

Clinical Development

FTY720D (fingolimod)

Clinical Trial Protocol CFTY720D2312 / NCT01633112

**A 12-month, randomized, rater- and dose-blinded study to compare the efficacy and safety of fingolimod 0.25 mg and 0.5 mg administered orally once daily with glatiramer acetate 20 mg administered subcutaneously once daily in patients with relapsing-remitting multiple sclerosis**

### **RAP Module 3 – Detailed Statistical Methodology**

**Author:**

[REDACTED]

**Document type:** RAP Documentation

**Document status:** Final Version 2.6 Addendum 1

**Release date:** 02Nov2018

**Number of pages:** 79

Property of Novartis

Confidential

May not be used, divulged, published or otherwise disclosed  
without the consent of Novartis

### Document History – Changes compared to previous version of RAP module 3.

Version	Date	Changes
Final V2.1	28MAY2014	Changes to follow updated protocol amendment 1
Final V2.2	06APR2015	Added sensitivity analysis to the primary analysis by excluding one misconduct site
	26AUG2015	Updated per protocol final amendment 2
Final V2.3	17FEB2017	Updated the sample size calculation per study early termination approved by FDA in December 2016
	24MAR2017	[REDACTED]
Final V2.4	20OCT2017	Added two Canada provinces (British Columbia and Manitoba) in Section 3.9 (Pooling of centers). Updated the sections related to C-SSRS per specific eCRF implemented via IVRS for this study. There is no visit information in IVRS data, so a visit window is added for C-SSRS.
	06DEC2017	Clarified nominal visits for dermatology and OCT assessment to be consistent with the TLF shells.
	28DEC2017	Updates to MRI model specifications to match with the protocol.
	10JAN2018	Vital signs: updated per the study data collection specifics (sitting for old CRFs and supine/standing for later CRFs). Data summaries are updated to present for each position separately.
	10JAN2018	Bradycardia event from first dose monitoring: Clarified that bradycardia events that do not have symptoms are not collected in the bradycardia CRFs and will be identified from the Adverse Event CRF.
	07FEB2018	Added a sentence to specify MedDRA version. Final sample number is updated to 1064, instead of 1063. Added wording about two discontinuation listings per FDA request.
	09FEB2018	Moved C-SSRS section to Safety Section 8.9.
	02MAY2018	Clarified last dose date definition. Clarified regular ECG summaries at Baseline and Month 12/EOT to

Version	Date	Changes
		align with RAP M7.1 and study protocol. Added the decision about following the old Novartis PD process, instead of the new Novartis PD process. Removed the figures for individual subjects meeting liver function safety criteria and those meeting hematology notable criteria as they do not provide any additional information to the existing abnormality tables per dry run 1 team decision. Removed the section about baseline comparability tests to be consistent with the recently completed FTY2201 study and considering this is a randomized study.
	04MAY2018	[REDACTED]
	10MAY2018	Updated Section 6 per [REDACTED] comments (NVS). Updated to remove 'unenhanced' in the label for T1 lesion to be consistent with source data label and also per [REDACTED] comment (NVS).
	14MAY2018	Updated Section 3.9 Pooling of Centers by added a few more US states per the final clinical document. Added a complete list of centers with geographic locations and region classification in Appendix 1.
Final 2.5	22Aug2018	For PFT section, a note is added to clarify how to handle a few outliers (unrealistic values) in the final analysis.
	28Aug2018	Section 8.1.2.1 Occurrence of SAEs and NSAEs: Added the entire section to add two tables to meet the CTSD requirement. Section 8.1.3 Data summaries for AEs: Added two bullets for the two new tables described in Section 8.1.2.1.
	10Sep2018	Section 3.6 Study Medication: Added data handling rule for partial end date for the last dosing interval. Section 3.8.7 Last dose day of randomized study medication: Added data handling rule for partial or completely missing last dose date on the CMP panel. Section 4.1.3 Time in study (years) for aggregate ARR: Updated to use the new definition by excluding FU visits and removing the truncation at 367 days. Section 4.2.1.2 Supportive analysis: Moved the old definition for time in study for ARR calculation to this section and kept it as one

---

Version	Date	Changes
		supportive analysis. Section 4.2.1.2 Supportive analysis: Added additional sensitivity analysis by defining confirmed relapse strictly based on EDSS scores. Section 4.2.2.1 MRI: Added sensitivity analyses to exclude baseline assessments > 45 days before randomization and end of treatment assessments > 45 days after last dose.
	13Sep2018	Section 8.1 Adverse Events: Added summaries for risk AEs and updated wording for occurrence rate of AEs to be consistent with study D2311.
Final 2.6	02Nov2018	Section 4.2.2.1: Added new analyses for annualized rate of brain atrophy (ARBA) at Month 12/end of treatment. This is to adjust for study duration and also other MRI covariates due to some differences between treatment groups.

---

## Table of contents

Table of contents.....	5
List of tables.....	8
1 Statistical methods planned in the protocol and determination of sample size .....	10
2 Statistical and analytical plans.....	10
3 Subjects and treatments, general rules and definitions.....	11
3.1 Analysis sets.....	11
3.2 Subject disposition.....	11
3.3 Protocol deviations.....	12
3.4 Background and demographic characteristics.....	13
3.5 MS disease and medical history .....	14
3.6 Study medication .....	15
3.7 Concomitant medication/therapy .....	16
3.8 Visit windows and cutoffs for efficacy and safety analyses.....	18
3.8.1 Study Day 1 and other study days .....	18
3.8.2 Screening/Baseline .....	19
3.8.3 Nominal visits and visit windows.....	19
3.8.3.1 Nominal visits.....	19
3.8.3.2 Visits windows .....	20
3.8.4 Time points for first dose monitoring ECG and Vital signs .....	22
3.8.5 Safety data cutoff and visit windows for summaries on follow-up set .....	22
3.8.6 Multiple assessments within visit windows .....	23
3.8.7 Last dose day of randomized study medication .....	23
3.9 Pooling of centers .....	24
4 Efficacy evaluation.....	24
4.1 Primary and key secondary efficacy variables.....	24
4.1.1 Primary efficacy variable .....	24
4.1.2 Key secondary efficacy variable.....	25
4.1.3 Time in study (years) for aggregate ARR .....	25
4.2 Statistical methodology .....	25
4.2.1 Primary analysis.....	25
4.2.1.1 Multiplicity adjustment for statistical hypothesis testing .....	26




4.2.1.2	Supportive analyses of primary endpoint.....	27
4.2.1.3	Handling dropouts in ARR calculations.....	29
4.2.1.4	Missing MS relapse dates .....	31
4.2.2	Analyses of secondary [REDACTED] variables .....	31
4.2.2.1	Magnetic resonance imaging (MRI).....	32
4.2.2.2	Relapses .....	34
[REDACTED]	[REDACTED] .....	36
[REDACTED]	[REDACTED] .....	37
4.2.2.5	EDSS Scores.....	38
[REDACTED]	[REDACTED] .....	38
[REDACTED]	[REDACTED] .....	38
7	Health-related Quality of Life (QoL) .....	40
[REDACTED]	[REDACTED] .....	40
[REDACTED]	[REDACTED] .....	40
[REDACTED]	[REDACTED] .....	40
[REDACTED]	[REDACTED] .....	41
[REDACTED]	[REDACTED] .....	41
7.3	TSQM (Version 1.4).....	41
8	Safety evaluation .....	42
8.1	Adverse events .....	43
8.1.1	Adverse event date imputation .....	43
8.1.2	Adverse events .....	44
8.1.2.1	Occurrence of SAEs and NAEs (non-serious AEs).....	45
8.1.3	Data summaries for AEs.....	45
8.2	Laboratory data .....	46
8.2.1	SI Units .....	46
8.2.2	Clinically notable criteria .....	46
8.2.3	Data Summaries for lab tests.....	50
8.3	Vital signs .....	51
8.3.1	Vital signs measurements.....	51
8.3.2	Vital signs notable criteria .....	51

---

8.3.3	Data summaries for vital signs.....	52
8.4	Electrocardiograms (ECG) .....	52
8.4.1	Corrected QT interval .....	53
8.4.2	Abnormality criteria for corrected QT interval.....	53
8.4.3	ECG findings .....	53
8.4.4	Data summaries for ECG data.....	53
8.5	Pulmonary function tests (PFT) .....	54
8.5.1	Predicted values and units of PFTs .....	54
8.5.2	Percent predicted PFTs.....	54
8.5.3	Data summaries for PFTs .....	55
8.6	Ophthalmic evaluations.....	55
8.6.1	Data summaries for ophthalmic evaluations.....	56
8.7	Dermatology assessment.....	56
8.8	First dose (second dose or restart dose) of fingolimod monitoring.....	57
8.8.1	Vital signs .....	57
8.8.2	ECG.....	57
8.8.3	Dose monitoring experience .....	57
8.8.4	Data summaries for first dose monitoring .....	57
8.9	Columbia Suicide Severity Rating Scale (C-SSRS).....	58
9	Interim analyses.....	60
10	Determination of sample size.....	60
11	Documentation of changes from the protocol.....	61
12	Clinical Study Report - Appendix 16.1.9 Documentation of statistical methods .....	62
12.1	Statistical methods and analysis outputs .....	62
12.1.1	Statistical methods .....	62
12.1.1.1	Negative binomial regression model.....	62
12.1.1.2	Rank Analysis of Covariance (ANCOVA) and Rank Analysis of Variance (ANOVA)	64
12.1.1.3	Logistic regression for proportion variables .....	66
12.1.1.4	Survival analysis for time to event variables .....	67
12.1.1.5	Between-treatment group comparisons for other efficacy variables .....	70
12.1.1.6	Analysis of Covariance (ANCOVA) for quality of life (QoL) variables.....	71

12.1.1.7	Wilcoxon sign test for quality of life (QoL) variables.....	72
12.1.2	Statistical analysis outputs .....	73
13	References .....	73
14	Appendices .....	74
14.1	Appendix 1: Geographic location (state/providence and country) of the study centers .....	74

## List of tables

Document History – Changes compared to previous version of RAP module 3.....	2
Table 1-1 The analysis sets and severity codes .....	12
Table 1-1a The derivation formula for creatinine clearance.....	13
Table 1-2 Display of drug names .....	15
Table 1-3 Definitions of time-intervals used to summarize exposure to study medication.....	16
Table 1-4 CMD date imputation.....	17
Table 2-1 Visit windows for vital signs and laboratory values .....	20
Table 2-2 Visit windows for PFTs.....	21
Table 2-3 Visit windows for ophthalmologic examination (excluding OCT) .....	21
Table 2-4 Visit windows for ECG .....	21
Table 2-5 Visit windows for C-SSRS.....	21
Table 3-1 Primary efficacy endpoint, supportive endpoints & analyses.....	27
Table 3-2 Example of imputation .....	30
Table 3-3 MRI endpoints & analyses.....	32
	35
	37
	39
Table 4-1 AE date imputation .....	44
Table 4-2 Criteria for clinically notable abnormalities .....	47
Table 4-3 The lower limit of detections of laboratory parameters.....	48
Table 4-4 Abnormal criteria .....	49
Table 4-5 Criteria for clinically notable vital signs.....	51
Table 4-6 Abnormality criteria for corrected QT interval .....	53
Table 12-1 Confidence Interval calculations .....	69





## 1 Statistical methods planned in the protocol and determination of sample size

Data will be analyzed by [REDACTED] statisticians and statistical programmers assigned to this study according to the data analysis section 9 of the study protocol and protocol amendments which is available in [Appendix 16.1.1 of the CSR](#). Important information is given in the following sections and details are provided, as applicable, in [Appendix 16.1.9 of the CSR](#).

## 2 Statistical and analytical plans

Unless otherwise stated, summary tables/figures/listings will be on all subjects included in the respective analysis set.

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum will be presented.

The contrast for between-treatment comparisons (ex. ARR, ARR-ratio) and corresponding p-values will be provided in section 16 SAS output tables. But between treatment comparison and corresponding p-value for the comparison between the two FTY doses (ex. ARR-ratio) will not be presented in the in-text and post-text tables. On in-text and post-text tables, the comparisons for FTY720 0.5 mg vs GA 20 mg, FTY720 0.25 mg vs GA 20 mg and the corresponding p-values will be displayed. In addition, the p-values generated from analysis models will be displayed up to 4 decimal places for reuse in publication as well as in CSR.

The primary objective of the study is to demonstrate that at least one dose (tested hierarchically 0.5 mg followed by 0.25 mg) of fingolimod is superior to glatiramer acetate 20 mg SC in reducing the ARR up to 12 months in subjects with relapsing-remitting MS.

[REDACTED]

A general note on considering country or region in an efficacy analysis model: An adjustment for country or region will be considered in the initial model. If the initial model can not generate estimates with region as a factor, then country or region will be removed from the model for further specified analysis.

### 3 Subjects and treatments, general rules and definitions

#### 3.1 Analysis sets

The following analysis sets will be used:

- **Randomized set (RS):** consists of all subjects who are assigned randomization numbers. The subjects in this set are called randomized subjects. This set will be used to summarize subject disposition, demographic and baseline characteristics, and protocol deviation information. Subjects will be grouped according to randomized treatment.
- **Full-analysis set (FAS):** Consists of all subjects who are randomly assigned and take at least 1 dose of study drug. Following the intent-to-treat principle, subjects will be grouped according to the assigned treatment at randomization. Efficacy analyses will be performed on the full-analysis set unless otherwise notified.
- **Per-protocol set (PPS):** Consists of all subjects in the FAS who do not have any major protocol deviations that could confound the interpretation of analyses conducted on the FAS. Major protocol deviations will be determined according to the predefined protocol deviation criteria before treatment unblinding. Any efficacy data after study drug discontinuation will be excluded. The per-protocol set will only be used for the supportive analyses of the primary efficacy variable.
- **Safety set (SS):** Consists of all subjects in the FAS who take at least 1 dose of study drug. Subjects will be analyzed according to the treatment they have actually taken. Safety and tolerability analyses will be performed on the safety set unless otherwise specified.
- **Follow-up set (FUS):** The follow-up set consists of all subjects in the safety set who have follow-up visit data or who have at least 1 safety assessment (AEs, laboratory test, vital sign measurement, PFTs or ophthalmology assessments) 46 days or more than 46 days after study drug discontinuation. Subjects will be grouped in the same way as previously described for the analysis on the safety set. Only the safety follow-up data analysis will be performed on the follow-up set.

#### 3.2 Subject disposition

Subject disposition will be summarized on the randomized set. The number and proportion of subjects who complete the study or discontinue the study prematurely along with the reason for discontinuation will be presented. The number and proportion of subjects who discontinue the study drug prematurely along with the reason for study drug discontinuation will be summarized and presented separately. A listing will also be provided with treatment group and subject number.

Additionally two types of subject discontinuation listings will be provided per FDA request.

- Discontinuation of study drug after which patients are still allowed to continue to have efficacy/safety assessment according to abbreviated schedule without receiving study drug treatment until the end of study.
- Discontinuation from the study in which no further safety/efficacy will be assessed except for follow-up.

Due to the non-specific nature of the HA communication, both types of listings are provided with the additional detail requested (under the heading “**Relevant information related to reason**”).

Reasons for screen failure will be summarized for all screened subjects.

### 3.3 Protocol deviations

Protocol deviations will be summarized by treatment group on the randomized set. The major protocol deviations (severity codes of 0, 1, and 8) resulting in subjects’ exclusion from analysis sets are defined in study CFTY720D2312 VAP Module 3 – Protocol Deviation. Please refer to study VAP M3 for details.

The analysis sets and severity codes used in the trial programming are listed in Table 1-1, which includes the analysis set codes 1-5 and the severity codes 0, 1, 5, 8 and 49 for the double-blind treatment period.

**Table 1-1 The analysis sets and severity codes**

<b>Analysis set codes</b>	<b>Severity Codes</b>				
1 = Randomized set 2 = Full analysis set 3 = Per protocol set 4 = Safety set 5 = Follow-up set	0 = Exclude subject from all efficacy analyses 1 = Exclude subject from per-protocol analysis 5 = Exclude subject from all safety analyses 8 = Exclude subject from all analyses 49 = Include subject in all analyses				
	<b>Deviation Severity Codes</b>				
<b>Analysis set (code)</b>	<b>0</b>	<b>1</b>	<b>5</b>	<b>8</b>	<b>49</b>
Randomized set (1)	Yes	Yes	Yes	Yes	Yes
Full analysis set (2)	No	Yes	Yes	No	Yes
Per protocol set (3)	No	No	Yes	No	Yes
Safety set (4)	Yes	Yes	No	No	Yes
Follow-up set (5)	Yes	Yes	No	No	Yes

Note: 'Yes' means include in the analysis set, 'No' means exclude from the analysis set.

Additionally, unblinding per protocol due to safety reasons may occur during the study. If a subject’s treatment code is revealed to any observer or study center personnel who are

intended to be blinded during the evaluation period (as marked on the study completion CRF page), that subject will also be excluded from the per-protocol set. A list of those subjects will also be presented.

**Note:** Novartis study team discussed and agreed that FTYD2312 should follow the old PD process due to the following key reasons:

- FTYD2312 Data Management has been following the old process and data is collected on paper forms
- The dates of the FTYD2312 Initial SAP (10/18/2011) and the 1st amendment of the SAP (6/3/2014) are prior to the roll-out date specified for the new PD Process (3/28/2016)

### 3.4 Background and demographic characteristics

Background and demographic characteristics include age at screening derived from date of birth and the screening assessment date, gender, race, and ethnicity collected on the demography CRF, height, and body weight recorded at baseline on the vital signs CRF, and body mass index (BMI) calculated as (body weight in kilograms) / (height in meters)<sup>2</sup>.

These variables recorded at baseline will be summarized on the randomized set by presenting frequency distributions (for categorical variables) or summary statistics (for continuous variables).

It is anticipated that in some cases baseline body weight may be missing. For baseline characteristics, the algorithm to derive baseline body weight is as follows: it is the last assessment of any evaluations done up to the time of first dose; if there is no such assessment, it will be the first assessment of any evaluations done after first dose. In the situation that baseline and post-baseline weight summaries are part of the vital sign summaries, missing body weights will be treated as missing.

In the computation of creatinine clearance using the Cockcroft-Gault formula, the last assessment of body weight made on or prior to the day when the subject takes the laboratory test will be used.

