A Multicenter, Follow-Up Study to Assess Long-Term Electrophysiologic and Clinical Outcomes in Subjects Previously Enrolled in Study 215ON201

Protocol 215ON203/NCT02657915 Statistical Analysis Plan

Study Phase: 2
Product Studied: BIIB033
Date of Protocol: 16 November 2015 (version 3)
Date of SAP: 02 Mar 2017 (version 2)

Key words: (Follow-Up study, FF-VEP, ANCOVA, multiple sclerosis, adverse events)

Approved By:

[redacted] M.S..	Date 02 Mar 2017
[redacted], Ph.D.	Date 2 Mar 2017
[redacted] MD	Date 2 Mar 2017
[redacted] MD	Date

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP)

Confidentiality Statement

A Multicenter, Follow-Up Study to Assess LongTerm Electrophysiologic and Clinical Outcomes in Subjects Previously Enrolled in Study 215ON201

Protocol 215ON203 Statistical Analysis Plan

Study Phase: 2
Product Studied: BIIB033
Date of Protocol: 16 November 2015 (version 3)
Date of SAP: 03 Feb 2017 (version 2)

Key words: (Follow-Up study, FF-VEP, ANCOVA, multiple sclerosis, adverse events)

Approved By:

[redacted], M.S.. Date

[redacted], Ph.D. Date

[redacted] MD Date

[redacted] Date

06 Mar 2017

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP)

Confidentiality Statement

The information in this document contains trade secrets and commercial information that are privileged or confidential and may not be disclosed unless such disclosure is required by applicable law or regulation. In any event, persons to whom the information is disclosed must be informed that the information is privileged or confidential and may not be further disclosed by them. These restrictions on disclosure will apply equally to all future information supplied to you which is indicated as privileged or confidential.

A Multicenter, Follow-Up Study to Assess LongTerm Electrophysiologic and Clinical Outcomes in Subjects Previously Enrolled in Study 215ON201

Protocol 215ON203 Statistical Analysis Plan

Study Phase: 2
Product Studied: BIIB033
Date of Protocol: 16 November 2015 (version 3)
Date of SAP: 03 Feb 2017 (version 2)

Key words: (Follow-Up study, FF-VEP, ANCOVA, multiple sclerosis, adverse events)

Approved By:

[redacted], M.S.. Date

[redacted], Ph.D. Date

[redacted], MD Date

[redacted] Date

06 Mar 2017

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP)

Confidentiality Statement

The information in this document contains trade secrets and commercial information that are privileged or confidential and may not be disclosed unless such disclosure is required by applicable law or regulation. In any event, persons to whom the information is disclosed must be informed that the information is privileged or confidential and may not be further disclosed by them. These restrictions on disclosure will apply equally to all future information supplied to you which is indicated as privileged or confidential.

TABLE OF CONTENTS

1 STUDY OBJECTIVES AND ENDPOINTS

1.1 Primary Objective and Endpoint

The primary objective of this study is to assess FF-VEP latency in subjects who were enrolled in Study 215ON201 2 years (+ 4 months) after the last study visit.

The primary endpoint that is related to this objective is the change in FF-VEP latency of the affected eye as compared to the baseline of the fellow eye at 2 years (+ 4 months) after the last study visit assessment (Week 32) in Study 215ON201.

1.2 Secondary Objectives and Endpoints

The secondary objective is to assess clinical progression and severity of central nervous system (CNS) demyelinating disease in subjects who were enrolled in Study 215ON201 2 years (+ 4 months) after the last study visit.

The endpoints that relate to this objective are as follows:

- Evaluate incidence of clinically definite multiple sclerosis (CDMS) and time to diagnosis of CDMS.
- Evaluate severity of CNS demyelinating disease with Expanded Disability Status Scale (EDSS), Symbol-Digit Modalities Test (SDMT), and multiple sclerosis functional composite (MSFC). The MSFC includes the following:
  - Timed 25-Foot Walk (T25FW)
  - 9-Hole Peg Test (9HPT) [dominant and nondominant hands]
  - (3-Second) Paced Auditory Serial Addition Test (PASAT)
- Evaluate change in disease activity from baseline with brain magnetic resonance imaging (MRI) with and without gadolinium (Gd). MRI analysis will include the following:
  - Number of Gd-enhanced lesions
  - Volume of T2 lesions

2 STUDY DESIGN

2.1 Study Overview

This international multicenter, follow-up study will determine the long-term electrophysiologic and clinical outcomes in subjects previously enrolled in Study 215ON201.

One group is planned, with eligible subjects (previously enrolled in Study 215ON201) having already been administered at least 1 study treatment (BIIB033 100 mg/kg or placebo) by IV infusion once every 4 weeks for 20 weeks (a total of 6 doses). Dosing was completed 2 years (+ 4 months) prior to enrollment in this follow-up study. A maximum of 82 subjects (determined from the number of subjects who received at least 1 dose of BIIB033 or placebo in Study 215ON201) will be included in this study. There is no formal sample size calculation. The number of subjects eligible for this study is determined by the number of subjects who participated in Study 215ON201.

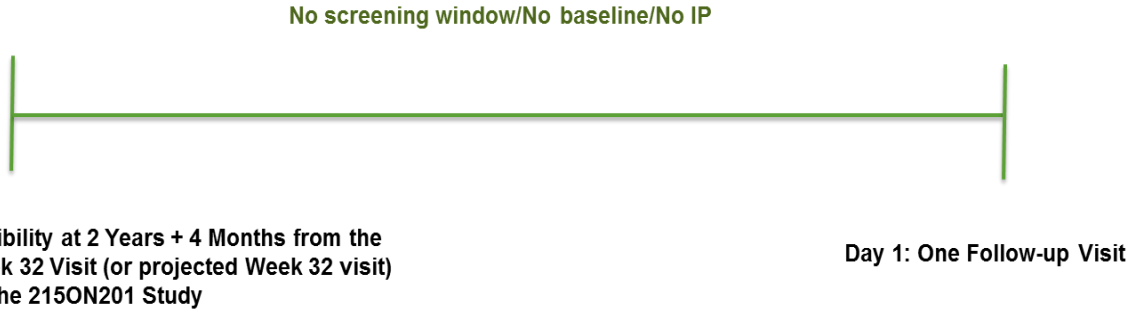
See Figure 1 for a schematic of the study design.

