safety follow-up	Safety follow-up	[Start day of safety follow-up, the end of study]

4.5.8 Electrocardiogram (ECG)

The ECG data will be obtained directly from the 12-lead ECG traces. These data include the PR interval, QTcB interval, QTcF interval, QRS duration, and pulse rate. The ECG data will be listed by subject. Descriptive statistics will be used to summarize ECG measurements at each visit for each treatment group. In addition, a shift table of ECG interpretation results will be presented at each visit by treatment group. Individual subject listings of ECG parameters at each timepoint will also be provided by treatment group. Plots of mean QTcB and mean QTcF over time will be also presented.

Visit windows for by visit summaries

For ECG by visit summaries, the analysis visit should be defined using window in the table below. If there is more than 1 value in the same analysis visit window for a certain parameter for a subject, the closest record to the target visit day will be used for the by visit analysis. If there are 2 values in the same analysis visit window with the same distance from the target visit day for a certain parameter for a subject, then the record with the later date will be used for the by visit analysis.

Table 12: Visit windows for ECG

Analysis visit	Target visit day	Analysis visit window
Baseline	1	Most recent non-missing pre-dose value
Week 2	15	[2, 21]
Week 4	29	[22, 35]
Week 6	43	[36, 49]
Week 8	57	[50, 70]
Week 12	85	[71, EOS]

4.5.9 COVID-19 Related Analyses

Any impacts to the study or study data due to COVID-19 will be summarised and/or listed. These include:

- Participant accountability of participants who discontinued due to COVID-19.
- COVID-19 adverse events.
- Concomitant non-drug treatment related to COVID-19.
- Protocol deviations due to COVID-19.
- Participants with protocol alternations due to COVID-19 (as reported on the COVID-19 Impact Log).

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

5. INTERIM ANALYSIS

An interim analysis (IA) for futility of the primary endpoint will be performed after approximately 50% of the subjects have completed the Week 12 Visit. This is to allow early termination of the study if it is evident that the efficacy of BIIB104 is unlikely to be achieved.

The futility criteria will be based on conditional power, which is the probability that the primary efficacy endpoint analysis will be statistically significant in favour of the treatment at the planned final analysis, given the data at the IA.

The conditional power will be calculated using PROC SEQTEST which allows for calculation under different hypothesised treatment effects and assumed variance. The treatment effects and variance to be used for calculating the conditional power will be determined as follows:

- The treatment effects will be assumed to be equal to the observed values at the IA, *i.e.* the maximum likelihood estimate (MLE).
- Prior to unblinding at the interim analysis, the pooled variance will be assessed to project the likely amount of statistical information at the final analysis. If the pooled variance is less than the value assumed in the study design the pooled variance will be used to calculate the assumed quantity of information at the final analysis. If the pooled variance is greater than the value assumed in the study design the value assumed in the study design will be used to calculate the assumed quantity of information at the final analysis. This approach ensures the futility analysis is conservative for falsely declaring futility.
- For additional context, conditional power under other hypotheses may be calculated.

No test for superiority will be performed at the IA, while the LS mean differences and corresponding standard errors will be derived from the primary analysis model.

The interim analysis results will be reviewed by a Biogen internal unblinded team. Given the insufficient knowledge of BIIB104's potential effects on various cognitive and functional endpoints at this clinical stage, selected analyses on key efficacy endpoints in addition to the pre-specified futility criteria will also be considered for the decision to stop one or both BIIB104 dose arms. Details of these analyses and criteria for futility are documented separately.

6. REFERENCES

Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. *J Biopharm Stat*. 2013;23(6):1352-71.

ICH E9 (R1): Addendum to statistical principles for clinical trials on choosing appropriate estimands and defining sensitivity analyses in clinical trials. Endorsed by the ICH Steering Committee. 2014.

ICH E9 (R1): Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials. European Medicines Agency. 2017.

Rubin, D. B. (1976) “Inference and missing data.” *Biometrika*, 63(3), 581-592.

Signature Page

Document Name: BIIB104 263CS201 SAP v2.0 Final

Document Title: A Phase 2, Randomized, Double-Blind, Multiple-Dose, Placebo-Controlled Study to Evaluate the Safety and Efficacy of BIIB104 in Subjects with Cognitive Impairment Associated with Schizophrenia (CIAS)

Signed by	Role	Date / Time (UTC)
[redacted]	Signing as Author	04-Apr-2022 15:46:05
	Signing as Approver	05-Apr-2022 11:57:08
	Signing as Approver	04-Apr-2022 18:29:41
	Signing as Approver	04-Apr-2022 16:12:31