#### Table 1-1a The derivation formula for creatinine clearance

Creatinine clearance [mL/min] using Cockcroft-Gault formula (Cockcroft and Gault, 1976) =  $(140 - A) \times W / (72 \times C) \times G$

Where

*A* is age [years]

*W* is body weight [kg]

*C* is the serum concentration of creatinine [mg/dL]

*G* is a gender: *G*=1 for males and *G*=0.85 for females.

Creatinine clearance [mL/min] using abbreviated MDRD formula (Levey et al, 2000) =  $186.3 \times C^{-1.154} \times A^{-0.203} \times E \times S$

Where

*C* is the serum concentration of creatinine [mg/dL],

*A* is age [years],

*E* is ethnicity:  $E=1.212$  if patient is black, else  $E=1$ ,

*S* is gender:  $S=0.742$  (if patient is female), else  $S=1$ .

### 3.5 MS disease and medical history

Any condition entered on the relevant medical history/current medical conditions CRF will be coded using the MedDRA dictionary.

Relevant medical history/current medical condition, pulmonary history, MS disease history and duration of the disease, MS related eye history, MS medication history (including history of disease-modifying drugs and other drugs), MS symptoms and number of relapses in the past 2 years before study enrolment, and MS disease baseline characteristics including baseline MRI assessments and baseline EDSS will all be summarized on the randomized set.

The relevant medical history and continuing medical conditions will be summarized by primary system organ class (SOC) and preferred term (PT). MS medical history of other drugs will be summarized by preferred term.

For the MS disease history summary, duration of MS since diagnosis (years) will be derived as follows: (the first dose date – the MS diagnosis date + 1)/365.25; duration of MS since first symptom (years) will be derived as follows: (the first dose date – the first MS symptom date + 1)/365.25; time since the onset of the most recent relapse (months) will be derived as follows: (the first dose date – the most recent relapse onset date + 1)/30.

For the MS diagnosis date and the first MS symptom date, the following rule will be used when imputing a partial date (complete missing date will not be imputed):

- a) Only year yyyy is available. The imputed date will be 1-July-yyyy.
- b) Both year yyyy and month mmm are available. The imputed date will be 15-mmm-yyyy.

For the most recent relapse onset date, the following rule will be used when imputing a partial date:

- a) Only year yyyy is available. The imputed date will be 1-Jan-yyyy.
- b) Both year yyyy and month mmm are available. The imputed date will be 01-mmm-yyyy.

For the MS medication history of disease-modifying drugs (DMDs), the treatment naïve subjects are defined as those who never took the approved MS medications or other MS treatment as specified in Table 1-2.

Any subject previously treated with any MS therapies (marked as ‘Yes’ on “Previous MS treatment” CRF page) will not be considered treatment naïve. When displaying these drugs in the tables, the names specified in Table 1-2 will be used.

**Table 1-2 Display of drug names**

<b>Drug name collected in database</b>	<b>Drug name to be used in display</b>
Azasan®, Azimune, Azoprin, Azathine, Azoran, Imuzat, Zymurine, Imuran®	Azathioprine
Interferon beta-1a (Avonex®)	Interferon beta-1a i.m.
Interferon beta-1a (Rebif®)	Interferon beta-1a s.c.
Interferon beta-1b (Betaferon®/Betaseron®/Extavia®)	Interferon beta-1b s.c.
Glatiramer acetate (Copaxone®)	Glatiramer acetate
Natalizumab (Tysabri®/Antegren®)	Natalizumab
Rheumatex®, Trexall®, Amethopterin, MTX	Methotrexate
Aubagio	Teriflunomide
Tecfidera	Dimethyl fumarate (BG12)
New approved drugs	
Other	Other MS medications

### 3.6 Study medication

Duration of exposure to randomized study medication will be summarized on the safety set by presenting the number (and percentage) of subjects being exposed for a minimum duration of time as specified in Table 1-3 Duration in terms of number of days as shown in the second column will be used in the summary for the display of duration.

Duration of exposure is the number of days on study drug during the core study phase. The days when the subject does not take the study drug will be excluded. That is, duration of exposure will exclude periods of temporary interruption of study medication. Note: for the last dosing interval entered on the Dosage Administration Record CRF, if the dosing end date is partial date, impute the first day for missing day and January for missing month.

Summary statistics of duration of exposure to study medication will also be presented, which included mean, SD, minimum, maximum, etc.

For each treatment group, the patient-years, which will be calculated as (the sum of the number of days on study drug for all subjects in the group)/365.25, will also be presented.

**Table 1-3 Definitions of time-intervals used to summarize exposure to study medication**

Duration	Duration in terms of number of days
Any Exposure	≥1 day
≥1 week	≥7 days
≥2 weeks	≥14 days
≥1 month	≥30 days
≥2 months	≥60 days
≥3 months	≥90 days
≥6 months	≥180 days
≥9 months	≥270 days
≥12 months	≥360 days

### 3.7 Concomitant medication/therapy

Free-text records on the Concomitant medications / significant non-drug therapies CRF will be coded using NOVDTD drug dictionary (Novartis updated version of WHODRUG).

All medications recorded on the Concomitant medications / significant non-drug therapies form will be classed as prior or concomitant medications and summarized separately, in alphabetical order, by ATC class and preferred term. Tables will show number of subjects (receiving at least one drug of a particular ATC class and at least one drug in a particular preferred term) and percentages.

Prior medications will be drugs taken prior to first dose of randomized study medication but not at any time after baseline; any medication given at least once between the day of first dose of randomized study medication and the last day of randomized study medication will be a concomitant medication. Concurrent medications could start before first dose of randomized study medication but continue through after the first dose of randomized study medication. Any medication started after the discontinuation of randomized study medication will not be considered concomitant medication. Medications will be categorized into one (and only one) of above classes based on imputed start and end dates.

Concomitant medication (CMD) date imputation will be based on a comparison between the partial CMD start date and the treatment start date and if necessary, on the CMD type as well which is also recorded on the CRF (i.e., the variable CMDTYP1C with 3 possible outcomes: prior medication, concomitant medication, or prior/concomitant medication). Date comparisons will be based on the year and month values only (i.e., day values will be ignored) as shown in Table 1-4.

1. If the CMD start date year value is missing, the following imputation rules will apply.



- a. If the CMD type is either prior medication or prior/concomitant medication, the CMD start date will be imputed as the date one day before the treatment start date.
  - b. If the CMD type is concomitant medication or missing, the CMD start date will be imputed as the date one day after the treatment start date.
2. If the CMD start date year value is less than the treatment start date year value, then the CMD must have started before treatment.
  - a. If the CMD month is missing, the CMD start date will be imputed as the mid year date (01JulYYYY).
  - b. If the CMD month is not missing, the CMD start date will be imputed as the mid month date (15MONYYYY).
3. If the CMD start date year value is greater than the treatment start date year value, then the CMD must have started after treatment.
  - a. If the CMD month is missing, the CMD start date will be imputed as the year start date (01JanYYYY).
  - b. If the CMD month is not missing, the CMD start date will be imputed as the month start date (01MONYYYY).
4. If the CMD start date year value is equal to the treatment start date year value, then the month values will be compared as follows.
  - a. If the CMD month is either missing or equal to the treatment start month, the following imputation rules will apply.
    - i. If the CMD type is either prior medication or prior/concomitant medication, the CMD start date will be imputed as the date one day before the treatment start date.
    - ii. If the CMD type is concomitant medication or missing, the CMD start date will be imputed as the date one day after the treatment start date.
  - b. If the CMD month is less than the treatment start month, the CMD start date will be imputed as the mid month date (15MONYYYY).
  - c. If the CMD month is greater than the treatment start month, the CMD start date will be imputed as the start month date (01MONYYYY).

**Table 1-4 CMD date imputation**

### **3.8 Visit windows and cutoffs for efficacy and safety analyses**

This section defines the data cutoffs within which data will be included in the analysis and time points for which data will be summarized.

Due to the study design, both data collected at Month 12 visit as end of the treatment (EOT) (also as end of study (EOS) visit) and data collected at the study drug discontinuation visit as EOT will be recorded in the CRF at EOT (corresponding to Visit 777 in the database). For subjects who discontinue study drug early but continue through the study per abbreviated schedule, the EOT visit and EOS visit may be different, in which case data collected at the EOS visit will be recorded in the CRF at study completion visit (corresponding to Visit 778 in the database). For subjects whose EOT visit and EOS visit are the same, data will be collected only once at the EOT visit. Therefore, there will be no corresponding Visit 778 in the database for these subjects. In the subject completion (CMP) data set, data collected from the study drug discontinuation CRF (Visit 777) and study completion CRF (Visit 778) will be differentiated by visit number although there is no actual study visit corresponding to Visit 778 for subjects whose EOT visit and EOS visit are the same. .

#### **3.8.1 Study Day 1 and other study days**

The first day of administration of randomized study medication (i.e., the first dose date of the study drug) is defined as Study Day 1 or Day 1.

All other study days will be labeled relative to Day 1. For event dates on or after Day 1, study day for an event date is calculated as (event date – first dose date +1), which could be Day 1, Day 2, Day 3, etc. For event dates before Day 1, study day for an event date is calculated as (event date – first dose date), which could be Day -1, Day -2, etc., referring to one day, two days, etc., before Day 1, respectively. Duration of an event will be calculated as (event end date – event start date + 1).

One month is defined as 30 days.

### 3.8.2 Screening/Baseline

Screening refers to any procedures (e.g. checking inclusion and exclusion criteria) performed prior to first dose date. Subject informed consent is generally obtained prior to any assessment. Any assessment obtained prior to first dose date will be called screening assessment.

In general, a baseline value refers to the last measurement made prior to the administration of the first dose of the study drug. Assessments made on Day 1 may occur before or after the first dose. The following rules will be applied to determine which record is the last measurement prior to the first dose (i.e., the baseline value).

- Consider all records with an assessment date before or on the first dose date. Baseline is the last record after sorting all above records with respect to assessment date, visit and repeat visit number.
- Some data specific rules for determining baseline will be applied as described below or in data specific sections if applicable.

For vital signs, the average of the 3 readings from Day 1 pre-dose at each position (sitting, supine or standing) will be used as baseline for that position. If unavailable, the last available average value from visits (including unscheduled) before the first dose will be used. For sitting pulse the last available single value before the first dose will be used as baseline.

For ECG, the Day 1 pre-dose value will be used as baseline. If unavailable, the last available (including unscheduled) value before the first dose will be used.

The pre-dose time point is the record at scheduled time point 0 (i.e., variable STM2N equals 0).

### 3.8.3 Nominal visits and visit windows

#### 3.8.3.1 Nominal visits

For EDSS, [REDACTED], PRO variables (except C-SSRS), dermatology exam, MRI and OCT, the nominal visit will be used in the by-visit summaries. Data from unscheduled visits for these assessments will not be summarized unless otherwise specified.

The end of treatment or early discontinuation visit (Visit 777) will be the nominal 12 month visit for subjects who complete the study on study drug. For subjects who discontinue the study drug prematurely, the end of treatment or early discontinuation visit (Visit 777), the follow-up visit (Visit 501) and the end of study visit (Visit 778) which are not timepoint specific visits will be remapped based on the following rules:

- If these visits fall into a specific visit window and the corresponding scheduled visit is missing, then these visits will be remapped to be that scheduled visit. The visit windows are defined as within 5 days of the scheduled visit (+/-5 days from the target day at each visit).

- If these visits fall into a specific visit window and the corresponding scheduled visit (e.g., Visit X) is not missing, then these visits will be assigned a visit number with a decimal place where the visit number depends on whether the corresponding scheduled visit occurs before (e.g., Visit X.1 or Visit X.2 if Visit X.1 already exist etc.) or after these visits (e.g., Visit X-1.1 or Visit X-1.2 if Visit X-1.1 already exist etc.)
- If these visits do not fall into a specific visit window but fall in between 2 consecutive scheduled visit windows (e.g., between visit windows of Visit X and Visit X+1), then these visits will be assigned a visit number with a decimal place (e.g., Visit X.1 or Visit X.2 if Visit X.1 already exist etc.).

The remapped visits (e.g., Visit X) will be treated as nominal visits in the by-visit summaries.

For dermatology exam and OCT the above visits (Visit 777, Visit 501 and Visit 778) will not be remapped.

### 3.8.3.2 Visits windows

For data not listed in Nominal visit [Section 3.8.3.1](#), visit windows will be defined and used in the by-visit summaries. Visit windows will only be defined for post-baseline visits and applied to post-baseline data (including both scheduled and unscheduled visits). Based on the study assessment schedule (see Table 6-1 and Table 6-2 in the study protocol), visit windows are defined by a set of days around each nominal visit target day. For any assessment, there are at most 6 scheduled visits around which visit windows will be created: Months 1, 3, 6, 9, 12 and Follow-up.

Visit windows for vital signs, lab, ophthalmologic examination (excluding OCT), pulmonary function tests (PFTs), ECG, and C-SSRS are provided in Table 2-1 to Table 2-5. These visit windows will not be applied to any data related to the first dose or second dose or restart dose monitoring which will be summarized separately.

When visit windows are used, all post-baseline visits will be re-aligned, i.e., they will be mapped into one of the visit windows. E.g., if a subject's *Month 1* visit is delayed and occurs on Day 47, then it will be re-aligned to visit window *Month 2*. As a result, it is possible that several assessments may fall into one particular visit window. Statistical approaches to handle multiple visits in a given visit window are described in Multiple assessments within visit windows [Section 3.8.6](#).

**Table 2-1 Visit windows for vital signs and laboratory values**

Visit	Start day	Target Day	End day
Month 1	1	30	60
Month 3	61	90	135
Month 6	136	180	225
Month 9	226	270	315
Month 12	316	360	last dose date day +45

	Last dose date day +	
Follow-up	46	450

**Table 2-2 Visit windows for PFTs**

Visit	Start day	Target Day	End day
Month 6	1	180	270
Month 12	271	360	last dose date day +45
	Last dose date day +		
Follow-up	46	450	

**Table 2-3 Visit windows for ophthalmologic examination (excluding OCT)**

Visit	Start day	Target Day	End day
Month 3	1	90	135
Month 6	136	180	270
Month 12	271	360	last dose date day +45
	Last dose date day +		
Follow-up	46	450	

**Table 2-4 Visit windows for ECG**

Visit	Start day	Target Day	End day
Month 12	1	360	last dose date day +45
	Last dose date day +		
Follow-up	46	450	

**Table 2-5 Visit windows for C-SSRS**

Visit	Start day	Target Day	End day
Screening	-45	NA	-1
Baseline	1	1	1
Month 1	2	30	60
Month 3	61	90	135
Month 6	136	180	225
Month 9	226	270	315
Month 12	316	360	last dose date day +45
	Last dose date day +		
Follow-up	46	450	

### **3.8.4 Time points for first dose monitoring ECG and Vital signs**

For the first dose monitoring ECG, data will be summarized for the following time points:

- Day 1 pre-dose: day 1 ECG value at 0 hour per time label
- Day 1 post-dose (6 hours): day 1 ECG value at 6 hours per time label.
- Day 1 post-dose (>6 hours): the mean of all day 1 ECG values after 6 hours per time label (including unscheduled values).

Unscheduled ECG measured between the day 1 pre-dose and the 6 hours post-dose will not be summarized but reported in the data listing only. For the second dose monitoring ECG, data will be summarized similarly.

For the first dose monitoring vital signs, data will be summarized for the following time points:

- Day 1 pre-dose: day 1 vital sign value at 0 hour per time label
- 1 hour post-dose
- 2 hours post-dose
- 3 hours post-dose
- 4 hours post-dose
- 5 hours post-dose
- 6 hours post-dose

The vital sign values at specified hours per time label post-dose will be used. Extended first dose monitoring after the 6 hours post-dose will not be summarized but reported in the data listing only. For the second dose or restart dose monitoring vital signs, data will be summarized similarly.

### **3.8.5 Safety data cutoff and visit windows for summaries on follow-up set**

For safety summaries on the safety set, only data which are within 45 days (5 times the half life of FTY720, 9 days) after the last dose of study drug will be considered.

For safety summaries on the follow-up set, all data (including assessments or events more than 45 days after the last dose of study drug) will be considered.

The visit windows below will be used in summaries on the follow-up set.

- 1) Last assessment on study drug
- 2) Day 1 to 45 after study drug discontinuation

- 3) Month 3 (day 46 to 135) after study drug discontinuation
- 4) Month 6 (day 136 to 270) after study drug discontinuation
- 5) Month 12 (day 271 to 540) after study drug discontinuation

For the visit window 1), last assessment on or before the last dose date will be used. For visit windows 2) to 5), assessments collected during the specified day intervals after the last dose date will be used.

### **3.8.6 Multiple assessments within visit windows**

For visit windows defined in [Visit windows Section 3.8.3.2](#), multiple records may exist in one particular visit window. The following rules (unless otherwise specified) will be applied to obtain one value per subject per visit window which will be used in the by-visit summaries.

- For quantitative variables (i.e., continuous variables for which summary statistics and change from baseline tables will be provided), the average of the multiple records will be used.
- For qualitative variables (i.e., binary or nominal variables for which incidence rates and frequency distribution tables will be provided), the worst record will be used. It is noted that in the analyses performed, worst case is always well defined.

Note that “Last assessment on study drug” is not like other visit windows but more like a single time point. The assessment to be summarized at that time point is the last on drug observation which is the last observation a subject has while he or she is on study drug. Therefore, the above multiple assessment rules do not apply.

In addition, the multiple assessment rules do not apply to the first dose/second dose monitoring data and the restart dose monitoring data.

For tables to display worst case scenario, such as shift tables, and on listings, multiple assessments will be used as they are and no average values will be taken.

### **3.8.7 Last dose day of randomized study medication**

The last dose date of the randomized study drug will be collected on three CRFs:

- 1) The last date subject took study drug on the study completion CRF (CMP panel),
- 2) The last date subject took study drug on the study drug discontinuation CRF (CMP panel),
- 3) The last end date on the study drug administration record CRF (DAR panel).

The last dose date recorded in the CMP panel will be used. If the last dose date in the CMP panel is partial date, impute the first day for missing day and January for missing month. If the last dose date in the CMP panel is completely missing, the last dose date recorded on the DAR panel will be used.

### 3.9 Pooling of centers

Total 6 countries participate in this study (Argentina, Brazil, Canada, Chile, Mexico and US) with majority of subjects from US. Region will be defined by pooling of centers as follows and will be used in the analysis as a factor.

