NOVARTIS

Clinical Development

FTY720D/fingolimod

CFTY720D2311 / NCT01892722

A double-blind, randomized, multicenter, active-controlled study of flexible duration (maximum 2 years) to evaluate the safety and efficacy of fingolimod administered orally once daily versus interferon β-1a i.m. once weekly in pediatric patients with multiple sclerosis with five-year fingolimod Extension Phase

RAP Module 3 – Detailed Statistical Methodology

(Core Phase)

Author: [redacted]

Document type: RAP Documentation

Document status: Amendment 2, Version 3.0 (Final)

Release date: August 4, 2017

Number of pages: 59

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

Version	Date	Changes
Final1.0	May 15, 2013	NA
Final 2.0	October 11, 2013	• Removed OPH follow up set analysis as data is not collected after study drug discontinuation • Corrected the formula for PFT predicted value per clinical team's confirmation that there's a typo in the protocol • [redacted]
Amendment 1	September 19, 2016	• Update due to the study design update. • Modified the analysis population section by listing the PDs that warrant the exclusion of patients from a particular analysis set following the new PD process • Added the sensitivity analysis for the primary/key secondary efficacy endpoint by excluding INF-beta patients who are Nabs positive at end of study visit • removed/added endpoints per protocol amendments • updated clinically notable criteria per protocol amendments
Amendment 2	August 4, 2017	• For all the efficacy analysis model adjusted by the stratification factors, the small regions will be pooled into one; also, only LSMean using the OM option will be provided given the unbalanced strata in this study • Specified that any data collected during follow up but after the core phase data base lock will be reported in the extension phase CSR in section 9.1 • Added a subgroup summary for primary and key secondary efficacy by baseline body weight • Modified table 3-1 by adding more PDs that may be excluded from PP set per PD spec update. also added footnote. • Updated the FU set definition • Added the algorithm on what CSSR source data should be used for analysis • Added analyses supporting RMP report • Due to the update on CCG that the CMP form should be entered at EOS visit for those who completed the study and entered the 3-M follow up. The last visit date in core phase is now always before the follow up period and can be used to calculate time in study for ARR. The algorithm in section 4.1.2 is simplified. • Replaced the term of compliance rate with drug accountability rate as well as the definition • Analysis on screening failures is based on number of unique failures and reason is provided • Added details/clarifications to address the questions during programming

Table of contents

List of abbreviations

AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ARBA	Annual rate of brain atrophy
ARR	annualized relapse rate
AST	aspartate aminotransferase
BMI	Body mass index
CI	Confidence interval
CMH	Cochran-Mentel-Haenszel
[redacted]	[redacted]
EOS	End of study, the visit when the study completion eCRF is entered.
EOT	End of treatment
FUS	Follow-up set
CFT	Central foveal thickness
[redacted]	[redacted]
C-SSRS	Columbia Suicide Severity Rating Scale
DAE	AE leading to permanent study drug discontinuation
DLCO	diffusion capacity for carbon monoxide
DMC	Data monitoring committee
ECG	electrocardiogram
eCRF	case report/record form
EDSS	expanded disability status scale
FAS	Full analysis set
FEV1	forced expiratory volume in 1 second
FVC	forced vital capacity
Gd	gadolinium
GGT	gamma-glutamyl-transferase
IFN	interferon
i.m.	intramuscular(ly)
IVRS	interactive voice response system
KM	Kaplan-Meier
LLOD	Lower limit of detection
MedDRA	medical dictionary for regulatory activities
MRI	magnetic resonance imaging
MS	multiple sclerosis
Nabs	Neutralizing antibodies

NB	Negative binomial
OCT	optical coherence tomography
[redacted]	[redacted]
PFT	pulmonary function tests
PPS	Per-protocol set
PRO	Patient reported outcome
PT	Preferred term
RAN	Randomized set
RBC	red blood cell
RMP	Risk management plan
RRMS	relapsing-remitting MS
SAE	serious adverse event
SAF	Safety set
SCR	All screened subjects
[redacted]	[redacted]
SI	standard international
SOC	System organ class
[redacted]	[redacted]
[redacted]	[redacted]
UNL	the upper limit of normal range
WBC	white blood cell
WHO	World health organization
[redacted]	[redacted]

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

This document is specifying the analysis plan on data from the core phase. Data will be analyzed by Novartis internal statisticians and statistical programmers assigned for this study according to the data analysis section 9 of the study protocol 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

The planned analysis is described in Section 9 of the protocol (or CSR Appendix 16.1.1).

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

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviations, median, 25th and 75th percentiles (for some MRI parameters only), minimum and maximum will be presented.

The primary objective of the study is to compare fingolimod (FTY720) with interferon beta-1a and to demonstrate that FTY720 is superior to interferon beta-1a in terms of annualized relapse rate for the pediatric patients with RRMS (i.e., children/adolescent MS patients aged 10 to less than 18 years) treated for up to 24 months. The FTY720 arm includes patients receiving FTY720 0.5mg once daily (if body weight >40 kg at randomization) and a small number of patients receiving 0.25 mg FTY720 once daily (if body weight <=40kg at randomization). The low dose may be switched to full dose if the 2-month PK data indicating insufficient serum level exposure or the patient's body weight increased to >40kg during the study.