Study Duration:

The study period will consist of an approximately 1-day visit (with a +5-day window of the 1-day visit) that is dependent on the scanner capabilities and equipment availability at each clinical site.

2.2 Study Schematic

Figure 1: Study Design



IP = investigational product; MRI= magnetic resonance imaging.
Note: All assessments should be performed on the same day where possible or within a + 5 day window of the 1-day visit.

Table 1: Schedule of Activities

Tests and Assessments[1]	**Day 1**[2]**:** **(2 Years + 4 Months After Week 32 Visit in Study 215ON201 or Projected Week 32 Visit if the Subject Did Not Complete All Visits)**
Informed Consent	X
Eligibility Criteria Check	X
Medical History[3]	X
Multiple Sclerosis Diagnosis	X
Multiple Sclerosis Signs and Symptoms	X
Concomitant Therapy and Procedures Recording	X
Vital Signs[4]	X
Physical Examination	X
Expanded Disability Status Scale[5]	X
Symbol-Digit Modalities Test	X
Multiple Sclerosis Functional Composite Including • Timed 25-Foot Walk • 9-Hole Peg Test (Dominant and Nondominant Hands) • (3-Second) Paced Auditory Serial Addition Test	X
[redacted]	
Full-Field Visual Evoked Potential[7]	X

Tests and Assessments[1]	Day 1[2]: (2 Years + 4 Months After Week 32 Visit in Study 215ON201 or Projected Week 32 Visit if the Subject Did Not Complete All Visits)
[redacted]	
[redacted]	
Adverse Event/Serious Adverse Event Recording[12]	X

[redacted]; MRI = magnetic resonance imaging.

[1] All assessments should be performed on the same day and in the order listed, when possible (or within a + 5-day window of the 1-day visit).

[2] Completion of the Day 1 visit assessments constitutes the end-of-study visit.

[3] Medical history will be taken from after last study visit in Study 215ON201.

[4] Vital signs include temperature, systolic and diastolic blood pressure, pulse rate, body weight, and respiratory rate (after sitting for at least 5 minutes).

[5] Expanded Disability Status Scale should be performed by a study certified rater. Where applicable, refer to study manual for instructions.

[redacted]

[7] Sites and technicians must be qualified and approved by the central reader to perform MRI, [redacted], full-field visual evoked potential, and [redacted].

[redacted]

[12] Adverse event/serious adverse event monitoring will be recorded after informed consent form sign-off through the end-of-study visit.

3 STATISTICAL ANALYSIS METHODS

3.1 General Considerations

Descriptive summary statistics will be presented for all study primary, secondary and [redacted] endpoints collected.

Unless stated otherwise, summary statistics will be presented by original treatment groups and overall. For continuous endpoints, the summary statistics will generally include number of subjects enrolled, number of subjects evaluated, mean, standard deviation, median, and range (minimum, maximum). For categorical endpoints, the summary statistics will generally include number of subjects enrolled, number of subjects evaluated, and the percent of subjects in each category.

Statistical testing for efficacy endpoints will be made between the BIIB033 group and the placebo group. There will be no multiple comparison adjustments.

Unless stated otherwise, all statistical tests will be 2-sided with statistical significance level of 0.05.

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

Analysis population

Intent-to-treat (ITT) population, per-protocol population, and safety population are defined as follows.

- *Intent-to-Treat Population*

The intent-to-treat population is defined as all randomized subjects who received at least 1 dose of study treatment (BIIB033 or placebo) in 215ON201 study, and completed at least one of the 1-day assessments in this study.

- *Per-protocol Population*

The per-protocol population is defined as subjects from the ITT population who completed the 215ON201 study, did not miss more than one dose of BIIB033 or placebo, and did not receive MS modifying therapies during study 215ON201. These subjects were included in the per-protocol population of 215ON201 study.

- *Safety population*

The safety population is defined as all subjects who received at least 1 dose of study treatment in 215ON201 and completed at least one of the 1-day assessments in this study.

Baseline data

Unless stated otherwise, baseline data are defined as the data collected prior to the time and/or on the date of first dose in 215ON201. If there are more than one value on or before

the date of first dose, the non-missing value closest to and prior to (including on) the date of first dose will be used as the baseline value.

Treatment group

No study drug dosing is planned for this follow-up study. The treatment group used in the descriptive summary and treatment comparison is determined by the prior treatment with the investigational drug in 215ON201.

– new in 215ON203

- baseline T2 volume in 215ON201 (0 and >0)
- EDSS total score (>=median and < median)
- EDSS FS visual score (>=median and < median)
- RGCL thinning in the affected eye by week 4 (>=median and < median)

3.2 Study Subjects

3.2.1 Analysis population

Unless stated otherwise, all study subjects data will be summarized by treatment groups and overall in the ITT population.

3.2.2 Accounting of Subjects

The summary of subjects disposition will include: subjects who completed the study; and subjects who withdrew from study early and the reasons for withdrawal. A listing of those subjects who withdrew from study and the associated reasons for withdrawal will be presented by treatment group. Number of subjects in each analysis population defined in Section 3.1 will also be summarized.

3.2.3 Protocol Deviations

Protocol deviations identified during site monitoring will be captured in a Protocol Deviation log and categorized as major or minor deviations by Biogen. Count and percentage of subjects with at least one major deviation will be summarized by category.

3.2.4 Demographics and Baseline Disease Characteristics

Demographic data, including age (in years), age category (in years; <18, 18-30, 31-40, 41-50, and 51-55, >55), sex and race category, will be summarized using descriptive statistics. Demographic data will be also summarized in per-protocol population.

Medical history between last study visit Study 215ON201 and Day 1 is categorized into pre-specified system organ categories. Number and percent of subjects for each category will be summarized.

3.2.5 Concomitant Medications and Non-drug Treatments

All concomitant medications will be coded using the World Health Organization (WHO) medication dictionary. A concomitant medication is any drug or substance administered between the time the subject has signed the ICF and the End of Study. A concomitant non-drug treatment is any therapeutic interventions or diagnostic assessment performed between the time the subject has signed the ICF and the End of Study.