- Centers from Latin American countries (i.e. Argentina, Brazil, Chile, and Mexico) will be pooled as a region called Latin American region.
- Centers from US/Canada will be pooled according to their geographic location and the number of subjects randomized as follows.
  1. Centers from US states of CA, WA, OR, NV, AZ, MT, UT and Canadian western provinces of British Columbia, Alberta, Saskatchewan, and Manitoba will be pooled as a region called West region.
  2. Centers from states of TX, CO, NM, KS, OK, MO will be pooled as a region called Southwest region.
  3. Centers from states of WI, IA, IL, MI, IN, OH, KY, WV will be pooled as a region called Midwest region.
  4. Centers from US states of NJ, PA, MD, DC, DE, VT, NH, MA, CT, NY and Canadian Eastern provinces of Quebec, Ontario and Nova Scotia will be pooled as a region called Northeast region.
  5. Centers from states of FL, TN, AL, GA, LA, VA, NC, SC, PR will be pooled as a region called Southeast region.

The pooling of centers is performed so that there will be sufficient number of subjects in each region. If a statistical model stratified by region does not converge due to small number of subjects or small number of events in one or more regions, then regions will be further combined as needed.

## 4 Efficacy evaluation

### 4.1 Primary and key secondary efficacy variables

#### 4.1.1 Primary efficacy variable

The primary efficacy variable is the annualized relapse rate (ARR) which will be analyzed on both the full analysis set (primary analysis) and the per-protocol set (supportive analysis). The ARR can be calculated in two ways:

1. Group level (aggregate ARR)



The ARR of a treatment group is calculated as the sum of the number of confirmed relapses of all subjects in the group divided by the sum of the number of days in study of all subjects in the group and multiplied by 365.25.

The group level ARR is the primary efficacy variable.

## 2. Subject level

The ARR of an individual subject is calculated as the number of confirmed relapses divided by the number of days in study and multiplied by 365.25. The ARR of a treatment group is the mean of ARRs of all subjects in the group.

### 4.1.2 Key secondary efficacy variable

Not applicable.

### 4.1.3 Time in study (years) for aggregate ARR

Time in study is the number of days from the first study drug dose date to the study completion visit date or the last study drug dose date whichever occurs later. It will be calculated as:

Time in study = Max (Study completion visit date, last study drug dose date) – first study drug dose date + 1

Note that the study complete visit date is defined as follows:

1. For subjects who completed study drug, or completed the abbreviated schedule of assessments after discontinuing the study drug prematurely, the study completion visit date is the Visit 7/Month 12 visit date.
2. For subjects who discontinued study drug prematurely and did not complete the abbreviated schedule of assessments, the study completion visit date is the end of study (EOS) visit date. If a subject did not have an EOS visit date, then the last study visit date excluding follow-up visit will be used

The time in study (years) will be calculated as the number of days in study/365.25.

The confirmed relapses within this time period will be counted for each subject and will be used in the ARR computation.

## 4.2 Statistical methodology

### 4.2.1 Primary analysis

The 2 doses of fingolimod will be tested hierarchically vs. glatiramer acetate in a step-down procedure. For each of the 2 fingolimod doses, the null hypothesis is that there is no difference in the

ARRs between subjects treated with fingolimod and those treated with glatiramer acetate versus the alternative hypothesis that there is a difference between the 2 treatment groups. In order to preserve the Type I experiment-wise error rate, the null hypothesis will be rejected if the observed  $p$ -value for the between-treatment comparison is less than the significance level as specified in the multiplicity adjustment procedure described later in this section.

$H_{01}: \mu_{FTY\ 0.5\ mg} = \mu_{glatiramer\ acetate}$  versus  $H_{A1}: \mu_{FTY\ 0.5\ mg} \neq \mu_{glatiramer\ acetate}$

$H_{02}: \mu_{FTY\ 0.25\ mg} = \mu_{glatiramer\ acetate}$  versus  $H_{A2}: \mu_{FTY\ 0.25\ mg} \neq \mu_{glatiramer\ acetate}$

No formal hypothesis will be tested between the 2 fingolimod doses because the study is not powered to detect a difference in treatment effect between the 2 fingolimod doses.

The hypotheses will be tested using a negative binomial regression model with log link, using treatment as main effect, using the number of relapses in the previous year before study enrollment, baseline EDSS, and baseline Gd + T1 lesion count and country (or region) as covariates.

In the negative binomial regression model, the response variable is the number of confirmed relapses for each subject. The subject's time in the study (natural log of time in years) is used as an offset variable to obtain the ARR, adjusted for the varying lengths of subject's time in the study (time in years). Study centers will be pooled to country or region (based on geographical proximity and/or known or expected regional differences in medical care) in order to minimize the impact of low enrolling centers/countries on the analysis (such as non-convergence of the analysis model). Details of pooling will be provided in the statistical analysis plan before database lock. The SAS procedure GENMOD (or other software with similar functionality) will be used to conduct the analysis.

The treatment effect of each dose of fingolimod versus glatiramer acetate will be presented as an ARR ratio ( $\mu_{FTY}^{\wedge} / \mu_{GA}^{\wedge}$ ) with corresponding 95% confidence intervals and  $p$  values. In addition, the relative reduction in ARR will be presented as "percentage change," calculated as  $(\mu_{FTY}^{\wedge} / \mu_{GA}^{\wedge} - 1)$ .

The primary analysis will be conducted on the full analysis set. Only confirmed relapses will be counted.

#### **4.2.1.1 Multiplicity adjustment for statistical hypothesis testing**

The primary hypotheses tests will be tested using a hierarchical step-down procedure to control the experiment-wise error rate.

This study is designed to compare each of 2 doses of fingolimod to glatiramer acetate based on ARR. There are 2 hypotheses being tested ( $H_{01}$ ,  $H_{02}$ ).

Because it is highly likely that the approved dose of fingolimod (0.5 mg) is more efficacious than the lower dose (0.25 mg), the approved dose of fingolimod is initially tested against glatiramer acetate at a 2-sided significance level of 0.05. Only if this initial test  $H_{01}$  is rejected, will the low dose of fingolimod, be tested against glatiramer acetate also at a 2-sided significance level of 0.05.

Different scenarios of the anticipated treatment effect between fingolimod 0.5mg or 0.25mg versus glatiramer acetate have been evaluated based on the available data on fingolimod and the cumulative literature on glatiramer acetate. A 30% treatment benefit of fingolimod 0.5mg over glatiramer acetate could be anticipated on the basis of double-blind, 24-months, placebo-controlled studies of glatiramer acetate ([Johnson et al, 2001](#)) and fingolimod [Study CFTY720D2301] alone. However, more recent studies on glatiramer acetate, in a less active population, in a total of over 3000 patients, suggest that the efficacy of glatiramer acetate on ARR and other endpoints is similar to that of IFN beta-1a SC (REGARDS) or IFN beta-1b SC (BEYOND). The observed relapse rates on glatiramer acetate were 0.29 in REGARDS and 0.34 in BEYOND, compared to 0.3 (INF beta-1a SC) and 0.33 (INF beta-1b SC), respectively. In Study FTY720D2302, a double-blind 1-year study, fingolimod 0.5 mg was directly compared to IFN beta-1a IM. The ARR (0.21) in patients treated with fingolimod 0.5 mg was reduced by 52% compared to the ARR (0.43) in patients treated with INF beta-1a IM. Overall, the reduction in ARR in patients treated with fingolimod 0.5 mg compared to those treated with glatiramer acetate can be expected in the range of 30% to 50%. For the purpose of this study, a 35% lower ARR in patients treated with fingolimod 0.5 mg (ARR=0.195) compared to those treated with glatiramer acetate (ARR=0.30) is assumed.

Fingolimod 0.25 mg has never been tested in a clinical trial. Based on PK/PD modeling a 14% higher ARR in patients on fingolimod 0.25 mg compared to those on fingolimod 0.5 mg is anticipated. However, the uncertainty of this estimate is substantial. A 95% confidence interval from the PK/PD model for the ARR in patients on fingolimod 0.25 mg ranges from ARR=0.183 to ARR=0.303. For the purpose of this study, an ARR of 0.225 is assumed in patients treated with fingolimod 0.25 mg, which corresponds to an increase in ARR of approximately 15% compared to fingolimod 0.5 mg and a relative treatment effect of 25% compared to glatiramer acetate.

#### 4.2.1.2 Supportive analyses of primary endpoint

Supportive analyses will be performed for the primary endpoint as well as some variations of the primary endpoint (for sensitivity analysis purpose). All primary and supportive analyses are summarized below and in Table 3-1. The full analysis set will be used unless otherwise specified.

**Table 3-1 Primary efficacy endpoint, supportive endpoints & analyses**

Endpoint	Analysis method	
Aggregate ARR (confirmed relapses only) (i.e., the primary efficacy variable)	Negative binomial model (treatment as main effect, number of relapses in previous year, and baseline EDSS, baseline Gd + T1 lesion count as covariates) (region will be considered)	Primary analysis
	Negative binomial model (treatment)	Supportive analysis 1
	Negative binomial model (treatment)	Supportive analysis 2

	as main effect, number of relapses in previous year, and baseline EDSS, baseline Gd + T1 lesion count as covariates) (region will be considered) on the full analysis set excluding subjects from site # [REDACTED]	
	Negative binomial model (treatment as main effect, number of relapses in previous year, and baseline EDSS, baseline Gd + T1 lesion count as covariates) (region will be considered) on the per-protocol set	Supportive analysis 3
Aggregate ARR (confirmed and unconfirmed relapses)	Negative binomial model (treatment as main effect, number of relapses in previous year, and baseline EDSS, baseline Gd + T1 lesion count as covariates) (region will be considered)	Supportive analysis 4
Subject-level ARR (with imputation; confirmed relapses only)	Rank ANCOVA (treatment as main effect, number of relapses in previous year, and baseline EDSS, baseline Gd + T1 lesion count as covariates) (region will be considered)	Supportive analysis 5
Aggregate ARR (confirmed relapses only and time in study truncated at 367 days)	Negative binomial model (treatment as main effect, number of relapses in previous year, and baseline EDSS, baseline Gd + T1 lesion count as covariates) (region will be considered)	Supportive analysis 6
Aggregate ARR (confirmed relapses derived by EDSS)	Negative binomial model (treatment as main effect, number of relapses in previous year, and baseline EDSS, baseline Gd + T1 lesion count as covariates) (region will be considered)	Supportive analysis 7

- 1) For the primary endpoint, a supportive analysis will be performed using a negative binomial regression model adjusted for treatment group only.
- 2) For the primary endpoint, a supportive analysis will be performed using the same negative binomial model as it is in the primary analysis but based on the full analysis set excluding all subjects from the site # [REDACTED]. Subjects in site # [REDACTED] were mismanaged by the primary investigator at this site.
- 3) For the primary endpoint, a supportive analysis will be performed using the same negative binomial model as it is in the primary analysis but on the per-protocol set. Relapses that start after

study drug discontinuation will be excluded and log (time on study drug in years) rather than log (time in study in years) will be used as the offset variable in the negative binomial model with log link. The number of days on study drug is calculated as (the last study drug dose date – 1<sup>st</sup> study drug dose date +1) which will be used. This is an efficacy analysis of on-treatment data for subjects without major protocol deviations.

4) Analyses described in the primary analysis section will be repeated for the ARR counting all relapses (including both confirmed and unconfirmed relapses).

5) For the subject level ARR with imputation, a supportive analysis will be performed using Rank Analysis of Covariance (ANCOVA) with treatment group as main effect, number of relapses in the previous year before study enrollment, baseline EDSS, and baseline Gd + lesion count as covariates [Stokes (2000)]. An adjustment for country or region will be considered. For subjects who do not complete the treatment phase, relapses will be imputed (see Handling dropouts in ARR calculations Section). Additionally for subject-level ARR without imputation, summaries by treatment group will be presented.

6) Analyses described in the primary analysis sections will be repeated for the confirmed relapses but the calculation for time in study is different. For this supportive analysis the time in study is the number of days from the first study drug dose date to the study completion visit date or the last study drug dose date whichever occurs later. If more than 367 (=12\*30+7) days is obtained, then it will be cutoff to be 367 days which corresponds to 12 months plus a 7-day window. If the study completion visit date is not available, the latest study visit date will be used instead. The time in study (years) will be calculated as the number of days in study/365.25. The confirmed relapses within this time period will be counted for each subject and will be used in the ARR computation.

7) Analyses described in the primary analysis sections will be repeated for the confirmed relapses where confirmation is derived based on EDSS/FS scores, not on the MS Relapses CRF. For each relapse reported on the MS Relapses CRF the confirmation status is derived as the following:

- If no EDSS is performed between start and end dates of the relapse, then the relapse is “unconfirmed” ;
- If an EDSS is performed between start and end dates of the relapse, then compare the EDSS/FS scores to the last available EDSS excluding EDSS recorded during the previous MS relapses. If there is an increase of 0.5 in EDSS or one point on two different FS or 2 points on one FS (all FS except bowel/bladder and Cerebral), the relapse is “confirmed”; otherwise it is “unconfirmed”.

#### 4.2.1.3 Handling dropouts in ARR calculations

1) Negative binomial regression model for aggregate ARR on the full analysis set: For subjects who discontinue the study prematurely, the number of relapses recorded in the study will be used irrespective of whether the subject is on or off study drug and no imputation will be applied.

2) Negative binomial regression model for aggregate ARR on the per-protocol set: For subjects who discontinued the study drug prematurely, the number of relapses that start prior to or on the same day as the last dose date will be used and no imputation will be applied.

3) Rank ANCOVA model for subject-level ARR with imputation on the full analysis set:

For subjects who discontinue the study prematurely, the number of relapses per month for the months after discontinuation will be imputed by the corresponding monthly mean number of relapses of subjects across all treatment groups who has data of either 0 (no relapse) or 1 (one relapse) in that month. That is, for each month with completely missing data, the imputed number of relapses for the month is the average number of relapses for the corresponding month based on the available data from all subjects.

Subjects may have data from part of the month during which they drop out. The number of relapses for that month will be the observed number of confirmed relapse (either 0 or 1) during that partial month. That is, if a subject drops out with one relapse during the dropout month, the subject will be counted as having one relapse for the entire month. If no relapse is observed in the dropout month, the subject is assumed to have no relapse for the entire month. This is accurate because a subject could only have at most one confirmed relapse in each month due to the definition of confirmed relapse in the protocol (Appearance of a new neurological abnormality or worsening of previously stable or improving pre-existing neurological abnormality, separated by at least 30 days from onset of a preceding clinical demyelinating event).

Only the relapse start date will be used to determine the month in which a relapse will be counted when calculating the number of relapses per month.

The ARR of these subjects will be calculated based on the observed number of relapses before dropping out and the imputed number of relapses after dropping out. For example, suppose that the average number of relapses for each month based on all subjects contributing data to the month is as shown in Table 3-2.

**Table 3-2 Example of imputation**

Month	Average # of relapses
1	0.20
2	0.19
3	0.18
4	0.17
5	0.16
6	0.15
7	0.14

8	0.13
9	0.11
10	0.11
11	0.11
12	0.11

If a subject discontinues the study on Day 130 (i.e., after 4 months and 10 days) after having one confirmed relapse, his number of relapses from Month 6 on will be imputed from the above table. His total number of relapses will be the sum of 1, 0.15, 0.14, 0.13, 0.11...0.11, which equals to 1.86. Note that the imputed portion of the number of relapses will be the same for subjects who drop out in the same month.

The imputation will only apply to the dropout subjects up to 360 days (i.e., 12 months). After the imputation, the number of days in study for all dropout subjects is 360 days. For those who complete the study, the number of days in study will remain the same as defined in Time in study for aggregate ARR [Section 4.1.3](#).

4) Summary of subject-level ARR without imputation on the full analysis set: For subjects who discontinue the study prematurely, the number of relapses recorded in the study will be used and no imputation will be applied.

#### 4.2.1.4 Missing MS relapse dates

The following obvious corrections will be implemented by programming if the relapse partial start/or end date occurs where only day is missing.

For an incomplete relapse start date where day is missing but month and year are available, the day will be entered as the first day of the month. In case the relapse partial start date is in the same month as the first study drug dose date, the relapse start date will be entered as the first study drug dose date.

For an incomplete relapse end date where day is missing but month and year are available, the day will be entered as the last day of the month. In case the relapse partial end date is in the same month as the last visit date, the relapse end date will be entered as the last visit date. For a relapse is still ongoing at the last visit the relapse end date will be entered as the last visit date.

#### 4.2.2 Analyses of secondary [REDACTED] variables

Secondary efficacy variables will be analyzed on the full analysis set unless otherwise specified. For most analyses, all three pair-wise treatment comparisons (fingolimod 0.5 mg vs. glatiramer acetate, 0.25 mg vs. glatiramer acetate, and 0.5 mg vs. 0.25 mg) will be provided even though no formal hypothesis will be tested between the 2 fingolimod doses.

#### 4.2.2.1 Magnetic resonance imaging (MRI)

The analyses on MRI parameters will be performed on the full analysis set in the subset of subjects who have MRI scans done during the study. For all MRI data, no visit mapping will be performed for any visit and all analyses will be performed according to the visit recorded in the MRI data set.

The number and volume of Gd + lesion data obtained less than 30 days after use of steroid for treatment of relapses will be excluded from the analysis.

Analyses of MRI variables are summarized in Table 3-3 below and details will be provided in the following section.