The statistical models specified in this document may be modified by including fewer covariates in the models in case the pre-specified models do not converge after fitting to the trial data from final database lock.

3 Subjects and treatments

3.1 Analysis sets

All screened subjects (SCR): The SCR set comprises all subjects who were screened. The SCR will be used only for the summary of screening failures.

Randomized set (RAN): The RAN set comprises all subjects who were assigned a randomization number. The RAN will be used for the summaries of subject disposition, protocol deviation, demographic, baseline characteristics, and medical history data.

Full Analysis Set (FAS): The FAS comprises all randomized subjects with assigned treatments who took at least one dose of study medication. Subjects will be analyzed according to the randomized treatment assignment following the intention-to-treat principle,

even if they actually received a different treatment. The FAS will be used for all efficacy analyses.

Safety Set (SAF): The safety set includes all subjects who received at least one dose of study medication. Subjects will be analyzed according to the actual treatment received. The Safety Set will be used for all safety analyses. Subjects will be analyzed according to the treatment they are assigned to at randomization unless a subject takes non-randomized study medications and does not take any study medication as randomized, in that case the subject will be analyzed according to the treatment actually taken.

Per-protocol set (PPS): The PPS is a subset of FAS, consists of all randomized subjects who take at least one dose of study medication and have no major protocol deviations that could confound the interpretation of analyses conducted on the FAS. Major protocol deviations will be determined according to the pre-defined protocol deviation criteria before treatment unblinding. For analyses performed on the PPS, any efficacy data assessed after permanent discontinuation of study medication will be excluded. The PPS will primarily be used for the supportive analyses of the primary efficacy variable and the key secondary variable.

Follow-up set (FUS): The FUS consists of all subjects in the safety set who have at least one visit or who have AE, MS relapse, or concomitant mediation starting after study drug discontinuation, but before first dose of the extension study medication (if applicable). Subjects will be grouped in the same way as described above for the analysis on the safety set. Only the safety follow up data analysis will be performed on the follow-up set.

Subject classification in the analysis sets is entirely based on protocol deviation and non-protocol deviation criteria. A list of protocol deviations are specified prospectively in the Edit Check Specification. A subset of these protocol deviation criteria results in the exclusion of the subjects from analysis.

Table 3-1 provides a list of pre-defined protocol deviations considered important enough to lead to exclusion of a subject from an analysis set. These are taken from the Edit Check Specifications.

Table 3-1 Specifications of important protocol deviations

Deviation ID	Description of deviation	Excl. from analysis dataset				
		SCR	RAN	FAS	SAF	PPS
I01	Informed consent was not obtained. However, study assessments were performed.	X	X	X	X	X
E01	Patient has progressive multiple sclerosis. However, patient was randomized.			X		X
E04	Patient fulfills either the definition of ADEM or the criteria for neuromyelitis optica. However, patient was randomized.					X
E04A	Any patient is positive for aquaporin 4 at Screening. However, patient was randomized					X
E05B	Patient was treated with high dose intravenous immunoglobulin within 2 months prior to randomization. However, patient was randomized.					X

Deviation ID	Description of deviation	Excl. from analysis dataset				
		SCR	RAN	FAS	SAF	PPS
E05E	Any patient previously treated with excluded immunosuppressive or DMT medications within per protocol (Amend 4 or higher criteria) specified time windows at any time. However patient was randomized.					X
E05F	Patient has been previously treated with fingolimod at any time. However patient was randomized.					X
E26	Patient took study drug or therapy within 180 days or 5 half‑lives of Randomization, whichever is longer. However, patient was randomized.					X
I03	Central MRI review completed and diagnosis of MS not confirmed. However, patient was randomized.			X		X
I04B	Number of relapses within: 1 yr prior to randomization is < 1, or 2 yrs prior to randomization is < 2; or 6 mths no Gd lesions. However patient was randomized.					X
I05	EDSS score at screening/baseline was greater than 5.5. However, patient was randomized.					X
M01	Patient used immunosuppressive medication during double blind treatment					X
O02[1]	Site did not follow per protocol blinding procedures such that the integrity of the study may have been compromised.					X
O02A[1]	Sponsor provided potentially unblinding information to the site that could have compromised the integrity of the study.					X
S03A	Compliance less than 80 percent with study medication based on drug accountability log at site over entire study.					X
S05[2]	Incorrect study medication kit given and does not match assigned randomized number for more than 20% of study duration.					X
S07	expired study medication kit given for more than 20% of study duration					X
S12[3]	Patients body weight increases to >40 kg sustained over 2 scheduled visits (3 months apart) and dose is not increased to 0.5 mg for more than 90 days					X
S14	Monitoring of drug compliance was inconsistent or incomplete, resulting in an inadequate documentation for drug accountability					X

[1] Findings with strong evidence of potential unblinding in the data (e.g. blinding information that belongs to the blinded database but is reported in the main database) after DBL will be defined and added as new entries for PD O02 or PD O02A in VIOPTO and excluded from PPS analyses.