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

3.3 Efficacy Analysis

3.3.1 General consideration for efficacy analysis

Analysis population

All efficacy endpoints will be evaluated in the ITT and per-protocol population as defined in Section 3.1.

The analyses performed in the per-protocol population will be considered the primary analyses, and the analyses based on ITT population will be considered as sensitivity analyses.

Analysis models

Analysis of covariance (ANCOVA) model will be used as the primary inferential analysis approach to evaluate changes from the baseline measurement in 215ON201 to Day 1 of 215ON203 in efficacy endpoints and to test the treatment differences between BIIB033 and placebo. ANCOVA model will include terms for the baseline value and treatment group. Based on Type III sum of squares, Least-squares (LS) adjusted mean of treatment difference, 95% confidence interval for the difference and the associated p-value will be presented.

Since this is a follow up study, due to loss of follow up, if the distribution of the responses is markedly non-normal (eg., highly skewed or contains outliers), the normality assumption appears dramatically violated, the Wilcoxon rank-sum test will be used to compare the treatment difference of efficacy endpoints. Normal probability plot, and QQ-plot with tests will be used to check normality assumption.

3.3.2 Primary Efficacy Endpoint

In order to be included in the statistical analyses of the FF-VEP data the subject had to have been evaluated using the same monitor type in both the 215ON201 and 215ON203 studies.

Subjects who were evaluated using different monitor types will have their FF-VEP data excluded from the statistical analyses. These data will appear in the subject listings.

Latency as determined by FF-VEP

P100 latency value from FF-VEP is the primary efficacy outcome and will be measured for affected and fellow eyes at Day 1.

The actual value for affected eye, fellow eye and the difference between the two eyes at Day 1 will be summarized using descriptive statistics by treatment group.

Change in latency value at Day 1 for the affected eye in the current study from the baseline of fellow eye in RENEW will be analyzed using ANCOVA as mentioned in Section 3.3.1. The baseline of the fellow eye in RENEW will be used as the baseline covariate in above ANCOVA model.

Sensitivity analysis

Change in latency value at Day 1 for the affected eye in the current study from Week 32 of fellow eye in RENEW will be analyzed using ANCOVA as mentioned in Section 3.3.1. Week 32 of the fellow eye in RENEW will be used as the covariate in ANCOVA model. If Week 32 value is not available, then Week 24 value will be used to calculate the change in FF-VEP latency.

Categorical Analyses on FF-VEP latency

In addition, each subject's latency values will be compared between baseline fellow eye in RENEW and affected eye at Day 1 in the current study. For subjects whose baseline affected eye data in RENEW are missing or more than 3% worse (higher) than baseline fellow eye, their affected eye data at Day 1 will be checked if they are within 10% of baseline fellow eye measurements in RENEW or worse than 10% (if the affected eye data are missing, they will be treated as worse than 10%). The 10% cutoff will be used to categorize latency delay as normal/mild (less or equal to 10%) or moderate/severe (>10%). Count and percent will be summarized by treatment and chi-squared test will be conducted for treatment comparison.

3.3.3 Secondary Efficacy Endpoints

The secondary endpoints will include the time to diagnosis of CDMS, severity of disease evaluated by EDSS, SDMT and MSFC, and brain MRI results. Number and percentage of subjects with diagnosis of CDMS will be summarized, and descriptive statistics of time to diagnosis of CDMS will be presented. EDSS, SDMT, MSFC, and brain MRI outcomes will also be summarized using descriptive statistics.

Clinically definite multiple sclerosis (CDMS) and time to diagnosis of CDMS

The number and percentage of subjects with diagnosis of CDMS will be displayed by treatment groups. The time to diagnosis of CDMS will be summarized using descriptive statistics .

Time to confirmed MS is the time from the diagnosis of Acute optic neuritis (AON) to the confirmed MS date.

Subjects who do not have a confirmed MS will be considered censored. The censor date for subjects is the time of the last study contact. The start date for calculation of day to censor is the date of diagnosis of AON in 215ON201.

The proportion of subjects with a confirmed MS from the diagnosis of AON in 215ON201 to Day 1 in 215ON203, and the time to diagnosis of CDMS in 215ON203 from the diagnosis of AON in 215ON201 will be analyzed by the Kaplan-Meier (KM) product limit estimate. The KM estimate for the proportion of patients with MS and 10^{th} percentile, 25^{th} percentile and 50^{th} percentiles of time to MS will be summarized by treatment group.

EDSS, MSFC and SDMT

For Expanded Disability Status Scale (EDSS), and its 7 functional system scores including Visual (Converted), Brainstem, Pyramidal, Cerebellar, Sensory, Bowel and Bladder (Converted), and Cerebral, each functional system score and EDSS total score will be summarized using descriptive statistics by treatment group.

For multiple sclerosis functional composite (MSFC), MSFC z-score [1] for T25FW, 9HPT and PASAT-3 will be derived as in Appendix 2 and will be summarized using descriptive statistics by treatment group.

The MSFC Z-score is calculated by creating Z-scores for each component of the MSFC, as explained below, and averaging them to create an overall composite score, i.e.,

MSFC Z-score = (Z25-foot-walk + Z9HPT + ZPASAT-3)/3, where Zj refers to Z-scores of component j.

A Z-score represents the number of standard deviations a subject's test result is higher ($Z > 0$) or lower ($Z < 0$) than the average test result ($Z = 0$) from the reference population. For this study, the reference population will be the subjects from Biogen 999MS005 MS COG norm healthy volunteer study.

For T25FW, 9HPT dominant hand, 9HPT non-dominant hand and PASAT-3, actual assessments as continuous variables will also be summarized using descriptive statistics by treatment group.

For Symbol Digit Mobility Test (SDMT), the SDMT scores will be summarized using descriptive statistics by treatment group .