**Table 3-3 MRI endpoints & analyses**

Endpoint	Analysis
Brain atrophy - Percent brain volume change (PBVC) from baseline at Month 12/end of treatment	Rank ANCOVA (treatment, age, baseline normalized brain volume (NBV), and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
	ANCOVA (treatment, age, baseline normalized brain volume, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
	ANCOVA (treatment, age, baseline Gd + T1 lesion count, and baseline T2 lesion volume (in cc). (region will be considered)
New/newly enlarged T2 lesion count (compared with baseline MRI scan) at Month 12/end of treatment	Negative binomial (treatment, age, baseline T2 lesion count, baseline Gd + T1 lesion count, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
	Rank ANOVA (treatment, region)
Proportion of subjects free of new/newly-enlarged T2 lesions compared to baseline at Month 12/end of treatment	Logistic regression (treatment, age, baseline T2 lesion count, baseline Gd + T1 lesion count, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
Change (and % change) from Baseline in T2 lesion volume at Month 12/end of	Rank ANCOVA (treatment, age, baseline T2 lesion volume, and the number of relapses



treatment	experienced in the previous year before study enrollment) (region will be considered)
Gd + T1 lesion count at Month 12/end of treatment	Negative binomial (treatment, age, baseline T2 lesion count, baseline Gd + T1 lesion count, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
Proportion of subjects free of Gd + T1 lesions at baseline, Month 12/end of treatment	Logistic regression (treatment, age, baseline Gd + T1 lesion count, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
Gd + T1 lesion volume at Month 12/end of treatment	Rank ANCOVA (treatment, age, baseline Gd + T1 lesion volume, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
Proportion of subjects free of new T1 lesions compared to baseline at Month 12/end of treatment	Logistic regression (treatment, age, baseline T1 lesion count, baseline Gd + T1 lesion count, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
Change and % change from baseline in T1 lesion volume at Month 12/end of treatment	Rank ANCOVA (treatment, age, baseline T1 lesion volume, and the number of relapses experienced in the previous year before study enrollment) (region will be considered)
Percent brain volume change (PBVC) from baseline at Month 12/end of treatment  New/newly enlarged T2 lesion count at Month 12/end of study  Gd + T1 lesion count at Month 12/end of treatment  Proportion of subjects free of new/newly enlarged T2 lesions and proportion of subjects free of Td + T1 lesions at Month 12/end of treatment	<b>Sensitivity analyses</b> to repeat the same analyses described above, but exclude baseline assessments performed > 45 days before randomization and end of treatment assessments performed > 45 days after the last dose of study drug.
Annualized rate of brain atrophy (ARBA) at Month 12/end of treatment	<b>Sensitivity analyses</b> to adjust for study duration and MRI covariates: Rank ANCOVA (treatment, age, baseline NBV, baseline Gd + T1 lesion count, and baseline T2

	lesion volume, and the number of relapses experienced in the previous year before study enrollment) (region was considered)  ANCOVA (treatment, age, baseline NBV, baseline Gd + T1 lesion count, and baseline T2 lesion volume, and the number of relapses experienced in the previous year before study enrollment (region was considered)
--	--

The number of new or newly enlarging T2 lesions at Month 12/end of treatment and number of Gd-enhancing T1 lesions at Month 12/end of treatment will be analyzed using a negative binomial regression model with log link, using treatment as main effect, age, baseline T2 lesion count, baseline Gd-enhancing T1 lesion count, and the number of relapses experienced in the previous year before study enrollment as covariates. An adjustment for region will be considered in the model. For the analysis of T2 lesions, an offset variable will be used in the negative binomial analysis to adjust for the time in years (since start of study drug).

The proportion type variables will be analyzed using a logistic regression model with treatment as main effect, age, corresponding baseline value, baseline Gd-enhancing T1 lesion count, and the number of relapses experienced in the previous year before study enrollment as covariates. An adjustment for country or region will be considered in the model.

The continuous variables (change, percent change, and total volume) will be analyzed using a rank ANCOVA model with treatment as main effect, adjusted for age, corresponding baseline value, and the number of relapses experienced in the previous year before study enrollment as covariates. An adjustment for country or region will be considered. ANCOVA analysis for the percent change in brain volume with the above variables as covariates will be considered as well. In addition, an ANCOVA model with treatment, baseline Gd + T1 lesion count and the T2 volume (in cc) will be performed.

**4.2.2.2 Relapses**

All relapse-related analyses use confirmed relapses only, unless otherwise specified. [REDACTED]

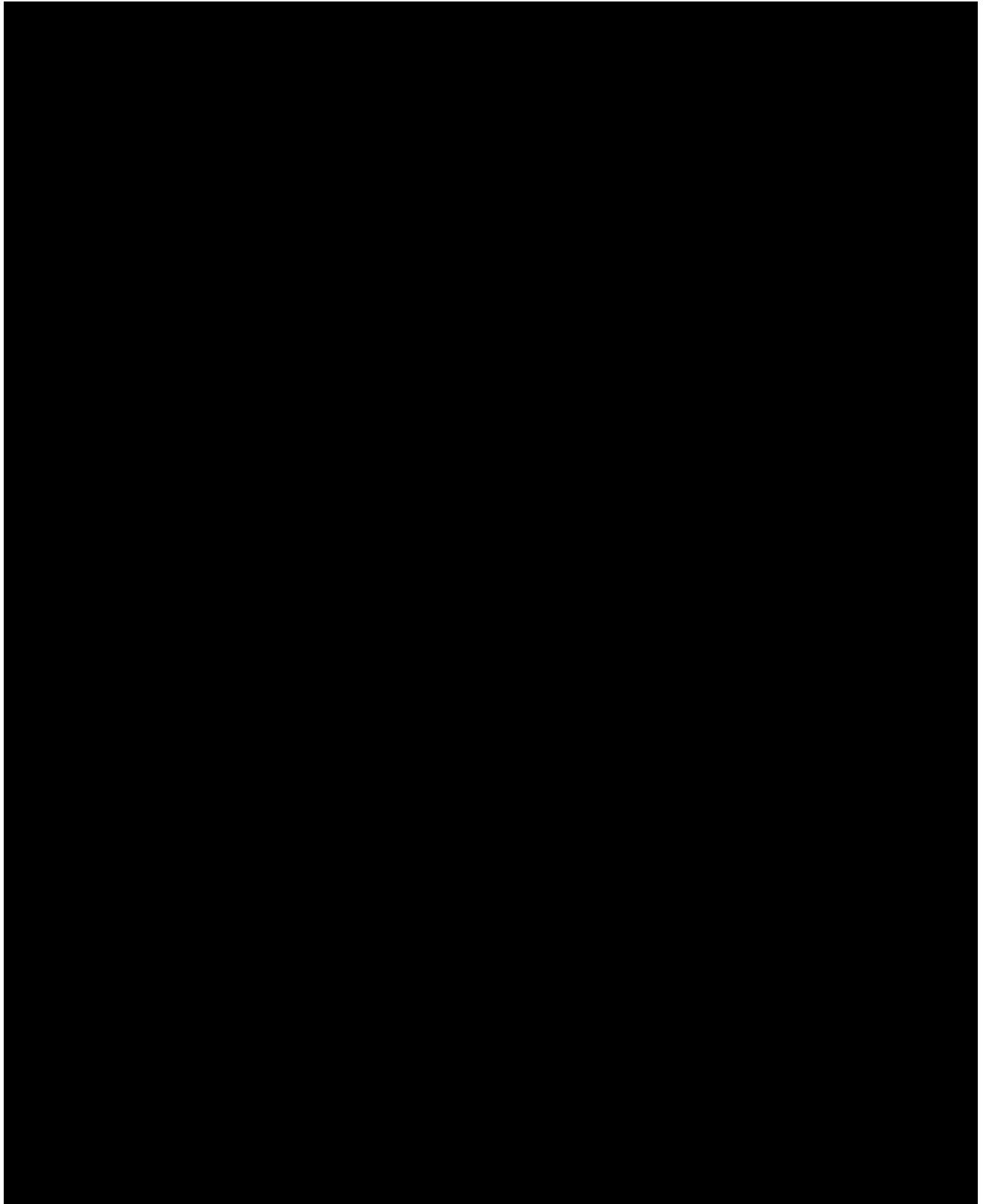
[REDACTED]

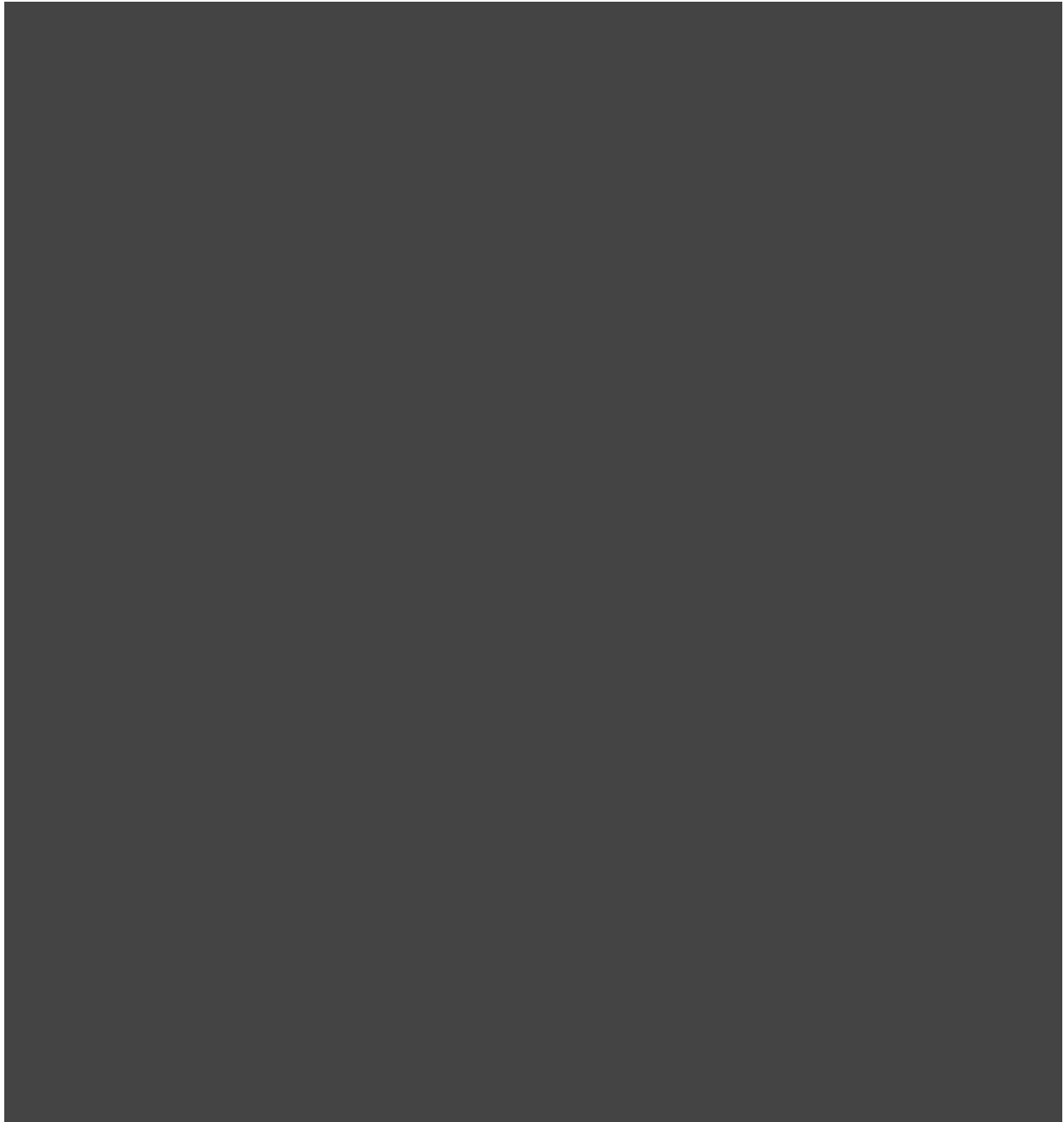
[REDACTED]

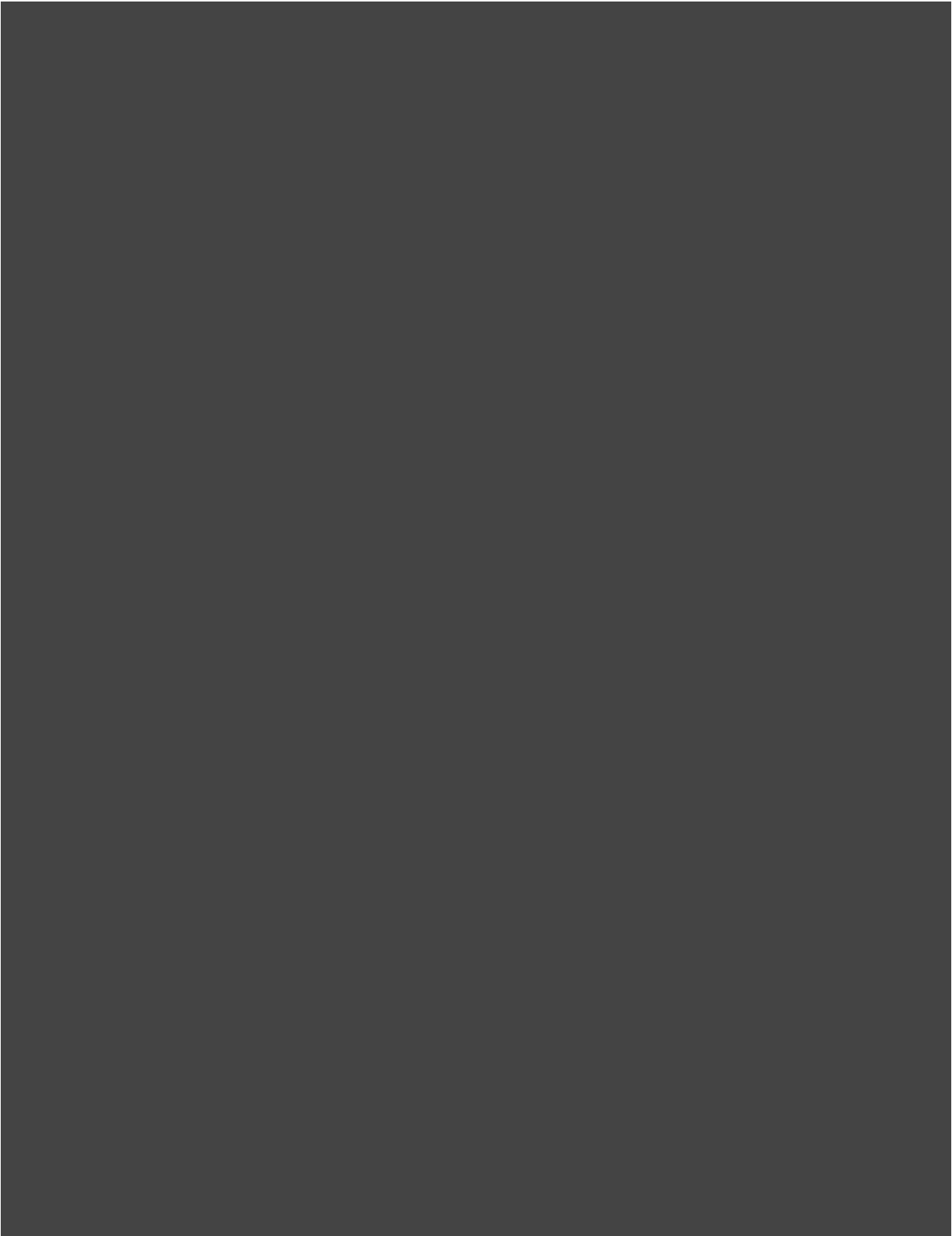
[REDACTED]

[REDACTED]

[REDACTED]



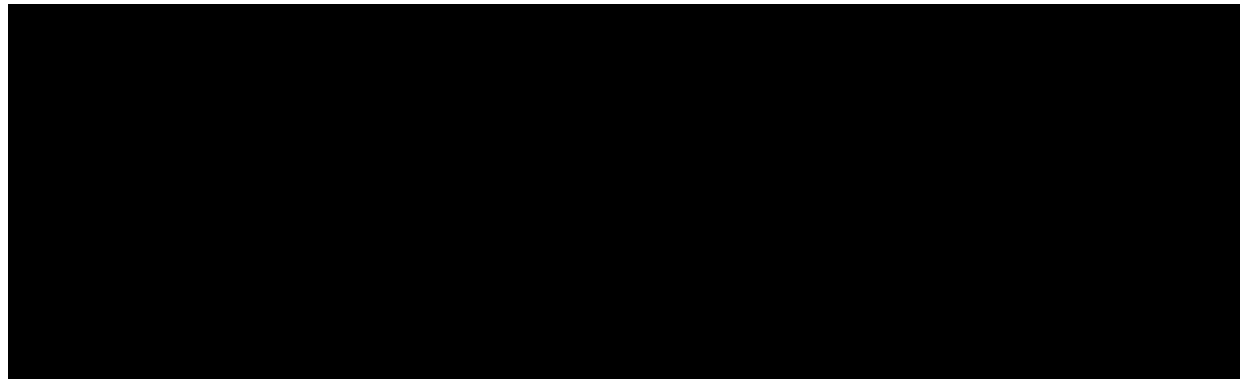
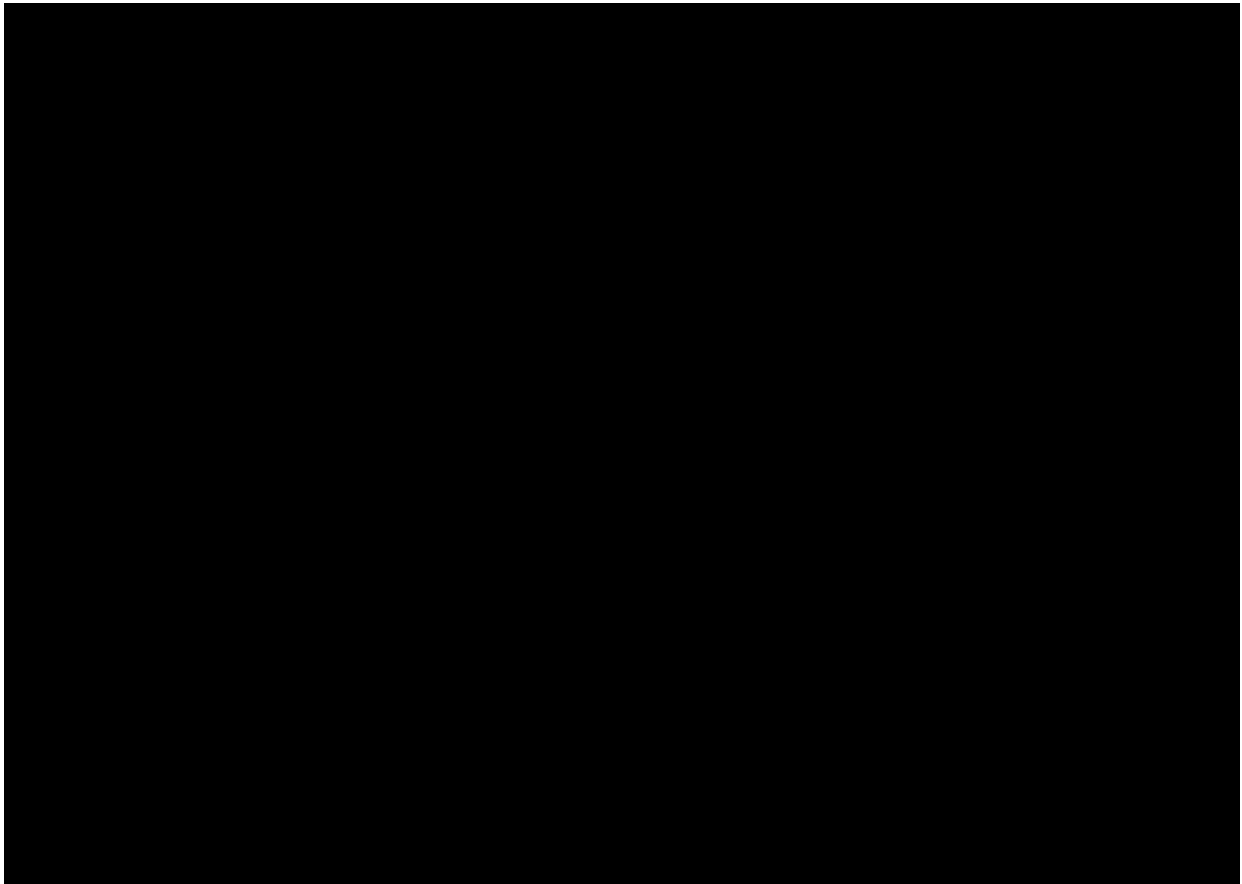