[2] Only exclude such PD from PPS if the actual treatment based on the kit numbers (available after DBL) does not match the randomized treatment.

[3] Only exclude such PD from PPS if the patient was randomize to FTY arm.

3.2 Subject disposition

According to the protocol, a subject who failed screening can be rescreened under a new subject ID. If a patient was randomized after rescreening, a rescreening log was completed to document the original subject ID the patient was assigned. However, if a patient failed a rescreening, no rescreening log was completed. As a result, there is not enough information from the clinical database to identify the number of unique patients who were screened/rescreened for this study. Only the number of unique subject IDs that had been assigned for each screening/rescreening are available. Therefore, the total number of screenings will be presented, along with the reason for each failed screening.

For patients who failed a screening and were later randomized, the reason for first failed screening will also be summarized.

A listing with all the screening/rescreening under a unique subject ID will be provided.

Subject disposition for study and study treatment will be summarized on the RAN set. The number and proportion of subjects who complete the study or discontinue the study prematurely along with the reason for discontinuation will be summarized by randomized treatment assignment. The number and proportion of subjects who discontinue the study drug prematurely along with the reason for study drug discontinuation will be summarized by randomized treatment assignment.

Subjects who discontinue early from the study treatment or those who complete the study treatment according to the study design but decide not to enter the extension study, will have a 3-month follow up visit. The number and proportion of the subjects who missed follow up visit, along with the reason, will also be summarized on the RAN set.

The number and proportion of subjects in each analysis set will be presented by treatment group.

3.3 Protocol deviation

Protocol deviations will be summarized by deviation categories and treatment for the RAN set. In addition, protocol deviations that led to exclusion from PPS will be summarized by deviation category, deviation terms and treatment groups for the RAN 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 eCRF page), that subject will also be excluded from the PPS. A summary of those subjects will also be presented for the RAN set.

3.4 Demographics and other baseline characteristics

Subject demographics and other baseline characteristics include age, gender, race, and ethnicity collected on the demography eCRF, height and body weight recorded at baseline on the vital signs eCRF, body mass index (BMI) calculated as (body weight in kilograms) / (height in meters) ^{2} at baseline, pubertal status based on Tanner stage assessed at screening

visit (pre-pubertal if Tanner staging score<2; pubertal if Tanner staging score >=2 or missing given an observation of bone age ≥16 yrs), and pubertal status (from interactive IVRS, i.e., I based on investigator assessment) at the randomization visit. Age will be calculated from date of first dose of study drug and date of birth. For subjects without a first dose date, date of randomization visit will be used.

The race will be reported using the designations (different from the categories used on the CRF) shown on Table 3-2.

Table 3-2 Designations for race in analysis

Race collected on CRF	Designations for Race in analysis	Race (shortened version in ASR)
Native American	American Indian or Alaska Native	Age/Sex/AI
Asian	Asian	Age/Sex/AS
Black	Black or African American	Age/Sex/BL
Pacific islander	Native Hawaiian or other Pacific Islander	Age/Sex/NH
Caucasian	White	Age/Sex/WH
Other	Other	Age/Sex/OT

ASR: age, sex, race variable.

These variables will be summarized on the RAN set by presenting frequency distributions (for categorical variables) or descriptive statistics of mean, standard deviation, minimum, median and maximum (for continuous variables).

For baseline comparability analysis, categorical variables will be analyzed by Cochran-Mantel-Haenszel (CMH) test stratified by region and pubertal status; continuous variables will be analyzed using ANOVA with covariates of treatment, region, and pubertal status. This analysis will be performed on the RAN set and will be presented in Appendix 16.1.9 only. For categorical variables, the number and percentage of subjects with missing data will be provided but those subjects who are in the missing category will not be included in the denominator for the CMH test.

Note that these tests of comparability are performed for descriptive purposes only, and will not serve as a basis for determining entry of explanatory variables into the respective models. However, when these tests yield significant results they can be used as supportive information in interpreting the statistical analyses performed on the primary and secondary efficacy variables.

3.5 MS disease and medical history

3.5.1 MS baseline disease characteristics

MS baseline disease characteristics, MS disease history, MS medication history, and MS related eye history will be summarized by treatment group for the RAN set.

MS baseline characteristics include baseline EDSS, key MRI parameters (e.g., number and volume of Gd-enhancing lesions, number and volume of T2 lesions, volume of T1 hypointense lesions, and normalized brain volume)

MS disease history includes duration of MS since diagnosis (years), duration of MS since first symptom (years), number of relapses as well as the number of relapses that required steroid treatment (in the last 12 months prior to randomization visit, between the last 12 to 24 months prior to randomization visit, and in the last 24 months prior to randomization visit), and time since onset of most recent relapse (months). Duration of MS since diagnosis (years) will be derived as [(first dose date – MS diagnosis date + 1)/365.25]; duration of MS since first symptom (years) will be derived as [(first dose date – first MS symptom date +1)/365.25]; and time since onset of most recent relapse (months) will be derived as [(first dose date – most recent relapse onset date + 1)/30]. For subjects without a first dose date, date of randomization visit will be used.