MS Signs and Symptoms

MS signs and symptoms are listed in 9 classes including Vision Problems, Cognitive Impairment, Coordination/Balance Problems, Bladder Dysfunction, Bowel Dysfunction, Sexual Function Problems, General Signs and Symptoms, Sensory Disturbances, and Motor Disturbances. The unlisted signs and symptoms are recorded under "Other Signs and Symptoms" in case report form (CRF).

Number and percentage of subjects having each listed MS signs and symptoms, and Other Signs and Symptoms will be displayed by treatment group.

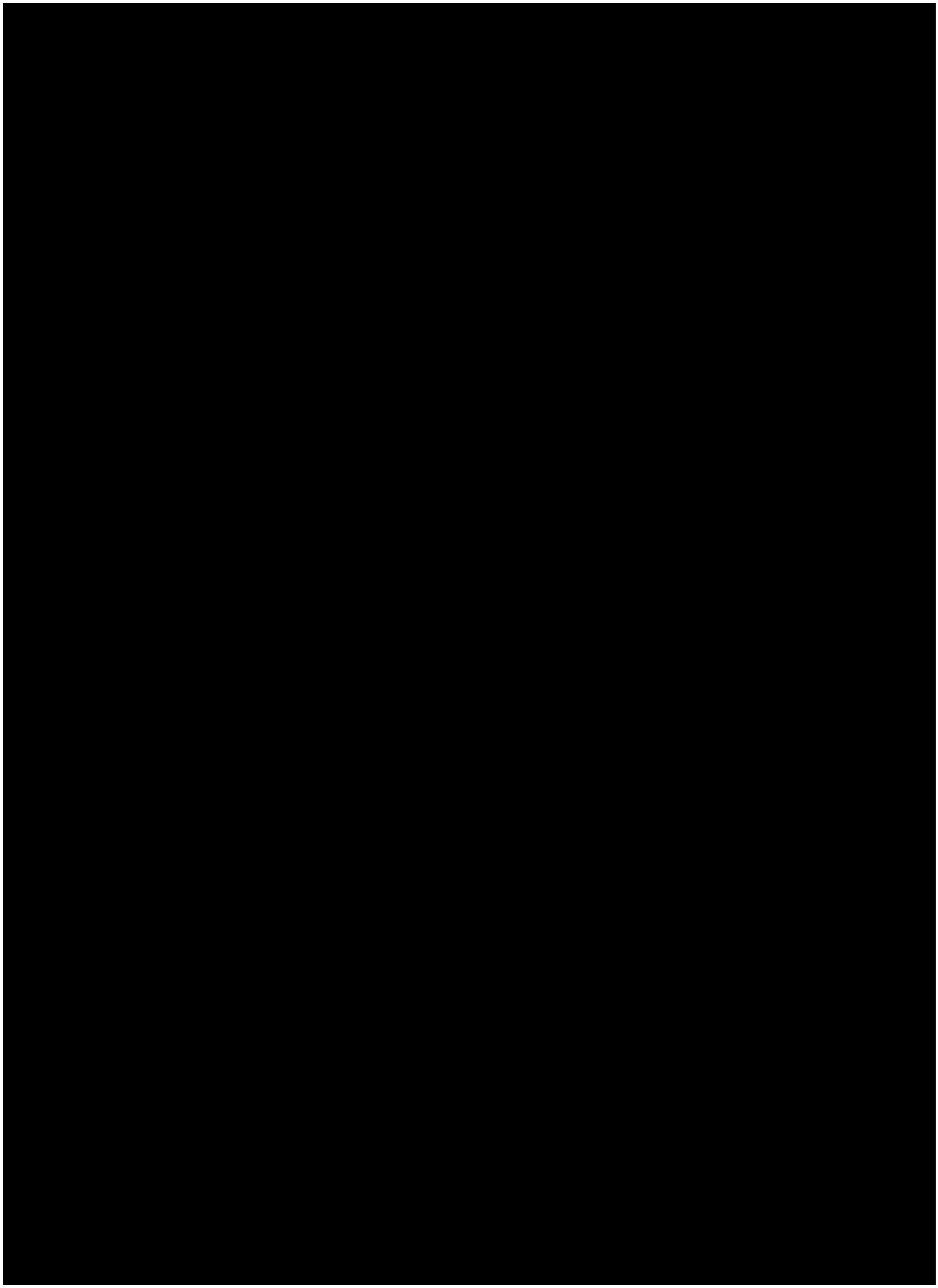
Disease Activity by Brain MRI Metrics

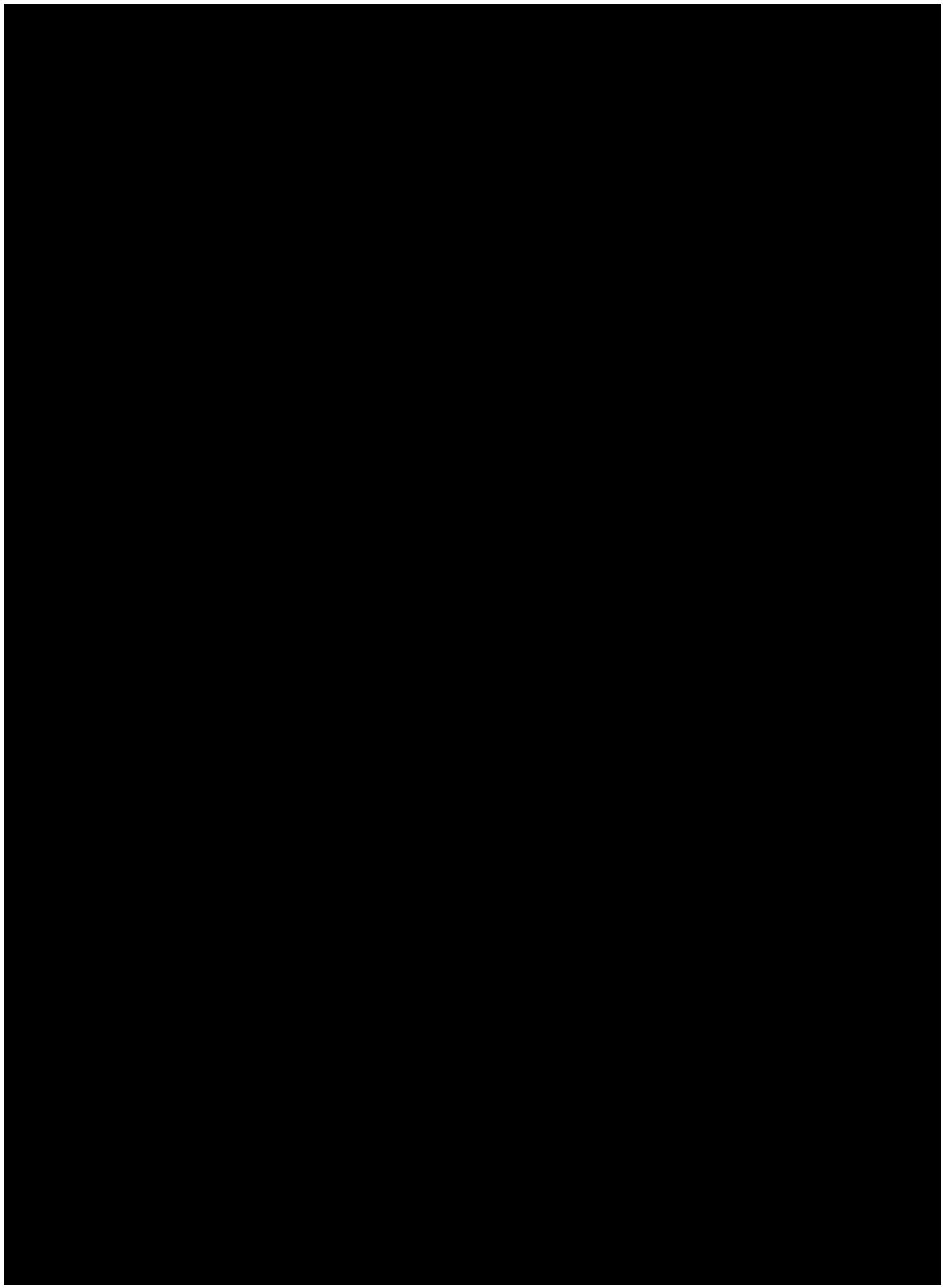
Disease activity is measured by consensus Gd lesion count, and T2 lesion volume.

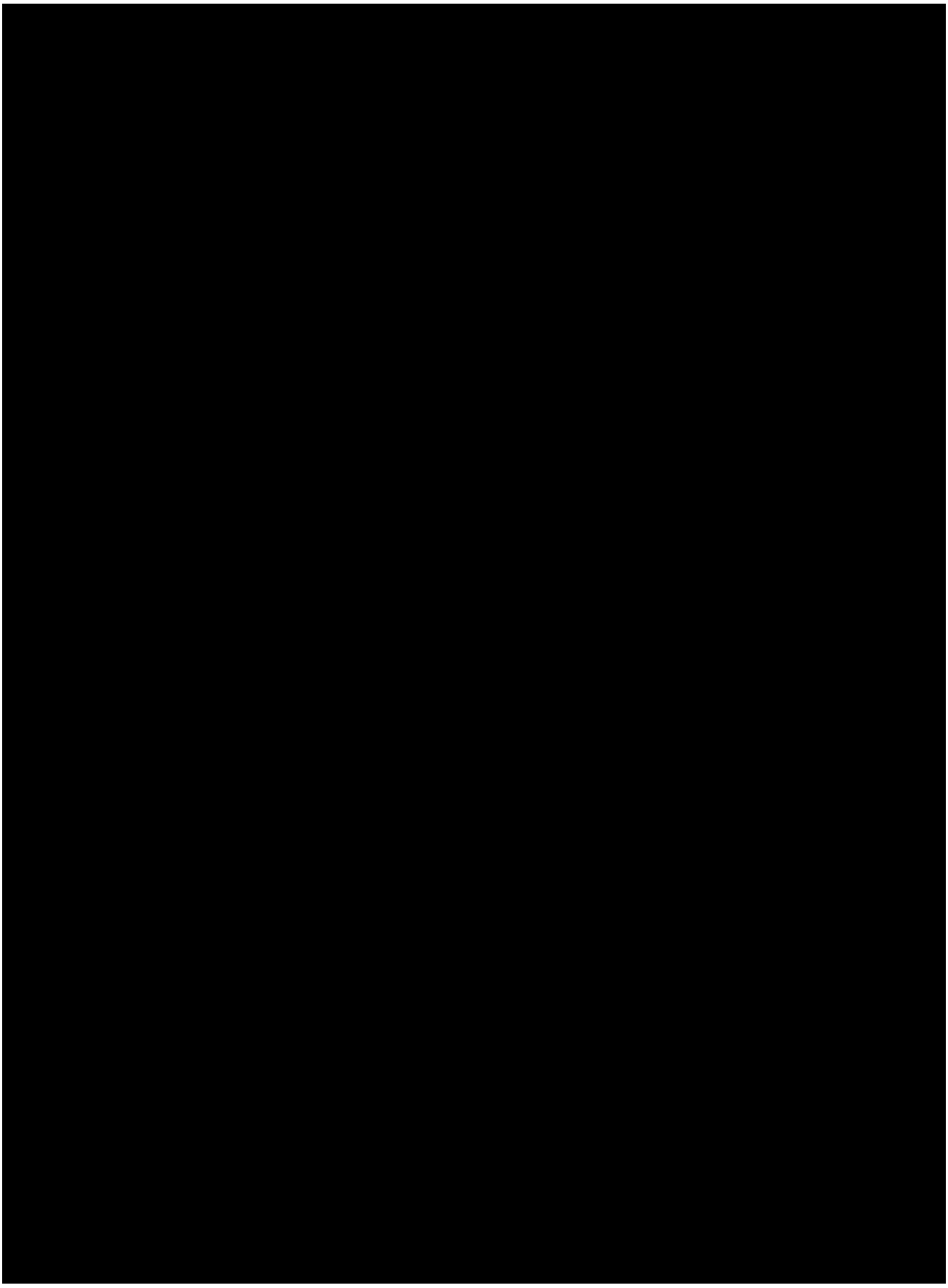
The number of consensus GD-enhancing lesions and volume of T2 Lesions (in mL), and change from RENEW baseline of them at Day 1 in the current study will be summarized by treatment group.