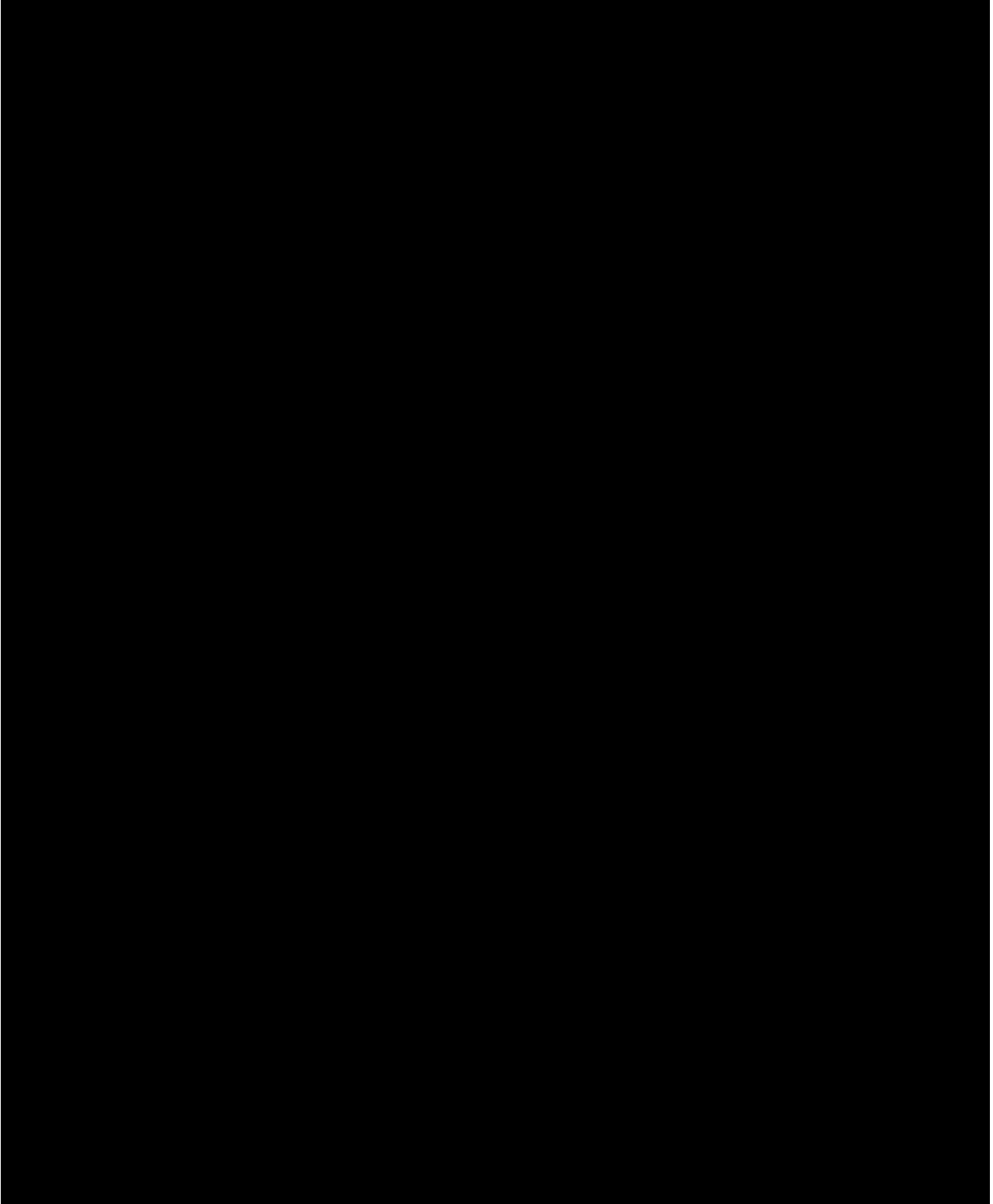




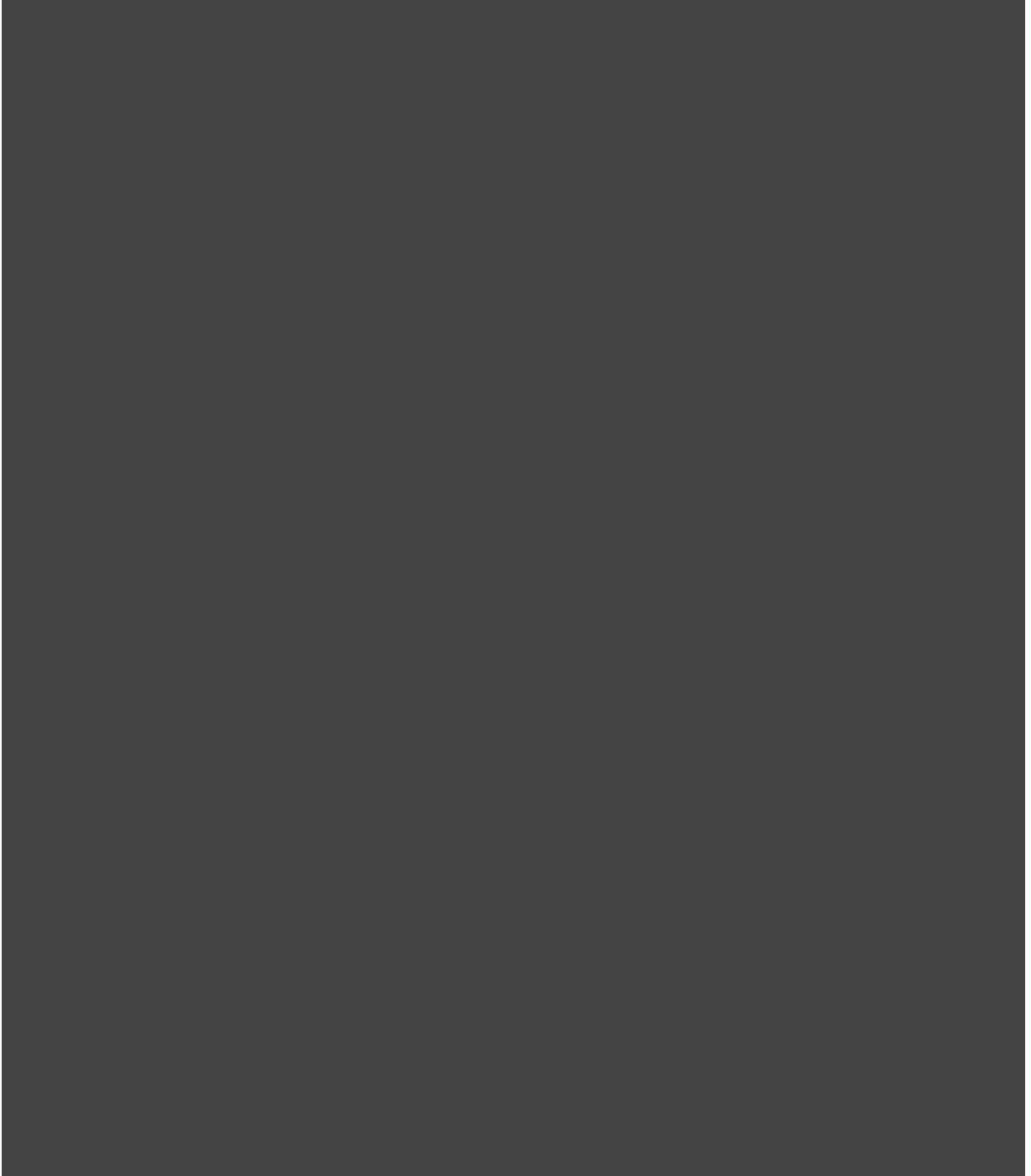
#### **4.2.2.5 EDSS Scores**

Summary statistics will be presented on the FAS for the observed EDSS scores by visit and treatment. The change from baseline to end of study in EDSS score will be summarized on the FAS as well.





## **7 Health-related Quality of Life (QoL)**







### 7.3 TSQM (Version 1.4)

Treatment Satisfaction Questionnaire for Medication (TSQM) was developed and validated as a general measure for treatment satisfaction. The TSQM version 1.4 modified will be used, which contains 14 items assessing the following 4 domains: Global Satisfaction, Effectiveness, Side Effects, and Convenience. This patient reported outcome (PRO) must be completed prior to any other study visit assessments at the specified visits according to Table 6-1 assessment schedule in study protocol.

Scores for the TSQM version 1.4 will be generated using the authors' scoring algorithm. Missing data will be treated in accordance with the authors' guidelines ([Atkinson et al 2004](#)) as follows:

#### EFFECTIVENESS:

$$\text{Score} = \{[(\text{sum of items 1-3}) - 3] / 18\} * 100$$

$$\text{If only one item is missing, Score} = \{[(\text{sum of available items in 1-3}) - 2] / 12\} * 100$$

#### SIDE EFFECTS:

If Question 4 is answered "No" then Score = 100.

Otherwise:

$$\text{Score} = \{[(\text{sum of items 5-8}) - 4] / 16\} * 100$$

$$\text{If only one item is missing, Score} = \{[(\text{sum of available items in 5-8}) - 3] / 12\} * 100$$

**CONVENIENCE:**

$$\text{Score} = \{[(\text{sum of items 9-11}) - 3] / 18\} * 100$$

$$\text{If only one item is missing, Score} = \{[(\text{sum of available items in 9-11}) - 2] / 12\} * 100$$

**GLOBAL SATISFACTION:**

$$\text{Score} = \{[(\text{sum of items 12-14}) - 3] / 14\} * 100$$

If only one item is missing:

$$\text{Item 12 or 13 missing: Score} = \{[(\text{sum of available items in 12-14}) - 2] / 10\} * 100$$

$$\text{Item 14 missing: Score} = \{[(\text{sum of available items in 12-14}) - 2] / 8\} * 100$$

The higher summary scores indicate better satisfaction with study drug.

Scores for each scale at scheduled visits will be summarized by treatment group with n, mean SD, minimum, Q1, median, Q3, and maximum. The change from baseline score will be compared within treatment group by Wilcoxon signed-rank test (also known simply as the Wilcoxon sign test).

## 8 Safety evaluation

Safety assessments include adverse events, infections, bradycardia events, laboratory tests, vital signs, ECG, pulmonary function tests (PFTs), ophthalmic examinations, and dermatology assessment.

All safety data will be summarized on the safety set. Some safety data (AEs, vital signs, labs, PFTs, and ophthalmic evaluations) will also be summarized on the follow-up set.

Unless otherwise specified, safety assessments more than 45 days after the study drug discontinuation will not be included in the summaries on the safety set but will be considered in the summaries on the follow-up set. The 45 days rule should be applied as follows:

45 days after the last study drug dose date will be the cutoff for safety data to be included in the summaries.

Adverse events starting/ending before the first dose date will be excluded from the adverse event summaries i.e. only treatment emergent AEs will be summarized by treatment.

Details of the safety data summaries will be provided in the sections below. All safety data will also be presented in the listings.

## 8.1 Adverse events

### 8.1.1 Adverse event date imputation

To ensure that all reported AEs are summarized, a conservative approach will be taken to impute the AE date when a partial date is available. AE date imputation is based on a comparison of the partial AE start date to the treatment start date (Table 4-1).

1. If the AE start date year value is missing, the AE start date will be imputed as the treatment start date.
2. If the AE start date year value is less than the treatment start date year value, the AE must have started before treatment. The following imputation rules will apply.
  - a. If the AE month is missing, the AE start date will be imputed as the mid year date (01JulYYYY).
  - b. If the AE month is not missing, the AE start date will be imputed as the mid month date (15MONYYYY).
3. If the AE start date year value is greater than the treatment start date year value, the AE must have started after treatment. The following imputation rules will apply.
  - a. If the AE month is missing, the AE start date will be imputed as the year start date (01JanYYYY).
  - b. If the AE month is not missing, the AE start date will be imputed as the month start date (01MONYYYY).
4. If the AE start date year value is equal to the treatment start date year value, the month values will be compared as follows.
  - a. If the AE month is missing or the AE month is equal to the treatment start month, the AE start date will be imputed as the date one day after treatment start date.
  - b. If the AE month is less than the treatment start month, the AE start date will be imputed as the mid month date (15MONYYYY).
  - c. If the AE month is greater than the treatment start month, the AE start date will be imputed as the start month date (01MONYYYY).

**Table 4-1 AE date imputation**

MON MISSING	MON < CFM	MON = CFM	MON > CFM
(F)=TRTSTD	(F)=TRTSTD	(F)=TRTSTD	(F)=TRTSTD
Uncertain	Uncertain	Uncertain	Uncertain
(D) = 01JULYYYY	(C) = 15MONYYYY	(C) = 15MONYYYY	(C) = 15MONYYYY
Before Treatment Start	Before Treatment Start	Before Treatment Start	Before Treatment Start
(B)= TRTSTD+1 Uncertain	(C) = 15MONYYYY Before Treatment Start	(A)= TRTSTD+1 Uncertain	(A)= 01MONYYYY After Treatment Start
(E)= 01JANYYYY After Treatment Start	(A)= 01MONYYYY After Treatment Start	(A)= 01MONYYYY After Treatment Start	(A)= 01MONYYYY After Treatment Start
Before Treatment Start	Partial indicates date prior to Treatment Start Date		
After Treatment Start	Partial indicates date after Treatment Start Date		
Uncertain	Partial insufficient to determine relationship to Treatment Start Date		
<b>LEGEND:</b>			
(A)	MAX(01MONYYYY,TRTSTD+1)		
(B)	TRTSTD+1		
(C)	15MONYYYY		
(D)	01JULYYYY		
(E)	01JANYYYY		
(F)	TRTSTD		

**8.1.2 Adverse events**

In this section, the term adverse event (AE) will refer to any entry on the adverse events CRFs. All AEs including infections will be coded using the MedDRA dictionary. The MedDRA version 20.0 or higher will be used depending when the study is completed.

Adverse events will be summarized by presenting, for each treatment group, the number and percentage of subjects having any AE by primary system organ class and preferred term. SAEs, death, drug-related AEs, the AEs leading to premature discontinuation of study drug, AEs requiring study-drug dose adjustment or interruption and AEs requiring additional therapy will be presented in a similar format as AEs. Notable events will include death, nonfatal SAEs (including infections), and AEs (including infections) leading to study drug discontinuation. AEs that fulfill the risk search terms (defined in the case retrieval sheet) with RMP risk will be summarized by risk name and lower level terms.

SAEs include death and non-fatal SAEs. All SAEs starting after the first dose date, including those starting more than 45 days after study drug discontinuation, will be included in the summaries on the safety set.

Additionally, occurrence rates of AEs, i.e. occurrence of AEs per 100 patient year by primary organ class and preferred term will be summarized. The occurrence rate of an event is calculated as the number (n) of all such events that patients experience divided by the patient years (Ny), where patient years are calculated as the sum of the number of days on study drug of all patients divided by 365.25. The occurrence of AEs per 100 patient years, which is calculated as  $n/Ny*100$ , will be presented.

AE summaries on the follow-up set will be conducted for the following time periods defined by days relative to study drug discontinuation: days 1 to 45 after study drug discontinuation, day 46 or later after study drug discontinuation.

#### **8.1.2.1 Occurrence of SAEs and NSAEs (non-serious AEs)**

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on on-treatment adverse events which are not serious adverse events with an incidence greater than X% and on on-treatment serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is  $\leq 1$  day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is  $> 1$  day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non SAE has to be checked in a block e.g., among AE's in a  $\leq 1$  day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

#### **8.1.3 Data summaries for AEs**

Summaries on the safety set by treatment group include:

- Incidence of AEs by SOC and PT
- Incidence of AEs that fulfill the risk search terms with RMP risk by risk name and level terms
- Incidence of AEs by SOC and PT and severity
- Incidence of deaths by SOC and PT
- Incidence of SAE by SOC and PT
- Incidence of SAEs that fulfill the risk search terms with RMP risk by risk name and level terms
- Incidence of AEs leading to study drug discontinuation by SOC and PT
- Incidence of AEs requiring study-drug dose adjustment or interruption by SOC and PT
- Incidence of AEs requiring additional therapy by SOC and PT
- Incidence of most frequent ( $\geq 1\%$  for any group) AEs by PT
- Incidence of AEs related to study drug by SOC and PT
- Occurrence of AEs per 100 patient-years by SOC and PT
- Occurrence of SAEs by SOC and PT
- Occurrence of NSAEs by SOC and PT

Summaries on the follow-up set by treatment group include:

- Incidence of AEs by SOC and PT (days 1 – 45 after study drug discontinuation, and more than 45 days after study drug discontinuation)

## **8.2 Laboratory data**

### **8.2.1 SI Units**

Data summaries will be provided in SI units. The standard conversion tool (UNITCONV) will be used to convert all values in different units to SI units.

### **8.2.2 Clinically notable criteria**

Clinically notable criteria are summarized in Table 4-2.

All applicable post-baseline laboratory results will be checked against the respective notable criteria. For a particular notable criterion, a subject will be counted in the notable abnormal category as long as one of the results meets the criterion. Note that a subject can be counted in both low and high abnormal categories.