For the MS diagnosis date and the first MS symptom date, a partial date will be imputed as follows: a) if only year yyyy is available, the imputed date will be 1-July-yyyy; b) if both year yyyy and month mmm are available, the imputed date will be 15-mmm-yyyy.

For the most recent relapse onset date, a partial date will be imputed as follows: a) if only year yyyy is available, the imputed date will be 1-Jan-yyyy; b) if both year yyyy and month mmm are available, the imputed date will be 01-mmm-yyyy.

MS medication history of previous disease-modifying drugs (coded by WHO drug dictionary) will be summarized by preferred term (PT). The number and proportion of treatment-naïve subjects (i.e., subjects who have not been treated with any disease-modifying drug before study enrolment) will also be presented.

MS related eye history includes history data collected at screening visit on the ophthalmologic examination eCRF.

3.5.2 Medical history

Any condition entered on the relevant medical history/current medical conditions eCRF will be coded using MedDRA. The relevant medical history and continuing medical conditions will be summarized by primary system organ class (SOC), preferred term (PT) and treatment group for the RAN set.

3.6 Treatments (study medication, rescue medication, other concomitant therapies)

3.6.1 Study medication

The duration (days) of intake to randomized study medication will be summarized on the SAF set by presenting the number (and percentage) of subjects taking the study medication for a minimum duration of time as specified in Table 3-3. Duration in terms of number of days as shown in the second column will be used in the summary tables. Summary statistics of duration to study medication will also be presented. The exposure to the actual study mediation taken will also be presented in the pre-pubertal (Tanner staging score <2) and young subjects (<=12 years) subgroups.

Duration of study medication intake is the number of days on study drug during the study. The days when the subject does not take the study drug will be excluded. That is, duration

will exclude periods of temporary interruption of study medication. Only data from the core phase will be included. See Section 3.8.3 for definition of core phase data.

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

3.6.2 Drug accountability

To measure compliance, the number of study drug capsules/ syringes dispensed and returned from the drug accountability page at each visit will be used to derive the drug accountability rate for each patient:

Drug accountability rate $(\%)=\frac{\#\text{ capsules/syringes dispensed}-\#\text{ capsules/syringes returned}}{\#\text{ capsules/syringes expected to be taken}} * 100$, where:

capsules/syringes expected to be taken is assuming the consumption of one capsule per day or one injection per week during the treatment period recorded on drug accountability form:

capsules expected to be taken = date of last dose from drug dispensed-date of first dose from drug dispensed+1

syringes expected to be used= ceil ((date of last dose from drug dispensed-date of first dose from drug dispensed+1)/7)

The Drug accountability rate will be summarized by treatment for fingolimod/placebo and IFN β-1a/placebo.

Table 3-3 Time-intervals used to summarize study medication intake

Duration	Duration in terms of number of days
Any	≥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
≥15 months	≥450 days
≥18 months	≥540 days
≥21 months	≥630 days
≥24 months	≥720 days
≥25 months	≥750 days

3.6.2.1 Overall Patient-year by subgroup

The overall Patient –year within each treatment arm by 1) age category (<=12 yrs vs. >12 yrs) and gender as well as 2) by race will be provided to support RMP report.

3.6.3 Concomitant and rescue therapy

Records on the Concomitant medications / significant non-drug therapies eCRF will be coded using WHO drug dictionary. All medications will be classified as *prior* and *concomitant* medication and summarized by treatment group and preferred term as follows:

- *Prior* medications are defined as drugs taken and stopped prior to first dose date of study medication and will be summarized for SAF.
- *Concomitant* medications are defined as drugs taken at least once on or between first dose date and last dose date of study medication (including those which were started prior to first dose and continued into the treatment period) and will be summarized for SAF.
- *Post-study drug discontinuation* medications will be drugs started after the discontinuation of randomized study medication and will be summarized for FUS.

Medications will be categorized into one (and only one) of above classes based on recorded or imputed start and end dates.

Incomplete or missing end date should be imputed as follows:

- If only day is missing (XXMONYEAR), it should be imputed with last day (LD) of the month (LDMONYEAR),
- If day and month are missing (XXXXXYEAR), it should be imputed with last day of December (LDDECYEAR,)
- If the whole date is missing (XXXXXXXXX), it should be imputed with last day of December of the cut-off year (LDDECCUTOFFYEAR). The cut-off year will be defined as the year of database lock.

Incomplete or missing start date should be imputed as follows.