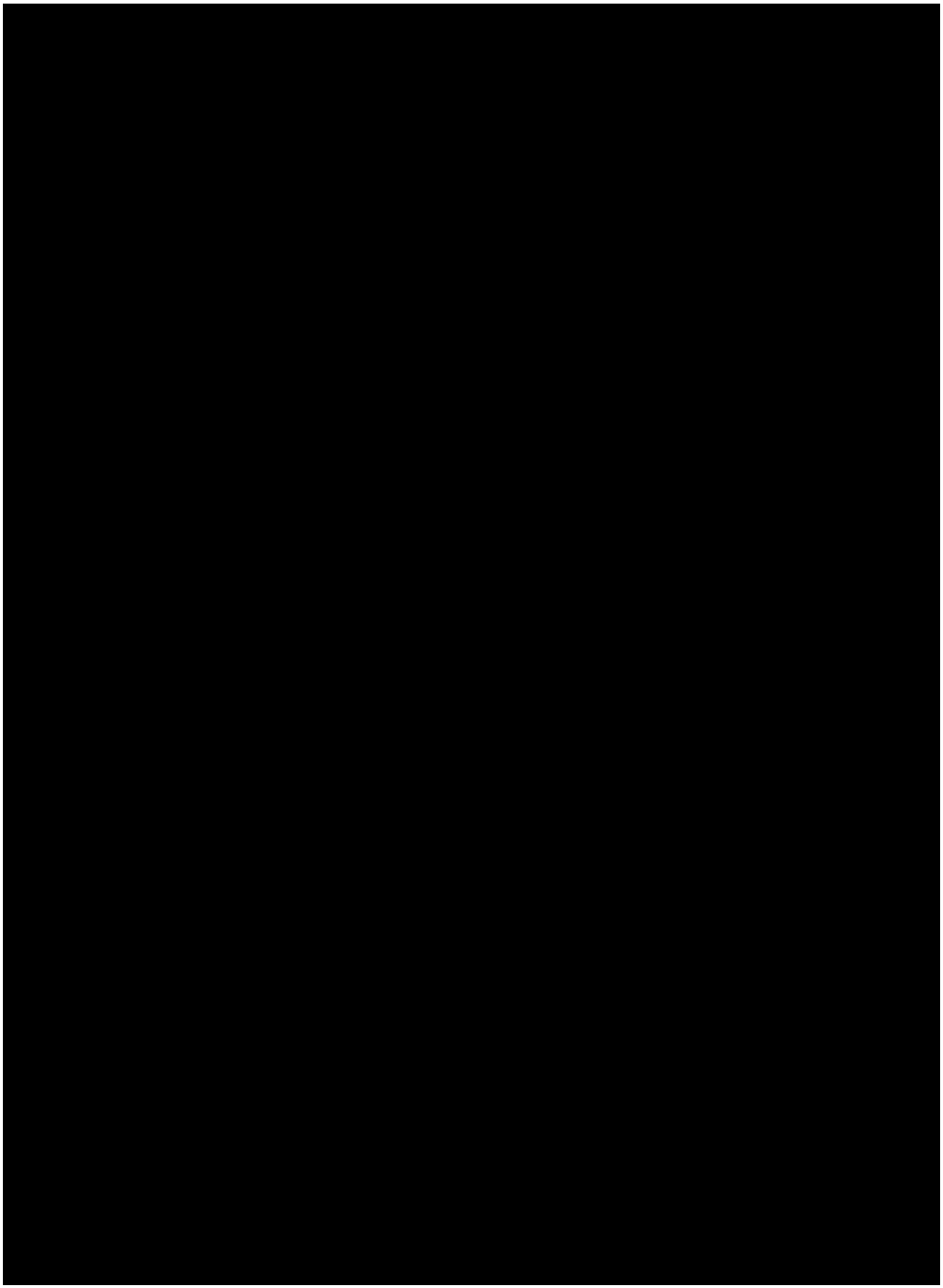
Change from baseline of the number of Gd-enhancing lesions at Day 1 will be compared between treatment groups using a Wilcoxon signed-rank test.

Change from baseline of the volume of T2 lesions at Day 1 will be also compared between treatment groups using the Wilcoxon signed-rank test.









[redacted]

3.4 Safety analysis

3.4.1 General consideration for safety analysis

Analysis population

All safety endpoints will be evaluated in the safety population as defined in Section 3.1.

Methods of Analysis

All adverse events (AEs) and serious adverse events (SAEs), vital sign measurements, physical examination findings, will be evaluated for safety. Safety data will be summarized using descriptive statistics by treatment group and overall.

3.4.2 Clinical Adverse Events

For this study, any AE experienced by the subject after ICF sign-off through the end-of-study visit is to be recorded on the CRF, regardless of the severity of the event or its relationship to the investigational drug (administered in Study 215ON201) or procedure(s) in this study. Any SAE experienced by the subject between the time the subject has signed the ICF and the End of Study Visit is to be recorded on an SAE Form, regardless of the severity of the event or its relationship to the investigational drug (administered in Study 215ON201) or procedure(s).

All AEs will be coded using the MedDRA. The overall summary table of AEs will present the number of subjects with the following events for each treatment group and overall. A subject is counted only once in each category.

- AE;
- AE with severity as “moderate” or “severe”;
- AE with severity as “severe”;
- investigational drug (administered in Study 215ON201) related AE
- study procedure(s) related AE;
- SAE;
- AE leading to withdrawal from study.

The incidence of AEs will be summarized using the primary system organ class and preferred term. Preferred terms are presented by decreasing incidence in the total column within each system organ class. The incidence of AEs will also be summarized by event severity and by relationship to the investigational drug (administered in Study 215ON201) or study procedure(s) using system organ class and preferred term. Within each system organ class or/and preferred term, the same subject will be counted only once. If a subject

has more than one AE in the same system organ class or/and preferred term, the most severe AE will be reported.

Additionally, the incidence of AEs will be summarized by using preferred term only, and preferred terms will be ordered by decreasing frequency of AEs in the BIIB033 100mg/kg column.