For some laboratory parameters (as listed in Table 4-3), subjects may have a character value of “<x.x” or “<x” instead of the actual numeric value which are below the lower limit of detections (LLOD). The 50% of the LLOD rule will be applied to transform a character value into a numerical value as specified in Table 4-3. This rule will be applied to all applicable lab values prior to checking against the notable criteria as well as computation of any statistics.

For the laboratory parameters listed in Table 4-4, results will be summarized based on the abnormal criteria which are defined by pre-specified level of abnormality in terms of either LLN or ULN as listed in Table 4-4.

In addition, the patients meeting modified Hy’s Law will be included in the summary of newly occurring or worsening post-baseline liver function test abnormalities. The modified Hy’s Law biochemical criteria is defined as at least one of (ALT or AST > 3 x ULN) and (total bilirubin  $\geq$  2 x ULN) and (alkaline phosphatase < 2 x ULN) at a same visit.

**Table 4-2 Criteria for clinically notable abnormalities**

Laboratory Variable	Standard Units	SI Units
<b>LIVER FUNCTION AND RELATED VARIABLES</b>		
SGOT (AST)	> 82 U/L	> 82 U/L
SGPT (ALT)	> 90 U/L	> 90 U/L
Gamma glutamyl transferase (GGT) (U/L)	> 130 U/L	> 130 U/L
Bilirubin (total)	$\geq$ 2.0 mg/dL	$\geq$ 34.2 $\mu$ mol/L
Alkaline phosphatase, serum	> 280 U/L	> 280 U/L
<b>RENAL FUNCTION / METABOLIC AND ELECTROLYTE VARIABLES</b>		
Glucose	$\geq$ 200 mg/dL	$\geq$ 11.11 mmol/L
Creatinine	$\geq$ 2.0 mg/dL	$\geq$ 176 $\mu$ mol/L
Amylase	$\geq$ 300 U/L	$\geq$ 300 U/L
Cholesterol (total)	$\geq$ 240 mg/dL	$\geq$ 6.21 mmol/L
Triglycerides	$\geq$ 300 mg/dL	$\geq$ 3.39 mmol/L
BUN	$\leq$ 2 mg/dL $\geq$ 30 mg/dL	$\leq$ 0.7 mmol/L $\geq$ 10.7 mmol/L
Sodium	< 125 mEq/L > 154 mEq/L	< 125 mmol/L > 154 mmol/L
Chloride	$\leq$ 85 mEq/L $\geq$ 119 mEq/L	$\leq$ 85 mmol/L $\geq$ 119 mmol/L
Potassium	$\leq$ 3.0 mEq/L $\geq$ 6.0 mEq/L	$\leq$ 3.0 mmol/L $\geq$ 6.0 mmol/L
Magnesium	$\leq$ 1.0 mg/dL	$\leq$ 0.40 mmol/L

	$\geq 3.00$ mg/dL	$\geq 1.23$ mmol/L
Calcium	$\leq 7.5$ mg/dL $\geq 11.6$ mg/dL	$\leq 1.87$ mmol/L $\geq 2.89$ mmol/L
Phosphate	$\leq 2.0$ mg/dL $\geq 5.3$ mg/dL	$\leq 0.65$ mmol/L $\geq 1.71$ mmol/L
<b>HEMATOLOGY VARIABLES</b>		
Haemoglobin	$\leq 10.0$ g/dL	$\leq 100$ g/L
Platelets (Thrombocytes)  <labeled as Platelet count (direct) in database>	$\leq 100$ k/mm <sup>3</sup> $\geq 600$ k/mm <sup>3</sup>	$\leq 100 \times 10^9$ /L $\geq 600 \times 10^9$ /L
Leukocytes (WBCs)  <labeled as WBC (total) in database>	$\leq 2.0$ k/mm <sup>3</sup> $\geq 15$ k/mm <sup>3</sup>	$\leq 2.0 \times 10^9$ /L $\geq 15 \times 10^9$ /L
<b>HEMATOLOGY VARIABLES: DIFFERENTIAL</b>		
Granulocytes (Poly, Neutrophils)  <labeled as Neutrophils (Seg. + Bands)>	$\leq 1,000$ /mm <sup>3</sup> $\geq 12000$ /mm <sup>3</sup>	$\leq 1 \times 10^9$ /L $\geq 12 \times 10^9$ /L
Lymphocytes	$< 200$ /mm <sup>3</sup> $\geq 8000$ /mm <sup>3</sup>	$< 0.2 \times 10^9$ /L $\geq 8 \times 10^9$ /L
Red blood cells (RBCs)	$< 3,300,000$ /mm <sup>3</sup> $> 6,800,000$ /mm <sup>3</sup>	$< 3.3 \times 10^{12}$ /L $> 6.8 \times 10^{12}$ /L

\*"Americas" refers to North (Canada, US), Central, and South America. It is applicable only to Canada and US in this study.

**Table 4-3 The lower limit of detections of laboratory parameters**

Laboratory test	Character value	50% Lower limit of detection (LLOD)
Absolute Monocytes	ABSMON < 0.1	ABSMON = $0.05 \times 10^9$ /L
Absolute Lymphocytes	ABSLYM < 0.1	ABSLYM = $0.05 \times 10^9$ /L
Alkaline Phosphatase	ALKPHS < 3	ALKPHS = 1.5 U/L



Amylase	AMY < 3	AMY= 1.5 U/L
SGOT (AST)	SGOT < 5	SGOT = 2.5 U/L
SGPT (ALT)	SGPT < 4	SGPT = 2 U/L
Gamma Glutamyltransferase (GGT)	GGT < 5	GGT = 2.5 U/L
Bilirubin (total)	TBIL < 2	TBIL = 1 µmol/L
Bilirubin (direct/conjugated)	DBIL < 3	DBIL = 1.5 µmol/L
HCG	HCG < 2	HCG = 1 IU/L
Herpes Simplex Virus 2 IgG	HSVIGG4 < 0.5	HSVIGG4= 0.25
IgG antibody to varicella zoster virus	IGGVZV2 < 10	IGGVZV2 =5 U/L
Glucose	GLUC < 1 mmol/L	GLUC= 0.5 mmol/L
Urine Protein Dipstick test	UPROTST < 0.1	UPROTST =0.05

**Table 4-4 Abnormal criteria**

Laboratory Variable (unit)	Criteria
Albumin (g/L)	<LLN
Bilirubin (direct/conjugated) (umol/L)	>ULN >2xULN
Gamma Glutamyltransferase (GGT) (U/L)	>ULN >3xULN
Glucose (mmol/L)	>ULN
Cholesterol (HDL) (mmol/L)	>ULN
Cholesterol (LDL) (mmol/L)	>ULN
SGOT (AST) (U/L)	>ULN >2xULN >3xULN >5xULN
SGPT (ALT) (U/L)	>ULN

	>2xULN >3xULN >5xULN
Cholesterol (total) (mmol/L)	>ULN

### 8.2.3 Data Summaries for lab tests

All lab parameters (excluding all hematology % values except % for white cell differential values) will be summarized.

Summaries on the safety set by treatment group include:

- Summary statistics of lab results (hematology, biochemistry, and urinalysis) by visit
- Changes from baseline in lab results (hematology, biochemistry, and urinalysis) by visit
- Summary statistics of percent of baseline in hematology results by visit
- Incidence rates of newly occurring or worsening abnormalities based on clinically notable laboratory abnormality criteria (based on criteria in Table 4-2)
- Incidence rates of clinically notable laboratory abnormalities by visit
- Incidence rates of laboratory results meeting abnormal criteria (in Table 4-4) by visit
- Frequency (%) distributions of liver function test results and hematology results in pre-defined categories (as specified in the table shells)
- Shift table (low, normal, or high as defined by the normal ranges) from baseline to post-baseline extreme values in hematology, biochemistry and continuous urinalysis results
- Shift table (negative, +, 2+, 3+, or 4+) from baseline to post-baseline extreme values in urinalysis results
- Frequency (%) distributions of serology results for subjects with serology test

Summaries on the follow-up set by treatment group include:

- Summary statistics of lab parameters (hematology, biochemistry, and urinalysis) by visit
- Change from baseline in lab parameters (hematology, biochemistry, and urinalysis) by visit

## 8.3 Vital signs

### 8.3.1 Vital signs measurements

The vital sign CRF collects systolic blood pressure (SBP) and diastolic blood pressures (DBP), pulse, body temperature, and body weight. The mean arterial blood pressure will be derived from the DBP and SBP using the formula below

- Mean arterial blood pressure =  $[2*DBP + SBP]/3$ .

Before the protocol amendment 1, the CRF collects vital signs only at sitting position at each visit. After the protocol amendment 1, the CRF collects vital signs at supine and standing position at each visit

For all visits (excluding the first or second or restart dose monitoring data) one measurement of pulse and 3 measurements of blood pressure (SBP and DBP) are collected before the protocol amendment 1 and three measurements of pulse and blood pressure (SBP and DBP) are collected.

At first or second or restart dose monitoring, pulse, SBP and DBP are measured 3 times at the scheduled pre-dose time point. The average of available values among three pulse ( or SBP or DBP) measurements at the same position will be calculated.

For post-baseline assessments (excluding the first or second or restart dose monitoring data), the pulse or blood pressure value will be the average of the non-missing values of the 3 measurements, if it is not a single measurement, at the same position. If more than one blood pressure assessment after the average of the 3 measurements at the same position at the same date (scheduled or unscheduled, except for the first dose or second dose or restart dose monitoring data) exists in a particular visit as defined by the visit windows, the blood pressure value will be the average of all assessments at the same position.

The vital sign (pulse, SBP and DBP) during the first dose administration on Day 1 or Day 2 and restart data will not be put into visit windows, but will be summarized separately by the hour in which the measurement was reported.

### 8.3.2 Vital signs notable criteria

Clinical notable criteria in vital signs are summarized in Table 4-5.

**Table 4-5 Criteria for clinically notable vital signs**

Vital Sign Variable	Notable Criteria
Pulse (beats/min)	>120bpm or Increase of $\geq 15$ bpm from baseline or < 50bpm or Decrease of $\geq 15$ bpm from baseline
Systolic BP (mmHg)	$\geq 160$ mm Hg or Increase of $\geq 20$ mm Hg from baseline or

	≤ 90 mm Hg or Decrease of ≥ 20 mm Hg from baseline
Diastolic BP (mmHg)	≥ 100 mmHg or Increase of ≥ 15 mm Hg from baseline or ≤ 50 mmHg or Decrease of ≥ 15 mm Hg from baseline
Temperature (°C)	>38.3 °C/101°F
Body weight (kg)	± ≥ 7% from baseline weight

### 8.3.3 Data summaries for vital signs

Pulse and blood pressure will be reported separately for sitting, supine and standing position. Incidence rates of clinically notable, post-baseline abnormalities and orthostatic hypotension will be provided by treatment. Orthostatic hypotension is defined as a reduction of systolic blood pressure of at least 20 mm Hg or diastolic blood pressure of at least 10 mm Hg from supine to standing position.

The following summaries will exclude hourly vital sign data from the first dose or second dose or restart dose monitoring.

Summaries on the safety set by treatment group include:

- Summary statistics of vital sign parameters by visit
- Changes from baseline in vital sign parameters by visit
- Frequency (%) distributions of highest SBP, DBP, and lowest pulse in pre-defined categories respectively
- Incidence rates of notable vital sign abnormalities based on the notable criteria in Table 4-5
- Incidence rates of orthostatic hypotension by treatment and visit

Summaries on the follow-up set by treatment group include:

- Summary statistics of vital sign parameters by visit
- Change from baseline of vital sign parameters by visit

## 8.4 Electrocardiograms (ECG)

The ECG data includes quantitative variables such as ventricular rate, PQ or PR interval, R-R interval, QRS duration, and uncorrected QT interval as well as categorical variables such as ECG interpretation (normal or abnormal), ECG evaluation type (rhythm, conduction, etc) and ECG findings (prolonged QTc, low voltage, etc).

#### 8.4.1 Corrected QT interval

The uncorrected QT interval will be corrected using the Bazett and Fridericia formulas. The Bazett formula corrects the QT interval by dividing by the square root of the R-R interval (secs). The Fridericia formula corrects the QT interval by dividing by the cubic root of the R-R interval (secs).

#### 8.4.2 Abnormality criteria for corrected QT interval

Abnormality criteria for the corrected QT interval (Bazett and Fridericia) are listed in Table 4-6.

**Table 4-6 Abnormality criteria for corrected QT interval**

	Male Subjects	Female Subjects
1	>450 msec	>470 msec
2	>500 msec	>520 msec
3	30 - 60 msec increase from Baseline	
4	>60 msec increase from Baseline	

#### 8.4.3 ECG findings

ECG findings are associated with the ECG evaluation type and thus will be summarized by the ECG evaluation type.

#### 8.4.4 Data summaries for ECG data

The following summaries will exclude ECG data from the first dose or second dose or restart dose monitoring.

Summaries on the safety set by treatment group include:

- Summary statistics of quantitative ECG parameters by visit
- Changes from baseline in quantitative ECG parameters by visit
- Incidence rates of abnormal corrected QT interval (Bazett and Fridericia) as defined by the abnormality criteria in Table 4-6
- Frequency (%) distribution of ECGs interpretation (normal/abnormal) by visit
- Frequency (%) distribution of ECG findings by ECG evaluation type by visit

Note: For this study the only post-baseline visit is Month 12/EOT (end of treatment). So, all summaries above will be present for Baseline and then Month 12/EOT.

## 8.5 Pulmonary function tests (PFT)

### 8.5.1 Predicted values and units of PFTs

For each subject, the predicted values of the PFT parameters (Forced expiratory volume in 1 second (FEV<sub>1</sub>), Forced vital capacity (FVC), FEV<sub>1</sub>/FVC, and Diffusion capacity of carbon monoxide (D<sub>L</sub>CO)) will be calculated based on the formulas below.

#### Reference equations to calculate the predicted values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC

- Forced expiratory volume in one second (FEV<sub>1</sub>) (Unit: Liter/second):  
male:  $(4.3 \cdot \text{height (m)}) - (0.029 \cdot \text{age (yr)}) - 2.49$  [ $\pm 0.51$ ]  
female:  $(3.95 \cdot \text{height (m)}) - (0.025 \cdot \text{age (yr)}) - 2.6$  [ $\pm 0.38$ ]
- Forced vital capacity (FVC) (Unit: Liter):  
male:  $(5.76 \cdot \text{height (m)}) - (0.026 \cdot \text{age (yr)}) - 4.34$  [ $\pm 0.61$ ]  
female:  $(4.43 \cdot \text{height (m)}) - (0.026 \cdot \text{age (yr)}) - 2.89$  [ $\pm 0.43$ ]
- FEV<sub>1</sub>/FVC (Unit: %):  
male:  $-0.18 \cdot \text{age (yr)} + 87.21$  [ $\pm 7.17$ ]  
female:  $0.19 \cdot \text{age (yr)} + 89.10$  [ $\pm 6.51$ ]

#### Reference equations to obtain predicted values of D<sub>L</sub>CO

- Single breath diffusion capacity (D<sub>L</sub>CO) (Unit: ml/min/mmHg):  
Male:  $(0.3319 \cdot \text{height (cm)}) - (0.1971 \cdot \text{age (yr)}) - 18.006$   
Female:  $(0.2441 \cdot \text{height (cm)}) - (0.1436 \cdot \text{age (yr)}) - 8.20$

In the above calculations, the height of a subject is from the screening visit and the age should be calculated based on the date when he or she performs the PFTs.

For D<sub>L</sub>CO, three units may be collected in the CRFs, which are ml/min/torr, ml/min/mmHg, and mmol/min/kpa. The PFT tables and listings will use the unit ml/min/mmHg. The other two units will be converted to ml/min/mmHg, prior to any calculations, as follows:

- 1 ml/min/torr = 1 ml/min/mmHg;
- 1 mmol/min/kpa = 2.986 ml/min/mmHg.

### 8.5.2 Percent predicted PFTs

The percent predicted PFT value, for each subject, is defined as (absolute PFT value / predicted PFT value) \* 100.

The change from baseline in percent predicted PFT value is defined as (post-baseline percent predicted PFT value - baseline percent predicted PFT value).

The percent of baseline in absolute value of a PFT is calculated as (post-baseline absolute PFT value / baseline absolute PFT value) \* 100.

The percent predicted PFT value will be categorized by pre-defined intervals. Details will be specified in the table shell document (RAP Module 7.1)

Relevant PFT changes are defined as below 80% of baseline and below 60% of baseline in absolute PFT values at 2 consecutive visits.

For non by-visit summaries, visits defined by visit windows will not be used. All available data from scheduled and unscheduled visits will be considered when checking 2 consecutive visits.

### 8.5.3 Data summaries for PFTs

Summaries on the safety set by treatment group include:

- Summary statistics in absolute values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO and in percent predicted values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO by visit
- Changes from baseline in absolute values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO and in percent predicted values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO by visit
- Number (%) of subjects meeting the following criteria as compared to the baseline (separately for FEV<sub>1</sub>, FVC, D<sub>L</sub>CO): below 80% at any visit; below 80% at 2 consecutive visits; below 60% at any visit; and below 60% at 2 consecutive visits
- Shift tables from baseline to the lowest post-baseline values in pre-defined categories of the percent predicted values of FEV<sub>1</sub>, FVC, and D<sub>L</sub>CO
- Number (%) of subjects with relevant PFT changes (for FEV<sub>1</sub>, FVC, and D<sub>L</sub>CO) by baseline percent predicted PFT categories

Summaries on the follow-up set include:

- Summary statistics in absolute values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO and in percent predicted values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO by visit
- Change from baseline in absolute values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO and in percent predicted values of FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC, and D<sub>L</sub>CO by visit

**Note:** Site [REDACTED] had 5 subjects with unrealistic values (>80L) reported for FEV<sub>1</sub>/FVC in the CRF. The site is not responsive. On August 22<sup>nd</sup>, 2018, the study team reviewed the issue and agreed to set them as missing (unknown) in the analysis.

### 8.6 Ophthalmic evaluations

The ophthalmic evaluation includes assessment of visual acuity, assessment of optical coherence tomography (OCT), fluorescein angiography, and assessment of macular edema.

For the assessment of visual acuity, if the decimal score is not available but the Corrected Snellen equivalent is available, the decimal score will be calculated by division (e.g., Corrected Snellen equivalent = 10/20 infers that decimal score = 0.5). All decimal scores will be converted to the LogMAR (log of the minimum angle of resolution) equivalent by taking the negative of the common logarithm (i.e.,  $-\log_{10}$  (decimal acuity score)) (Holladay 1997). Summaries of the visual acuity will be performed on the converted scores (i.e., LogMAR equivalent).