- If the year value is missing, it will be imputed as the date one day after the treatment start date.
- If the year value is less than the treatment start date year value, it will be imputed as the midmonth date (15MONYYYY) or the midyear date (01JulYYYY) if the month is missing.
- If the year value is greater than the treatment start date year value, it will be imputed as the month start date (01MONYYYY) or the year start date (01JanYYYY) if the month is missing.
- If the year value is equal to the treatment start date year value, it will be imputed as the midmonth date (15MONYYYY) if the month is less than the treatment start month or as the start month date (01MONYYYY) if the month is greater than the treatment start month. In case the month is either missing or equal to the treatment start month, it will be imputed as the date one day after the treatment start date.

3.6.3.1 Steroid for treatment of MS relapses during the study

Corticosteroids used for the treatment of multiple sclerosis relapses that started after the initiation of study medication will be listed.

3.7 Pooling of centers

This study will be conducted in centers worldwide. In order to minimize the impact of low-enrolling countries on the analysis (such as non-convergence of the analysis model), countries are pooled into regions based on geographical proximity and regions will be used in the analysis as a stratification factor.

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, data collected at the time of study drug discontinuation or treatment completion will be recorded in the eCRF at end of treatment visit (corresponding to Visit 777 in the database). In the subject completion (CMP) data set, data collected from the study drug discontinuation eCRF (Visit 777) and study phase completion eCRF (Visit 778) will be differentiated by the visit number although there is no study visit corresponding to Visit 778.

3.8.1 Study Day 1 and other study days

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

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. Day 0 will not be used.

3.8.2 Baseline

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 the first dose date or with an assessment date on the same day of the first dose date and from Day 1 visit or before (i.e., visit\repeat visit number < 3). 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 mean of 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.

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.

3.8.3 Definition of core phase data

After completing the core phase, subjects have the option to enter the extension study. All core phase data are entered in the database for core phase. All extension data will be entered in the database for extension phase.

For subjects who do not enter the extension study, all collected data will be considered as core phase data.

For subjects who enter the extension study, any measurement or event prior to the first dose of the extension study drug that is recorded on the core phase eCRF in the core phase database. For running records (e.g., adverse events), any event starting prior to the first extension study drug that is recorded on core phase eCRF in the core phase database.

3.8.4 Nominal visits and visit windows

As a general rule for by-visit summaries, nominal visits should be used to present efficacy data (including QoL and PRO assessments) and visit windows to present safety data.

3.8.4.1 Nominal visits

Nominal visits (i.e., the visits as collected in the database) will be used for by-visit summaries of EDSS, MRI, [redacted] and ophthalmic exam (OCT). Data from unscheduled visits will not be summarized in by-visit summaries, unless otherwise specified.

For subjects who complete the study drug before the nominal 24 month visit, the EOT visit (Visit 777) will be remapped to the next scheduled visit.

For subjects who discontinue the study drug prematurely, the EOT visit (Visit 777) and the follow-up visit (Visit 501) are not timepoint specific visits and 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 monthly scheduled visit (i.e., Visits 4, 5, 6 and 7), or within 10 days of the quarterly scheduled visit (i.e., Visits 8, 9, 10, 11, 12, 13 and 14).
- 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 exists etc.) or after these visits (e.g., Visit X-1.1 or Visit X-1.2 if Visit X-1.1 already exists 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 exists etc.).

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

For subjects who discontinue study drug prematurely and have the abbreviated visits, all the abbreviated visits (if different visit numbering is used) will be mapped to the corresponding regular visits and will be treated as nominal visits in the by-visit summaries.

3.8.4.2 Visits windows

For data not listed in Section 3.8.4.1, visit-windows will be defined and used for 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 visit). Based on the study assessment schedule as specified in the study protocol, visit-windows are defined by a set of days around each nominal visit target day.

Visit-windows for vital signs, laboratory assessments, ophthalmologic examination (excluding OCT), pulmonary function tests, and ECG are provided in Table 3-4 to Table 3-9. 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 Section 3.8.6. Tables displaying summary statistics "by visit" will also use the term visit-window as column header; this is to remind the reviewer that multiple assessments of a subject might be summarized.

Table 3-4 Visit-windows for vital signs, physical development, and laboratory values (hematology/blood chemistry other than liver panel)

Visit	Start day	Target Day	End day
Week 2	1	14	22
Month 1	23	30	45
Month 2	46	60	75
Month 3	76	90	135
Month 6	136	180	225
Month 9	226	270	315
Month 12	316	360	405
Month 15	406	450	495
Month 18	496	540	585
Month 21	586	630	675
Month 24	676	720	765

Table 3-5 Visit-windows for urinalysis

Visit	Start day	Target Day	End day
Month 6	1	180	270
Month 12	271	360	540
Month 24	541	720	765

Table 3-6 Visit-windows for liver panel

Visit	Start day	Target Day	End day
Week 2	1	14	22
Month 1	23	30	37
Month 1.5	38	45	52
Month 2	53	60	75
Month 3	76	90	105
Month 4	106	120	135
Month 5	136	150	165
Month 6	166	180	225
Month 9	226	270	315
Month 12	316	360	405
Month 15	406	450	495
Month 18	496	540	585
Month 21	586	630	675
Month 24	676	720	765

Table 3-7 Visit-windows for ophthalmologic examination (excluding OCT)

Visit	Start day	Target Day	End day
Month 3	1	90	135
Month 6	136	180	225
Month 9*	226	270	315
Month 12*	316	360	450
Month 18*	451	540	630
Month 24	631	720	765

*Only for a subset of subjects.