The incidence of SAEs will also be summarized by primary system organ class and preferred term and treatment group as well.

Listings of all AEs, SAEs, and AEs that led to study withdrawal will be presented.

3.4.3 Vital Sign Measurements

The analysis of vital signs will focus on the incidence of potentially clinically relevant abnormalities. The criteria for potentially clinically relevant abnormalities are given in Table 2.

Table 2: Criteria to Determine Potentially Clinically Relevant Abnormalities in Vital Signs

Vital Sign	Criteria for Abnormalities
Temperature	>38°C
Pulse	>100 beats per minute (bpm) <40 bpm
Systolic Blood Pressure	>160 mmHg <90 mmHg
Diastolic Blood Pressure	>100 mmHg <45 mmHg

A summary table for subjects with any potentially clinically relevant abnormalities will be provided. In the summary table, entries are numbers of subjects with an abnormality divided by number of subjects evaluated followed by corresponding percentages. Number evaluated is the number of subjects who had a baseline assessment and at least one post-baseline assessment for that vital sign.

The count and percentage of subjects with weight change of greater than 7% from baseline at Day 1 will be summarized by treatment group. A subject listing of weight change of greater than 7% from baseline at Day 1 will be presented.

For temperature, pulse, systolic blood pressure, diastolic blood pressure, and weight, actual values and changes from baseline will also be summarized using descriptive statistics by treatment group at day 1.

A subject listing will also be presented for subjects with any abnormalities in vital signs. In this listing, each subject's complete vital sign values including the baseline values will be listed with abnormalities labeled.

3.4.4 Physical Examination

The result collected on CRF at Day 1 will be listed as performed or not performed, and abnormality will be reported as AE.

4 INTERIM ANALYSES

An interim analysis will be performed when approximately 50% of potentially projected available subjects have been enrolled into the study to provide information to assist the interpretation and design of ongoing clinical development of CNS remyelinating therapies.

The interim analyses will only include the following specified analyses:

- Demography as described in Section 3.2.4 including age, age category, sex and race category
- Analysis of latency as determined by FF-VEP as described in Section 3.3.2
- Analysis of time to diagnosis of CDMS as described in Section 3.3.3
- Descriptive statistics of EDSS, T25FW, 9HPT (dominant/non-dominant), PASAT 3 and SDMT
- [redacted]
- [redacted]
- Correlation analysis between new assessments at Day 1 (EDSS, T25FW, 9HPT (dominant/non-dominant), PASAT 3, SDMT, [redacted]) and clinical and electrophysiological recovery from baseline (FF-VEP, [redacted]) at Day 1.

5 SAMPLE SIZE DETERMINATION

The sample size is based on the number of subjects who participated in Study 215ON201 (approximately 82) rather than statistical considerations. The subjects eligible for this study must have participated in Study 215ON201.

6 REFERENCE

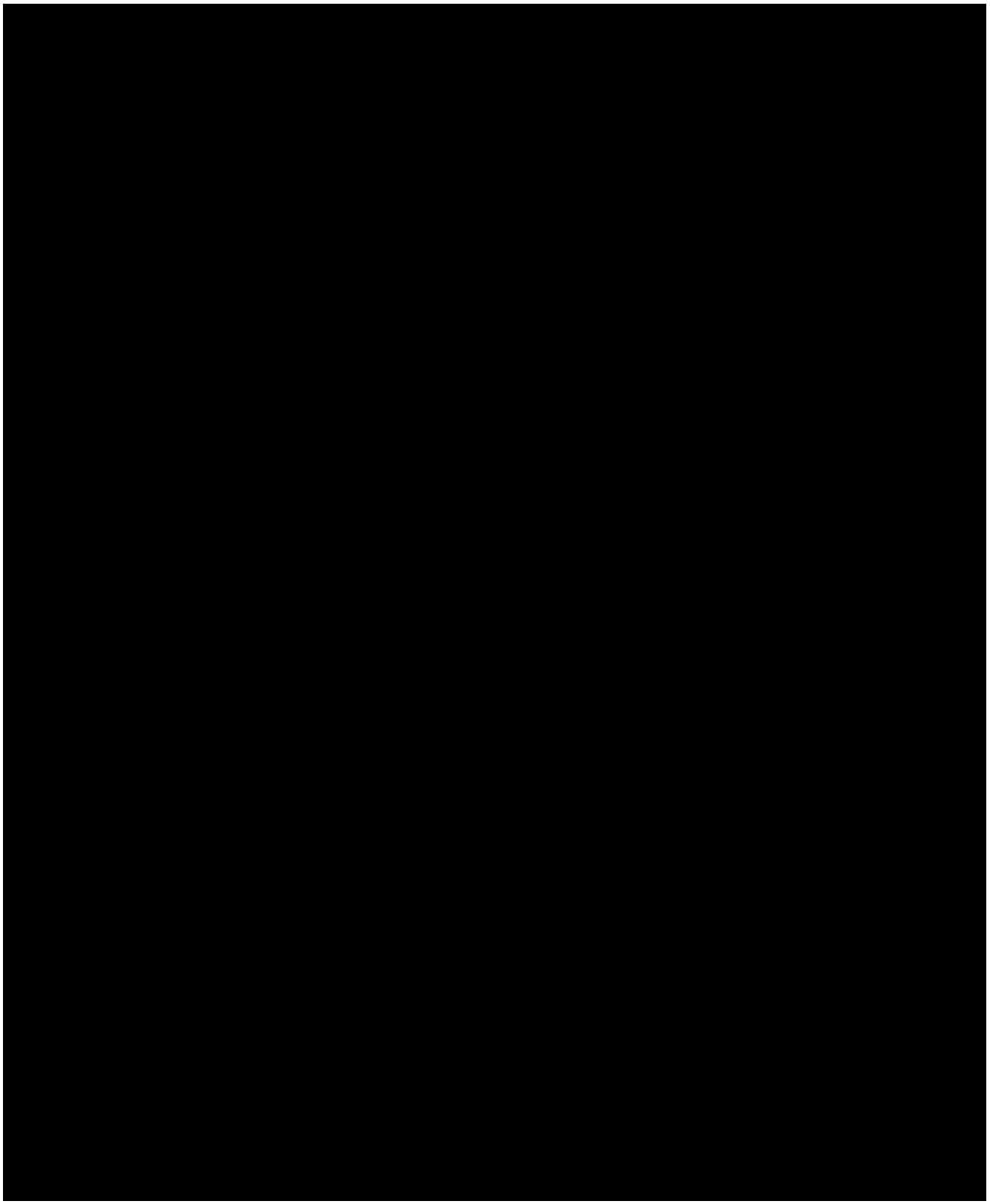
[1] Fischer, J., Jak, A.J., Kniker, J.E., Rudick, R.A. (2001). **MSFC administrative and scoring manual**. (http://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Brochures/10-2-3-31-MSFC_Manual_and_Forms.pdf)