If two or more visual acuity values for the same eye of a subject are available at a post-baseline visit window, the worst of all values will be used in the summaries. Note that the worst value for visual acuity refers to the smallest value in the decimal score (or equivalently refers to the largest value in the logMAR scale).

Visual acuity and change from baseline in visual acuity will be categorized by pre-defined intervals. Details will be specified in the table shell document (RAP Module 7.1).

The central foveal thickness (CFT), macular volume data from OCT and fluorescein angiography data will be displayed in listings.

### **8.6.1 Data summaries for ophthalmic evaluations**

Summaries on the safety set by treatment group include:

- Summary statistics of visual acuity by visit
- Changes from baseline in visual acuity by visit
- Frequency (%) distribution of categorized visual acuity by visit
- Frequency (%) distribution of categorized change from baseline in visual acuity by visit
- Summary of diagnosis of macular edema

### **8.7 Dermatology assessment**

Dermatology examinations will be performed at screening, and Month 12/end of study visits according to a protocol. Results of the dermatology exams are categorical (normal or abnormal with abnormal types) and will be summarized by treatment group for each visit as well as for the unscheduled visit. For unscheduled visit summary, if a subject has 2 or more unscheduled assessments, only the worst result will be summarized.

One shift table from baseline assessment to Month 12/end of treatment assessment will also be presented.



## **8.8 First dose (second dose or restart dose) of fingolimod monitoring**

### **8.8.1 Vital signs**

Hourly vital signs including pulse, SBP and DBP are collected at pre-dose and 6 hours after the first dose, second dose (if necessary), and restarting of dose after interruption (if applicable). The resting pulse and its change from pre-dose and percent change from pre-dose will also be categorized by pre-specified cutoffs as given in the table shells. The bradycardia events, bradycardia symptoms, and bradycardia treatment during the first dose, second dose (if necessary), and restarting of dose after interruption (if applicable) monitoring will be collected and summarized as well.

The bradycardia events, bradycardia symptoms, and bradycardia treatment will be collected on the corresponding CRFs. Note: Bradycardia events that do not have symptoms or require treatment are not collected in those CRFs and they are recorded in the Adverse Event CRF.

### **8.8.2 ECG**

ECG will be performed at pre-dose and 6 hours post dose during the first dose, second dose (if necessary), and restarting of dose after interruption (if applicable) monitoring.

### **8.8.3 Dose monitoring experience**

The overall dose monitoring experience refers to the following information: whether subjects are discharged after 6 hours post-dose or extended monitoring is required after 6 hours post-dose, whether subjects are hospitalized, whether subjects are required monitoring on the next day, whether subjects have symptomatic and/or treated bradycardia, whether subjects discontinue the study drug permanently, whether SAE is reported, and whether subjects are discharged but returned to clinic.

### **8.8.4 Data summaries for first dose monitoring**

Summaries on the safety set by treatment group include:

- Summary statistics of vital signs by hour (1<sup>st</sup> dose, 2<sup>nd</sup> dose, and restarts)
- Changes from pre-dose in vital signs by hour (1<sup>st</sup> dose, 2<sup>nd</sup> dose, and restarts)
- Frequency (%) distribution of categorized supine and standing pulse or change (or percent change) from pre-dose supine and standing pulse during first dose administration, respectively
- Summary of the overall dose monitoring experience (1<sup>st</sup> dose, 2<sup>nd</sup> dose, and restarts)
- Incidence rates of bradycardia events (1<sup>st</sup> dose and 2<sup>nd</sup> dose)
- Incidence rates of medications used for bradycardia (1<sup>st</sup> dose and 2<sup>nd</sup> dose)

- Incidence rates of bradycardia symptoms by SOC, PT, and severity (1<sup>st</sup> and 2<sup>nd</sup> dose)
- Incidence rates of notable vital sign abnormalities based on the notable criteria in Table 4-5 (1<sup>st</sup> dose and 2<sup>nd</sup> dose)
- Summary statistics of ECG parameters by time point during 1<sup>st</sup> and 2<sup>nd</sup> dose administration
- Changes from pre-dose in ECG parameters by time point during 1<sup>st</sup> and 2<sup>nd</sup> dose administration
- Incidence rates of abnormal QTc interval (Bazett and Fridericia) as defined by the abnormality criteria in Table 4-6 during 1<sup>st</sup> and 2<sup>nd</sup> dose administration
- Frequency (%) distribution of ECGs interpretation (normal/abnormal) by time point during 1<sup>st</sup> and 2<sup>nd</sup> dose administration
- Frequency (%) distribution of ECG findings by ECG evaluation type by time point during 1<sup>st</sup> and 2<sup>nd</sup> dose administration

## 8.9 Columbia Suicide Severity Rating Scale (C-SSRS)

### General

The Columbia Suicide Severity Rating Scale (C-SSRS) is a questionnaire that prospectively assesses suicidal ideation and behavior (SIB). There are two versions of the questionnaire:

- The “baseline/screening” version, which is to be performed at the first visit (Visit 1). It assesses SIB during the subject’s lifetime.
- The “since last visit” version to be used at subsequent visits (Visit 2 to Visit 7), assessing SIB since last visit.

A copy of the questionnaires can be found in protocol amendment [Appendix or CRF package](#).

The 11 categories described in [Table 7-1](#) have been adopted as standard by FDA to report SIB ([\[Food and Drug Administration, 2012\]](#), [\[Nilsson et al, 2013\]](#)). These categories include five levels of suicidal ideation, five levels of suicidal behavior and the category self-injurious behavior, no suicidal intent. Each category has a binary response (yes/no). The C-SSRS questionnaire allows suicidal ideation and behavior to be classified into these 11 preferred categories.

Note: the categories in [Table 7-1](#) have been reordered (based on [\[Nilsson et al 2013\]](#)) compared to the categories in the actual C-SSRS (paper) scale and the categories listed in the FDA guidance, in order to facilitate the definitions of endpoints and enable clarity in the presentation of results.

**Table 7-1 Standard SIB events**

Category number	C-SSRS category
<b>Suicidal Ideation</b>	

Category number	C-SSRS category
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
<b>Suicidal behavior</b>	
6	Preparatory acts or behavior
7	Aborted attempt
8	Interrupted attempt
9	Actual attempt
10	Completed suicide
<b>Self-injurious behavior, without suicidal intent</b>	
11	Non-suicidal self-injurious behavior

In this study, the C-SSRS questionnaire will be administered electronically (eC-SSRS), via an interactive voice response (IVR) system with the patient entering responses directly into the IVRS. Caregivers will not be allowed to answer the C-SSRS questions on behalf of the patient. Sites must review reports received from the system for any answers indicative of suicidal ideation and adverse event AEs. Adverse events ascertained through the administration of the C-SSRS will be documented. No safety cut-off will be applied for the SIB data reporting, i.e. all collected data will be used in the analysis including data collected off-treatment, if any.

#### Definition of 'all prior history'

*All prior history* will be defined as the SIB results obtained from the *lifetime* assessment at the first visit (Visit 1).

Worst case is defined by an answer 'yes' to the SIB category.

#### Data summaries at the study level

SIB data will be summarized for the Safety Set. The number and percentage of subjects with suicidal ideation, suicidal behavior and self-injurious behavior without suicidal intent will be presented by analysis-period (all prior history and any time post-baseline) and treatment group. The following 14 events will be included in the summary table:

- Each of the 11 categories listed in [Table 7-1](#), separately
- Any suicidal ideation or behavior (a 'yes' answer to at least one of the 10 suicidal ideation and behavior questions in analysis-period of interest)
- Any suicidal ideation (answered 'yes' to at least one of the 5 suicidal ideation questions in analysis-period of interest)

- Any suicidal behavior (answered 'yes' to at least one of the 5 suicidal behavior questions in analysis-period of interest).

For those analyses, each subject can only be counted once for each event. However, a subject can be counted in several different events.

Suicidal ideation and behavior data will be listed. Detailed answers to C-SSRS items will be listed separately for subjects with any suicidal ideation at any time post-baseline (i.e. a 'yes' answer to at least one of the five suicidal ideation questions at any time post-baseline) and for subjects with any suicidal behavior at any time post-baseline (i.e. a 'yes' answer to at least one of the five suicidal behavior questions at any time post-baseline).

## 9 Interim analyses

No efficacy interim analyses are planned. On an as-needed basis, safety interim analysis for the DSMB will be performed by an independent statistician and independent statistical programmer. The analysis plan for these safety interim analyses will be provided in a separate document.

## 10 Determination of sample size

The study will randomize a total of 1960 patients, and is planned to provide approximately 90% power for the comparison of fingolimod 0.5 mg to glatiramer acetate at a 2-sided significance level of 0.05.

The sample size calculation is based on simulations from a negative binomial distribution with a constant dispersion parameter  $k$ .

The power of the study was evaluated under various ARR assumptions and various dropout patterns based on the cumulative literature on glatiramer acetate and the Novartis data on fingolimod. The basis of the assumptions for this study and their level of uncertainty are presented in Section 9.4.2 under the sub-header "Multiplicity adjustment for statistical hypothesis testing". The anticipated overdispersion parameter ( $k=0.2231$ ) was observed in the Month-12 analysis of Study CFTY720D2301. The anticipated ARR for subjects treated with 0.5 mg fingolimod is  $\mu_{\text{FTY 0.5 mg}} = 0.195$ , the ARR for those treated with glatiramer acetate is  $\mu_{\text{glatiramer acetate}} = 0.30$ . Therefore, the estimated ARR reduction for fingolimod 0.5 mg versus glatiramer acetate is 35%.

The total sample size of 1960 randomized is predetermined but the exact sample size of each arm will depend on when the randomization ratio switch occurs. For example, if the randomization ratio is switched when a total of 800 patients have been randomized from an original ratio of 1:1:1 for fingolimod 0.25 mg, fingolimod 0.5 mg, or glatiramer acetate, respectively to a ratio of 5:3:2, then 847 patients will be randomized to fingolimod 0.25 mg, 615 patients to fingolimod 0.5 mg and 498 patients glatiramer acetate, which will provide more than 90% power to demonstrate superiority of fingolimod 0.5 mg dose versus glatiramer acetate in terms of ARR at a 2-sided significance level of

0.05 assuming a 15% drop-out rate. The calculations take into account that patients who discontinue prematurely from the study can participate with partial data to the primary endpoint.

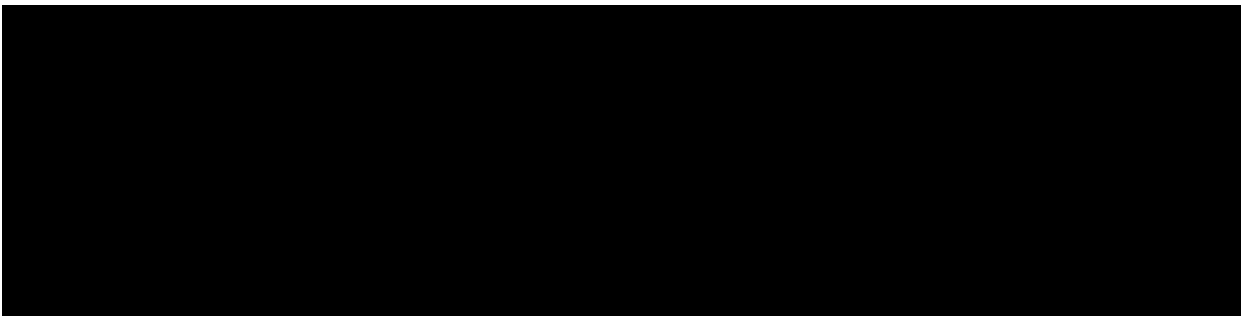
Fingolimod 0.25 mg has never been studied in a clinical trial in MS. Based on PK/PD modeling results, it is anticipated that the ARR in subjects treated with fingolimod 0.25 mg is approximately 15% higher than in those treated with fingolimod 0.5 mg. However, the uncertainty of this estimate is high; the 95% confidence interval of the estimated ARR in fingolimod 0.25 mg group ranges from 0.18 to 0.30. It is therefore anticipated that fingolimod 0.25 mg is less efficacious than fingolimod 0.5 mg, but more efficacious than glatiramer acetate. In line with the PK/PD modeling the anticipated ARR in subjects treated with fingolimod 0.25 mg is  $\mu_{\text{FTY 0.25 mg}} = 0.225$ , which corresponds to a reduction in ARR of 25% in subjects treated with fingolimod 0.25 mg compared to those treated with glatiramer acetate.

Following the multiplicity adjustment procedure in Multiplicity adjustment section 4.2.1.1, the power to detect a 25% reduction in ARR for subjects treated with fingolimod 0.25 mg compared with subjects treated with glatiramer acetate is approximately 68% at a 2-sided significance level of 0.05, if the primary objective for the fingolimod 0.5 mg dose can be rejected first; the corresponding marginal power with multiplicity adjustment is approximately 61%. If fingolimod 0.25 mg is similarly efficacious as fingolimod 0.5 mg (i.e., ARR reduction versus glatiramer acetate is 35% rather than 25%), the power to detect this treatment effect at a 2-sided significance level of 0.05 is approximately 88%. If fingolimod 0.25 mg is similarly efficacious as glatiramer acetate (i.e., estimated ARR in the fingolimod 0.25 mg group is at the high end of range proposed by the PK/PD model), there will be no significant difference at a 2-sided significance level of 0.05.

Due to the study early termination approved in December 2016 the final number of randomized patients is 1064 with 370 for fingolimod 0.25 mg, 352 for fingolimod 0.5 mg and 342 for glatiramer acetate. The statistical power based on the final sample size is approximately 71% for comparison of fingolimod 0.5 mg vs glactiramer acetate. For the comparison of fingolimod 0.25 vs glactiramer acetate the statistical power is approximately 44% without multiplicity adjustment and 31% with multiplicity adjustment.

The statistical software R (Version 2.13.1, open source) and the R library packages “MASS” and “PSCL” were used for sample size calculations and power analysis.

## 11 Documentation of changes from the protocol





## **12 Clinical Study Report - Appendix 16.1.9 Documentation of statistical methods**

### **12.1 Statistical methods and analysis outputs**

#### **12.1.1 Statistical methods**

Statistical Analysis System (SAS) version 9.2 or higher will be used to perform all the statistical analyses in the report.

Pair-wise treatment comparisons will be performed using separated data for treatment groups (A vs. C, B vs. C, and A vs. B) for the following analyses: rank ANCOVA/ANOVA, survival analyses, Wilcoxon rank-sum test, and Fisher's exact test. Data from all treatment groups (A, B, and C) will be used for the following analyses: negative binomial regression, logistic regression, and ANCOVA.

##### **12.1.1.1 Negative binomial regression model**

The primary efficacy variable (aggregate ARR) will be analyzed using a negative binomial regression model. This analysis will be performed on the full analysis set and the per-protocol set. All 3 treatment groups will be used to fit the regression model in the analysis. In addition, the negative binomial regression model will be used for several supportive analyses for aggregate ARR and number of new or newly enlarged T2 lesions.

Negative binomial regression is used for modeling count variables, usually for overdispersed count outcome variables, while adjusting for one or more covariates. In general, Negative binomial regression analyzes the distributed data in the following form

$$Y_i \sim NB(r_i, p_i), \text{ for } i = 1, \dots, n.$$

where the **negative binomial distribution** is a [discrete probability distribution](#) of the number of successes in a sequence of [Bernoulli trials](#) before a specified (non-random) number of failures (denoted  $r$ ) occur with a probability  $p$ .

The log of the count variables are modeled as a linear function of the  $X_i$ .

$$\log(y_i) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i}$$

Below is an example of the SAS codes used to perform the analysis of the negative binomial regression. The independent variables and covariates in the model may change per analysis defined in this document.

\*\*\*\*\*

\* SAS Codes: *Negative binomial regression* model

\* Variables in the model:

\* relnum = number of relapses

\* trt = treatment group code

\* reg1a = region code

\* n2 = number of relapses in previous year

\* bledss = baseline EDSS

\* sid1a = subject ID

\* lnday = log(days in study/365.25)

\*\*\*\*\*;

```
proc genmod data=datal;
```

```
  class trt reg1a;
```

```
  model relnum=trt reg1a n2 bledss / dist=nb
```

```
          link=log
```

```
          offset=lnday maxiter=500 lognb;
```

```
  lsmeans trt /cl;
```

```
  estimate 'A - B' trt 1 -1 0/exp;
```

```
  estimate 'A - C' trt 1 0 -1/exp;
```

```

estimate 'B - C' trt 0 1 -1/exp;

ods output estimates=est;
ods output lsmeans=lst;

run;

```

### 12.1.1.2 Rank Analysis of Covariance (ANCOVA) and Rank Analysis of Variance (ANOVA)

#### Rank Analysis of Covariance (ANCOVA)

The supportive efficacy variable, subject-level ARR (confirmed relapse only), will be analyzed using the rank ANCOVA. Other secondary efficacy variables, subject-level ARR (confirmed and unconfirmed relapse), change from baseline in [REDACTED] MRI variables, will be analyzed similarly on the full analysis set.

The rank ANCOVA is a non-parametric statistical method described in Stokes, Davis, and Koch (2000). It can be considered as an extension to the Wilcoxon rank-sum test with the ability to adjust for covariates in the model.

The analysis can be easily implemented using SAS by the following three steps:

- 1) compute the ranks of the response variable and covariate in the combined group of treatment using PROC RANK;
- 2) calculate the residuals from the linear regression of the response variable ranks vs. ranks of the covariates (without treatment) using PROC REG;
- 3) use the CMH mean score statistics to compare the mean values of the residuals in treatment by using TABLE scores (default scores) in PROC FREQ.

Below is an example of the SAS codes used to perform this rank ANCOVA. The independent variables and covariates in the model may change per analysis defined in this document.