Table 3-8 Visit-windows for pulmonary function tests

Visit	Start day	Target Day	End day
Month 1	1	30	60
Month 3	61	90	135
Month 6	136	180	270
Month 12	271	360	450
Month 18	451	540	630
Month 24	631	720	765

Table 3-9 Visit-windows for ECG

Visit	Start day	Target Day	End day
Month 1	1	30	195
Month 12	196	360	540
Month 24	541	720	765

3.8.4.3 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 post-dose (6 hours)
- Day 1 post-dose (>6 hours) (derived as 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
- 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 as recorded in the database will be used. For the second dose/dose increase/restart first and second dose monitoring vital signs, data will be summarized similarly.

Additional vital sign for extended monitoring after the 6 hour time-point will not be summarized but reported in the data listing only.

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

For safety summaries on the safety set, only core phase data (defined in 3.8.3) which are within 45 days (5 times the half-life of FTY720, 9 days) after the last dose of study drug. For safety summaries on the follow-up set, all core phase data, including assessments or events more than 45 days after the last dose of study drug in core phase, will be considered.

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

1) Day 1 to 45 after drug discontinuation
2) >Day 45 after drug discontinuation

For each of the time-window, assessments collected or events started during the specified day intervals after the last dose date will be used. The last dose date defined in Section 3.8.7 will be used to determine day windows.

3.8.6 Multiple assessments within visit windows

For visit windows defined in Sections 3.8.4.2 and 3.8.5, multiple records may exist in one particular visit window. All results (scheduled and unscheduled) will be displayed in listings, but only one value (observed or derived) "representing" the subject will be selected for summary statistics by 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 (from scheduled and unscheduled visits) 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 where multiple assessments could occur. 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/dose increase/restart dose monitoring data.

3.8.7 Last dose day of randomized study medication

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

1) The last date patient took study drug on the study phase completion eCRF,
2) The study drug discontinuation date on the study drug discontinuation eCRF, and
3) The last end date on the study drug administration record eCRF.

The last dose date recorded in the study phase completion eCRF (included in the CMP data panel with visit number 778) will be used.

4 Efficacy evaluation

Unless otherwise specified, all efficacy analyses will be based on the Full Analysis Set, following the intention-to-treat principle. Analysis of the primary efficacy variable will also be conducted on the per-protocol set, as well as the analyses on the key secondary variable. Treatment group estimates will be determined using least-squares means estimates.

All statistical models will be adjusted by the two randomization stratification factors: region and pubertal status reported in IVRS based on investigator's assessment. Pooling of regions with small patient count will be performed for better model fitting based on patient count in the region and geographic proximity. As a result, there will be three regions: East Europe, West Europe, rest of the regions (pooled by three small regions with <20 patients for each: Australia, Middle and South America, and North America). A stratification factor may also be considered to be removed from the model in presence of extremely small stratum with scarce information or given model converge problem.

4.1 Primary and key secondary variables

4.1.1 Primary variable

The primary efficacy variable is the annualized relapse rate (ARR), which is defined as the average number of confirmed relapses per year (i.e., the total number of confirmed relapses divided by the total number of days in the core phase multiplied by 365.25). The total number of days in the core phase is defined in section 4.1.2; the number of relapses will include all the confirmed relapses experienced during the corresponding observation period. The ARR can be calculated in two ways:

1. ARR (time-based)

The ARR (time-based) 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 on study of all subjects in the group and multiplied by 365.25. The time-based ARR is the primary efficacy variable.

2. ARR (subject-based)

The ARR (subject-based) is calculated as the number of confirmed relapses divided by the number of days on 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 Time in study for primary analysis

Time in study (days) (i.e., the number of days on study) for each subject will be calculated as the final study phase visit date– 1st core dose date +1.

Final study phase visit date is recorded on the study phase completion eCRF.

4.1.3 Key secondary efficacy variable

The key secondary variable is the annualized rate of new/newly enlarged T2 lesions (n/neT2).

4.2 Statistical methodology

The primary analysis pertaining to the primary efficacy objective will be based on the FAS, following the intention-to-treat principle.

4.2.1 Primary analysis

The null hypothesis for the primary analysis states that: there is no reduction in the annualized relapse rate (ARR) in the subjects treated with fingolimod compared to the ARR in subjects treated with IFN β-1a, while the alternative hypothesis states there is a reduction in the ARR in fingolimod treated subjects.

The test of the null hypothesis will be based on a negative binomial regression model with log link, using treatment, pubertal status (the stratification factor in IVRS), and region as factors and number of relapses within the previous two year before randomization as covariate. The primary analysis will be conducted on the FAS.