[2] Mangione, C.M. (2000). **The National Eye Institute 25-item visual function questionnaire (VFQ-25): Version 2000. Scoring Manual**. (Accessible at: www.nei.nih.gov/resources/visionfunction/manual_cm2000.pdf.)

[3] Raphael, B. A., Galetta, K. M., Jacobs, D. A., Markowitz, C. E., Liu, G. T., Nano-Schiavi, M. L., Galetta S. L., Maguire, M. G., Mangione, C. M., Globe, D. R., Balcer, L. J. (2006). **Validation and test characteristics of a 10-item neuro-ophthalmic supplement to the NEI-VFQ-25**. *American journal of ophthalmology*, 142(6), 1026-1035.

[4] Suner, I. J., Kokame, G. T., Yu, E., Ward, J., Dolan, C., Bressler, N. M. (2009). **Responsiveness of NEI VFQ-25 to changes in visual acuity in neovascular AMD: validation studies from two phase 3 clinical trials**. *Investigative ophthalmology & visual science*, 50(8) 3629-3635.

7 APPENDIX 1: SELECTIVE EXAMPLE SAS CODES FOR THE STATISTICAL ANALYSIS IN SECTION 3.3.1 AND 3.3.3



8 APPENDIX 2: MULTIPLE SCLEROSIS FUNCTIONAL COMPOSITE (MSFC)[1]

There are 3 components in the MSFC (Fischer et al., 2001):

(1) The average time from the 4 trials on the 9-hole peg test (HPT) (the 2 trials for each hand are averaged, converted to the reciprocals of the mean times for each hand and then the 2 reciprocals are averaged);

(2) The average time of the 2 trials in Timed 25-Foot Walk;

(3) The number of correct answers on the Paced Auditory Serial Addition (PASAT) 3 test.

The MSFC Z-score is calculated by creating Z-scores for each component of the MSFC, as explained below, and averaging them to create an overall composite score, i.e.,

MSFC Z-score = $(Z_{25\text{-foot-walk}} + Z_{9HPT} + Z_{PASAT\text{-}3})/3$, where Z_j refers to Z-scores of component j.

A Z-score represents the number of standard deviations a subject's test result is higher (Z >0) or lower (Z <0) than the average test result (Z = 0) from the reference population. For this study, the reference population will be the subjects from Biogen 999MS005 MS COG norm healthy volunteer study, and the overall scores are given in Table 3.

Table 3: The overall scores of T25FW, 9HPT and PASAT 3 from Biogen 999MS005 MS COG study

	T25FW	1/9HPT	PASAT-3
mean	4.2	0.052	45.6
SD	0.95	0.0057	11.05

The Z-score for each component is defined as below:

Timed 25-Foot Walk Z-score Calculation

The Z-score is calculated as in the following formula:

$$Z_{25\text{-foot-walk}} = (-1) \times \frac{((t_1 + t_2)/2) - MEAN(reference)}{SD(reference)},$$

where t_1 and t_2 are the time (in seconds) from the two trials, and the MEAN (reference) and SD (reference) are the mean and standard deviation of the averaged time from two trials for the reference population.

If a subject is missing only one trial due to reasons other than MS related physical limitation, then the time from the non-missing trial will be used to calculate the Z-score. If both trials are missing due to reasons other than MS related physical limitation then the Z-score is regarded as missing.

If either trial is missing due to MS related physical limitation, then the subject will be given a Z-score -13.7 for T25FW.

9HPT Z-score Calculation

Z-score for the 9HPT is calculated as in the following formula:

$$Z_{9HPT} = \frac{((\frac{1}{(t_{11}+t_{12})/2} + \frac{1}{(t_{21}+t_{22})/2})/2) - MEAN(reference)}{SD(reference)},$$

where $t_{11}, t_{12}, t_{21}, t_{22}$ are the time (in seconds) taken to complete for dominant hand trial 1, dominant hand trial 2, non-dominant hand trial 1, and non-dominant hand trial 2, respectively. The MEAN (reference) and SD (reference) are the mean and standard deviation of averaged reciprocals of time from two hands for the reference population.

If one trial of either hand is missing due to reasons other than MS related physical limitation, then the time from the non-missing trial will be used to calculate the Z-score. If both trials of one hand are missing due to reasons other than MS related physical limitation, then the time from this hand is set as missing, and the non-missing hand will be used to calculate the Z-score. For example, if only one trial is performed per hand, then the Z-score is calculated as an average of the reciprocals of the two times; if both trials of the non-dominant hand are missing, then the Z-score is calculated based on the reciprocal of the mean time from the dominant hand.

If either trial from one hand is missing due to MS related physical limitation, then 777 seconds are assigned to the disabled hand, and Z-score is calculated using the average of the inverse of each hand.

PASAT-3 Z-score Calculation

For the PASAT-3, the z-score for each subject will be calculated by subtracting the reference population's mean score from the subject's PASAT-3 score and dividing by the SD of the reference population, i.e.

$$Z_{\text{PASAT-3}} = \frac{\text{PASAT3 raw score} - MEAN(reference)}{SD(reference)}.$$

If a subject is missing the PASAT-3 score due to MS disease limitations, the subject will be given a Z-score calculated using a raw score of 0. If missing score is due to other reasons, then the Z-score is regarded as missing.