\*\*\*\*\*

```

* SAS Codes: Rank ANCOVA model
* variables in the model:
* relann = subject-level annualized relapse rates
* trt = treatment group code
* reg1a = region code
* n2 = number of relapses in previous year
* bledss = baseline EDSS

```



```
* other covariates as well
* sid1a = subject ID
*****;
Data eff;
  Set eff;
  If reg1a=' ' or n2=. or bledss=. or relann=.then delete;
run;

proc rank data=eff nplus1 ties=mean out=ranks;
  by reg1a;
  var n2 bledss relann more covariates...;
run;

proc reg data=ranks noprint;
  by reg1a;
  model relann=n2 bledss more covariates...;
  output out=residual r=resid;
run;

proc freq data=residual;
  tables reg1a*trt*resid / noprint cmh2;
  where trt in ('A','B');
  output out=rancova cmh2;
  *** SELECT P_CMHRMS;
run;
```

### Rank Analysis of Variance (ANOVA)

Number of new or newly-enlarged T2 lesions will be analyzed using the rank ANOVA adjusted for treatment and region on the full analysis set. Unlike Rank ANCOVA where continuous covariates exist, the PROC RANK and PROC REG steps are not necessary. The van Elteren's test will be used for treatment comparisons. The van Elteren's test is a rank based test and can be implemented in PROC

FREQ with options CMH2 and score equal to MODRIDIT. The van Elteren's test statistic corresponds to the row mean scores statistic labeled "Row Mean Scores Differ" in the CMH output.

Below is an example of the SAS codes used to perform this rank ANOVA. The independent variables and covariates in the model may change per analysis defined in this document.

```
*****
* SAS Codes: Rank ANOVA model
* variables in the model:
* RSLVAL1N = number of new or newly-enlarged T2
* trt = treatment group code
* reg1a = region code
* sid1a = subject ID
*****;
proc freq data= eff;
  where trt in ('A','B');
  tables reg1a*trt*RSLVAL1N / noprint cmh2 scores=MODRIDIT;
  output out=ranova cmh2;
  *** SELECT P_CMHRMS;
run;
```

### 12.1.1.3 Logistic regression for proportion variables

The other secondary/other efficacy variables ( ), proportion of subjects free of new or newly enlarging T2 lesions, proportion of subjects free of Gd + lesions, proportion of subjects free of new T1 lesion) will be analyzed using the logistic regression on the full analysis set. All 3 treatment groups will be used in the model.

Logistic regression is a model for prediction of the probability of occurrence of an event. It can be used to analyze the dichotomous response data while adjusting for one or more covariates. Usually, Logistic regression analyzes binomially distributed data of form

$$Y_i \sim B(n_i, p_i), \text{ for } i = 1, \dots, n.$$

The logits of the unknown binomial probabilities (*i.e.*, the logarithms of the odds) are modeled as a linear function of the  $X_i$ .

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i}.$$

Below is an example of the SAS codes used to perform the analysis of the LOGISTIC regression. The independent variables and covariates in the model may change per analysis defined in this document.

\*\*\*\*\*

\* SAS Codes: *LOGISTIC* model

\* Variables in the model:

\* mrifree = proportion of subjects free of Gd + lesions

\* reg1a = region code

\* blmri = baseline number of Gd + lesions

\* more covariates could be added as needed

\*\*\*\*\*;

```
proc logistic data=eff descending;
  class reg1a trt /param=ref;
  model mrifree = reg1a trt blmri (more covariated as needed);

  contrast ' A - B ' trt 1 -1 0/estimate=exp;
  contrast ' A - C ' trt 1 0 -1/estimate=exp;
  contrast ' B - C ' trt 0 1 -1/estimate=exp;

  ods output Contrastestimate=est;
run;
```

For this study, proc logistic is performed including all 3 treatment levels with the contrast specifying 2 levels.

#### 12.1.1.4 Survival analysis for time to event variables

Log-rank test will be used to compare the survival distributions of two treatment groups for secondary efficacy time to event variables, [REDACTED]. These analyses will be performed on the full analysis set.

The logrank test is a [hypothesis test](#) to compare the [survival](#) distributions of two samples. It is a [nonparametric](#) test and appropriate to use when the data are right skewed and [censored](#) (technically, the censoring must be non-informative). It is constructed by computing the observed



```

*****;
***include all 3 treatment arms data in one model fitting***
proc phreg data=datatn nosummary;
  class trt regla
  model reltm*relcens(1)=trt n2 bledss regla (more covariates as needed)/rl
  ties=exact;
  contrast "A vs C" trt 1 0 -1/estimate=exp;
  contrast "B vs C" trt 0 1 -1/estimate=exp;
  contrast "A vs B" trt 1 -1 0/estimate=exp;
  ods output parameterestimates=pest;
  ods output contrastestimate=est;
run;
**Note that data set pest contains Hazard ratio estimates from model
statement where only A vs C and B vs C are available. If parameter='trt',
then SELECT probchisq hazardratio hrlowercl hruppercl for results;
**However, data set est contains Hazard ratio estimates from contrast
statement where all 3 pairwise comparisons are available (result variable
names: estimate lowerlimit upperlimit probchisq);

```

#### 12.1.1.4.1 Estimation of 95% confidence intervals

Approximate 95% confidence intervals will be generated for several estimates: (1) at selected time-points, the (Kaplan and Meier 1958) estimates of the proportion of subjects experiencing a specific event, (2) at selected time-points, the between-treatment group difference in Kaplan-Meier estimates of the proportion of subjects experiencing a specific event, (3) for certain study periods, the between-treatment group difference in the proportion of subjects experiencing a specific event, and (4) at selected time-points, the between-treatment group difference in means. Details of such confidence intervals are displayed in Table 5-1.

**Table 12-1 Confidence Interval calculations**

	Approximate Confidence Interval (CI)	Lower Limit	Upper Limit
1	x% CI for the Kaplan-Meier estimate $KM$	$KM - c_x * SE$	$KM + c_x * SE$
	Notation: $SE$ = estimated standard error of $KM$ ; estimation of standard error uses Greenwood's formula, $c_x = \Phi^{-1}(\frac{1}{2} + x/200)$ , where $\Phi$ is the normal distribution function.		
2	x% CI for the difference in Kaplan-Meier estimates $KM_2 - KM_1$	$KM_2 - KM_1 - c_x * SE_d$	$KM_2 - KM_1 + c_x * SE_d$

Notation:  $KM_i$  = Kaplan-Meier estimate for group  $i = 1,2$ ,

$SE_d = \sqrt{(SE_1^2 + SE_2^2)}$ , where  $SE^2$  = estimated standard error of  $KM_i$ ,  $i = 1,2$ ;

estimation of standard error uses Greenwood's formula,

$c_x = \Phi^{-1}(1/2 + x/200)$ , where  $\Phi$  is the normal distribution function.

Examples: for  $x = 95$ ,  $c_x = \Phi^{-1}(0.975) = 1.95996$ ; for  $x = 97.5$ ,  $c_x = \Phi^{-1}(0.9875) = 2.24140$

### 12.1.1.5 Between-treatment group comparisons for other efficacy variables

Generally, non-stratified, non-parametrical tests will be used to perform between-treatment group comparisons on MS confirmed relapse characteristics and change from baseline in disability, MRI variables (Gd + T1 lesion count, T2 lesion volume, Gd + lesion volume, T1 hypointense lesion volume, and changes from baseline in T2 lesion volume and Gd + lesion volume, percentage change from baseline in brain volume). Tests applied will be the rank ANCOVA. For a small number of comparisons, Wilcoxon rank-sum test for continuous variables will be utilized; Fisher's exact test will be used for categorical variables; for variables with more than five categories it will be replaced with the chi-square test. Wilcoxon rank-sum test is also used for comparison between the treatment groups [REDACTED].

Below is an example of the SAS codes used to perform these analyses.

\*\*\*\*\*

\* SAS Codes: *Wilcoxon rank-sum* test

\* Variables in the model:

\* [REDACTED]

\* trt = treatment group code

\*\*\*\*\*;

```
proc nparlway data=eff wilcoxon noprint;
```

```
  where trt in ('A','B');
```

```
  class trt;
```

```
  var C_MSACT;
```

```
run;
```

\*\*\*\*\*

\* SAS Codes: *Fisher's exact* test

\* variables in the model:

\* pevt = 1-subject with confirmed relapses;

0-subject without confirmed relapses;

\* trt = treatment group code

\*\*\*\*\*;

proc freq data=eff noprint;

    where trt in ('A','B');

    tables trt\*pevt / exact;

run;

\* Select p-value for normal approximation from Wilcoxon Two-Sample Test

### 12.1.1.6 Analysis of Covariance (ANCOVA) for quality of life (QoL) variables

The quality of life variables:

\*\*\*\*\*

\* SAS Codes: *ANCOVA model*

\* variables in the model:

\* c\_msphs =

\* trt = treatment group code

\* reg1a = region code

\* age\_1n = age at baseline

\* =

\*\*\*\*\*;

proc mixed data=input dataset;

    class reg1a trt;

    lsmeans trt /pdiff cl;

run;

### One sample t-test for quality of life (QoL) variables

The regular one-sample t-test will be used to perform within-group comparisons of continuous (or quantitative) efficacy variables where the normality assumption is held reasonably well. The within-

group analyses for continuous / quantitative variables will compare assessments for the end of treatment vs. baseline within each treatment group. Examples of such analysis include the within group comparisons of the quality of life variable: change from baseline in TSQM total score.

Below is an example of the SAS codes used to perform the analysis. All 3 treatment groups will be used in the model.

\*\*\*\*\*

\* \* SAS Codes: *One sample t-test*

\* V1 = First timepoint --- e.g.:(baseline);

\* V2 = Second timepoint --- e.g.:(post-baseline visit, end of treatment);

\* PID = Subject ID or number of observations;

\* diff= difference between two timepoints (higher scores indicate better results);

\* NOTE: First, prepare and extract the data for each treatment group.

\* Then, in each treatment group, break up the data into the 2 timepoints.

\* Ensure that each subject has a value at both timepoints (V1 and V2).

\*\*\*\*\*;

```
data one;
```

```
  input PID V1 V2;
```

```
  diff = V2 - V1;
```

```
proc ttest data = one sides=u;
```

```
  by treat;
```

```
  var = diff;
```

```
run;
```

### 12.1.1.7 Wilcoxon sign test for quality of life (QoL) variables

Wilcoxon signed-rank test (also known simply as the Wilcoxon sign test) can be thought of as the non-parametric equivalent of the regular one-sample t-test. It will be used to perform all within-group comparisons of continuous (or quantitative) efficacy variables. The within-group analyses for continuous / quantitative variables will compare assessments for the end of treatment vs. baseline within each treatment group. Examples of such analysis include the within group comparisons of the quality of life variable: change from baseline in TSQM total score.



Below is an example of the SAS codes used to perform the analysis. All 3 treatment groups will be used in the model.

```
*****
* * SAS Codes: Wilcoxon signed-rank test (or wilcoxon sign test)
* V1 = First timepoint --- e.g.:(baseline);
* V2 = Second timepoint --- e.g.:(post-baseline visit, end of treatment);
* PID = Subject ID or number of observations;
* diff= difference between two timepoints;
* NOTE: First, prepare and extract the data for each treatment group.
* Then, in each treatment group, break up the data into the 2 timepoints.
* Ensure that each subject has a value at both timepoints (v1 and v2).
*****;
data one;
    input PID V1 V2;
    diff = V2 - V1;

proc univariate data = one;
    by treat;
    var = diff;
run;
```

### 12.1.2 Statistical analysis outputs

Results from the primary and secondary efficacy analyses will be presented in [Section 16.1.9](#) of the clinical study report.

The list of outputs that will be generated is provided in RAP Module 7.1.

## 13 References

Atkinson MJ, Sinha A, Hass SL et al. Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire (TSQM) using a national panel study of chronic disease. *Health and Quality of Life Outcomes* 2004; 2:12

Bretz, F, Maurer, W, Brannath, W et al (2009) A graphical approach to sequentially rejective multiple test procedures *Stat Med* 28(4), 586–604.

[REDACTED]

FDA (2012) Guidance for industry. Suicidal Ideation and Behavior: Prospective Assessment of Occurrence in Clinical Trials (last draft August 2012, available under <http://www.fda.gov/downloads/Drugs/Guidances/UCM225130.pdf>)

Gassmann-Mayer et al (2011) Clinical and statistical assessment of suicidal ideation and behavior in pharmaceutical trials. *Clin Pharmacol Ther.* 2011 Oct;90(4):554-60.

[REDACTED]

Kaplan, EL, Meier, P. (1958) Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assn.* 53: 457-481.

Keene ON, Jones MRK, Lane PW, Anderson J., et al. (2007) Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: example from the TRISTAN study. *Pharmaceutical Statistics* 6: 89-97.

Nilsson M. et al (2013) Columbia-Suicide Severity Rating Scale. Scoring and data analysis guide. (available at [http://www.cssrs.columbia.edu/C-SSRS%20Scoring%20and%20Data%20Analysis%20Guide\\_Cross-Industry%20Working%20Group\\_v2.0Final.pdf](http://www.cssrs.columbia.edu/C-SSRS%20Scoring%20and%20Data%20Analysis%20Guide_Cross-Industry%20Working%20Group_v2.0Final.pdf))

Stokes ME, Davis CS, Koch GG (2000). *Categorical data analysis using the SAS system.* Second edition. SAS Institute, Inc., Cary, N.C.

## 14 Appendices

### 14.1 Appendix 1: Geographic location (state/providence and country) of the study centers

Refer to clinical document CFTY720D2312 Global Qualified Sites list for study centers and their geographic location information. Below is the list of study centers with at least one randomized subject and also their region classification.

Center Number	Country	State/Province	Region
[REDACTED]	Argentina	CABA	Latin American region

Center Number	Country	State/Province	Region
████	Argentina	CABA	Latin American region
████	Brazil	MG	Latin American region
████	Brazil	RS	Latin American region
████	Brazil	SC	Latin American region
████	Brazil	RJ	Latin American region
████	Brazil	RS	Latin American region
████	Brazil	GO	Latin American region
████	Brazil	SP	Latin American region
████	Brazil	PR	Latin American region
████	CANADA	Nova Scotia	Northeast region
████	CANADA	Ontario	Northeast region
████	CANADA	Alberta	West region
████	CANADA	Quebec	Northeast region
████	CANADA	Quebec	Northeast region
████	CANADA	Quebec	Northeast region
████	CANADA	British Columbia	West region
████	Chile	Santiago	Latin American region
████	Chile	Santiago	Latin American region
████	Mexico	Chihuahua	Latin American region
████	Mexico	DF	Latin American region
████	Mexico	Chihuahua	Latin American region
████	Mexico	Nvo Leon	Latin American region
████	Mexico	Nvo Leon	Latin American region
████	Mexico	Aguascalientes	Latin American region
████	Mexico	San Luis Potosi	Latin American region
████	Mexico	DF	Latin American region
████	UNITED STATES	MI	Midwest region

Center Number	Country	State/Province	Region
████	UNITED STATES	IN	Midwest region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	OH	Midwest region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	AZ	West region
████	UNITED STATES	GA	Southeast region
████	UNITED STATES	NJ	Northeast region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	TX	Southwest region
████	UNITED STATES	MI	Midwest region
████	UNITED STATES	CA	West region
████	UNITED STATES	OH	Midwest region
████	UNITED STATES	CO	Southwest region
████	UNITED STATES	OK	Southwest region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	NM	Southwest region
████	UNITED STATES	TX	Southwest region
████	UNITED STATES	IL	Midwest region
████	UNITED STATES	WA	West region
████	UNITED STATES	NJ	Northeast region
████	UNITED STATES	NY	Northeast region
████	UNITED STATES	UT	West region
████	UNITED STATES	KS	Southwest region
████	UNITED STATES	TN	Southeast region
████	UNITED STATES	TX	Southwest region
████	UNITED STATES	MI	Midwest region
████	UNITED STATES	TN	Southeast region

Center Number	Country	State/Province	Region
████	UNITED STATES	KS	Southwest region
████	UNITED STATES	VA	Southeast region
████	UNITED STATES	OH	Midwest region
████	UNITED STATES	IL	Midwest region
████	UNITED STATES	CO	Southwest region
████	UNITED STATES	NY	Northeast region
████	UNITED STATES	AZ	West region
████	UNITED STATES	NC	Southeast region
████	UNITED STATES	OH	Midwest region
████	UNITED STATES	AZ	West region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	TN	Southeast region
████	UNITED STATES	DE	Northeast region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	NY	Northeast region
████	UNITED STATES	CA	West region
████	UNITED STATES	NY	Northeast region
████	UNITED STATES	VA	Southeast region
████	UNITED STATES	CO	Southwest region
████	UNITED STATES	NY	Northeast region
████	UNITED STATES	OK	Southwest region
████	UNITED STATES	TX	Southwest region
████	UNITED STATES	DC	Northeast region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	IL	Midwest region
████	UNITED STATES	CO	Southwest region
████	UNITED STATES	FL	Southeast region

Center Number	Country	State/Province	Region
████	UNITED STATES	MO	Southwest region
████	UNITED STATES	NJ	Northeast region
████	UNITED STATES	MO	Southwest region
████	UNITED STATES	VA	Southeast region
████	UNITED STATES	AL	Southeast region
████	UNITED STATES	IN	Midwest region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	NY	Northeast region
████	UNITED STATES	NC	Southeast region
████	UNITED STATES	IA	Midwest region
████	UNITED STATES	IL	Midwest region
████	UNITED STATES	NV	West region
████	UNITED STATES	OK	Southwest region
████	UNITED STATES	MO	Southwest region
████	UNITED STATES	OH	Midwest region
████	UNITED STATES	MA	Northeast region
████	UNITED STATES	MA	Northeast region
████	UNITED STATES	KY	Midwest region
████	UNITED STATES	OH	Midwest region
████	UNITED STATES	CO	Southwest region
████	UNITED STATES	PA	Northeast region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	TX	Southwest region
████	UNITED STATES	MO	Southwest region
████	UNITED STATES	CT	Northeast region
████	UNITED STATES	NY	Northeast region
████	UNITED STATES	TN	Southeast region

Center Number	Country	State/Province	Region
████	UNITED STATES	AZ	West region
████	UNITED STATES	WA	West region
████	UNITED STATES	VA	Southeast region
████	UNITED STATES	MI	Midwest region
████	UNITED STATES	TX	Southwest region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	LA	Southeast region
████	UNITED STATES	SC	Southeast region
████	UNITED STATES	NC	Southeast region
████	UNITED STATES	MI	Midwest region
████	UNITED STATES	WA	West region
████	UNITED STATES	MT	West region
████	UNITED STATES	PR	Southeast region
████	UNITED STATES	OR	West region
████	UNITED STATES	WI	Midwest region
████	UNITED STATES	FL	Southeast region
████	UNITED STATES	VA	Southeast region
████	UNITED STATES	OH	Midwest region