The response variable for this analysis is the number of confirmed relapses for each subject, and the quadratic variance estimate will be used. The natural log of (time in study in years for each subject) will be used as the offset to account for the varying lengths of subjects' time in the study. The null hypothesis will be rejected if the observed p-value is less than the significance level of 0.05.

Model-based estimates of the ARR (Least-squares mean, LSMean) and its 95% confidence intervals, ARR ratio and its 95% confidence intervals and p-value comparing fingolimod vs. IFN β-1a will be provided. The LSmean will be estimated with the OBSMARGINS option (abbreviated OM). The OM option allows the LSMEANS estimation to use the observed marginal distribution of the categorical covariates (region, pubertal status) rather than assuming equal weighting among the levels for each categorical covariate. Specification of the OM option will not affect the between-group comparison, i.e., p-value.

To control for the overall type-I error rate, a multiplicity adjustment as specified in Section 4.2.2.1 will be applied to the primary and key secondary endpoints hypothesis testing. The 95% confidence intervals will not have any further adjustment for multiplicity.

4.2.1.1 Supportive of primary analyses

Supportive analyses will be performed for the primary variable. The FAS will be used unless otherwise specified. All supportive analyses are summarized below and in Table 4-1.

1) To assess the impact of study drug discontinuation to the results observed, a supplementary analysis using the same negative binomial model as in the primary analysis will be performed on relapses on study drug (i.e., relapses will be counted only up to study drug discontinuation) in the FAS. The number of days on study drug is calculated as (the last core dose date – 1st core dose date +1).

2) A supplementary analysis using the same negative binomial model as in the primary analysis will be performed on all relapses (i.e., confirmed and unconfirmed relapses) in the FAS.

3) Analyses on confirmed relapse and on all relapse will be repeated on the PPS.

Table 4-1 Primary efficacy endpoint, primary & supportive/sensitivity analyses

Endpoint	**Analysis method**	
ARR (confirmed relapses only) **(**time-based**)**	Negative binomial model (treatment, region, pubertal status (the stratification factor in IVRS), and number of relapses within the previous two years before randomization) on FAS	**Primary analysis**
	Negative binomial model (treatment, region, pubertal status (the stratification factor in IVRS), and number of relapses within the previous two years before randomization) excluding data after drug discontinuation on FAS	Supportive analysis
	Negative binomial model (treatment, region, pubertal status (the stratification factor in IVRS), and number of relapses within the previous two years before randomization) on PPS	
ARR (confirmed and unconfirmed relapses) (time-based)	Negative binomial model (treatment, region, pubertal status (the stratification factor in IVRS), and number of relapses within the previous two years before randomization) on FAS	Supportive analysis
	Negative binomial model (treatment, region, pubertal status (the stratification factor in IVRS), and number of relapses within the previous two years before randomization) on PPS	
ARR (confirmed relapses only) (time-based)	Negative binomial model (treatment, region pubertal status (the stratification factor in IVRS), and number of relapses within the	Sensitivity analysis

	previous two year before randomization) on FAS but exclude patients in the IFN-beta arm who are neutralizing-antibody (Nabs) positive at end of study visit	

In addition, ARR (time-based) and ARR (subject-based) will be calculated and summarized by treatment group. Summaries will be provided for ARR counting confirmed relapses only and for ARR counting both confirmed and unconfirmed relapses.

4.2.1.2 Handling dropouts/censorings in ARR calculations

The primary NB model with an offset for the time in study adjusts for various treatment duration and missing information (drop-out or censoring at the time of study completion) under the assumption of non-informative drop-out, information missing at random, and constant relapse rate over time. According to the protocol, subjects who discontinue study treatment should remain in the study and follow the assessment schedule. All subjects in the FAS will be included in the primary analysis, i.e., the number of confirmed relapses observed up to the end of subject's participation in the study. Relapses will be counted regardless of whether a subject is on or off study drug.

4.2.1.3 Sensitivity analysis

To assess the impact of patients positive for NAbs on the primary efficacy outcome (i.e., confirmed relapse), the negative binomial model will be rerun on the FAS population but exclude patients in the IFN-beta arm who are Nabs-positive at the end of study visit (Table 4-1).

4.2.2 Secondary analyses

Secondary efficacy variables will be analyzed on the FAS except for the key secondary efficacy variable, annualized rate of the number of new/newly enlarged T2 lesions, which will be analyzed additionally on the PPS as supportive analyses as well as on the FAS population but exclude patients in the IFN-beta arm who are Nabs-positive at the end of study visit as a sensitivity analysis.

Secondary analyses (including both the key secondary analysis and other secondary analyses) are summarized in Table 4-2 and Table 4-4 (for MRI analysis in section 4.2.2.4) and details will be provided in the following sections. The FAS will be used unless otherwise specified. P-values for the treatment comparisons will be provided. Each inferential analysis will be performed at a significance level of 0.05 with no adjustment for multiple analyses.

Table 4-2 Key and Other secondary efficacy variables & analyses

Endpoint	**Analysis method**	
Key secondary		
Annualized rate of the number of new/newly	Negative binomial regression (treatment, region, pubertal status (the stratification factor in IVRS), and baseline T2 lesions), offset: time in	**Key secondary analysis**

enlarged T2 lesions up to Month 24	study on FAS	
	Negative binomial regression (treatment, region, pubertal status (the stratification factor in IVRS), and baseline T2 lesions), offset: time in study on PPS	Supportive analysis
	Negative binomial model (treatment, region pubertal status (the stratification factor in IVRS), and baseline T2 lesions), offset: time in study on FAS but exclude patients in the IFN-beta arm who are Nabs-positive at end of study visit	Sensitivity analysis
Other Secondary		
Time to first confirmed relapse and proportion of patients free of relapse	Log-rank test stratified by treatment group	Supportive analysis
	Cox's proportional hazard model (treatment, region, pubertal status (the stratification factor in IVRS), and number of relapses within the previous two years before randomization)	Supportive analysis
Number of T1 Gd-enhancing lesions per scan up to Month 24	Negative binomial regression (treatment, region, pubertal status and baseline T1 Gd-enhancing lesions), offset: number of scans. The response variable is the cumulative number of T1 Gd-enhancing lesions from each visit up to Month 24	Supportive analysis

4.2.2.1 Multiplicity Adjustment

The primary analysis of the primary variable and key secondary analysis of the key secondary variable will consist of 2 hypothesis testing of treatment comparisons. Multiplicity adjustment will be applied to control the Type-I error rate by conducting the hypothesis testing in the following hierarchical order [CPMP (2002)]:

1. fingolimod vs. IFN β-1a testing treatment difference for ARR;
2. fingolimod vs. IFN β-1a testing treatment difference for annualized rate of the number of new/newly enlarged T2 lesions;

Each hypothesis testing will be performed at a significant level of 0.05. However, the lower-ranked testing will be performed only when the higher-ranked testing preceding it is statistically significant at 0.05.

4.2.2.2 New/newly enlarged T2 lesions

The key secondary variable is the annualized rate of the number of new/newly enlarged T2 lesions (n/neT2) from baseline to end of the core phase, with a duration up to 24 months.

The null hypothesis for the key secondary analysis states that: there is no reduction in then/neT2 lesion rate in subjects treated with fingolimod compared to subjects treated with

IFN β-1a, while the alternative hypothesis states there is a reduction in the n/neT2 lesion rate, in fingolimod treated subjects.

4.2.2.2.1 Key secondary analysis

The test of the null hypothesis will be based on a negative binomial regression model adjusted for treatment, number of T2 lesions at baseline, pubertal status (the stratification factor in IVRS), and region. The natural log of time in study for each patient will be used as the offset.

For the key secondary analysis, there will be no imputations for missing data. The response variable for this analysis is the cumulative number of n/neT2 lesions up to Month 24 (i.e., Month 0 to EOS), for each subject. The quadratic variance estimate will be used and the natural log of time (in years) from first dose of study treatment to the last available T2 MRI assessment date for subject will be used as the offset to account for the varying lengths of each subject's time in the study.

The above null hypothesis will be rejected if the observed p-value for treatment comparison (based on the n/neT2 rate ratio) is less than the significance level of 0.05. Model-based estimates (LSmean estimated with the OM option) of the annualized rate of the number of n/neT2 lesions from baseline to end of the core phase, with a duration up to 24 months and their 95% confidence intervals, estimate of the n/neT2 rate ratio (i.e., relative treatment effect) and its 95% confidence intervals and p-value for comparing fingolimod vs. IFN β-1a i.m. will be provided.

The number of n/neT2 relative to baseline and the number of n/neT2 relative to previous visit will also be summarized (mean, standard deviation, median, minimum and maximum) by visit.

4.2.2.3 Relapses

All relapse related analyses in this section will consider confirmed relapses only unless otherwise specified and will be conducted on the FAS.

4.2.2.3.1 Time to first relapse and proportion of relapse-free subjects

Time to first relapse and proportion of relapse-free subjects are considered as other secondary endpoints related to relapses.

For subjects with at least one event (i.e., relapse), time to event is calculated as (event start date – 1st core dose date + 1). For subjects without event, time to event will be calculated the same as the time in study for ARR in Section 4.1.2.

The log-rank test of the treatment difference in the Kaplan–Meier estimates of the survival function of the time to first relapse will be performed.

Kaplan-Meier (KM) estimates of the survival functions will be constructed by treatment group. KM estimates of the relapse free patients at 12 and 24 months, together with 95% confidence intervals will be presented. Two-sided 95% confidence intervals of the difference in KM estimates will also be presented. The corresponding KM plot will be provided.

Cox's proportional hazards model adjusted for treatment, region, pubertal status (the stratification factor in IVRS), and number of relapse within the previous two year before randomization will be performed as a supportive analysis. Estimates of the hazard ratio and its 95% confidence intervals, and the p-value for treatment comparison from the Cox model will be provided.

4.2.2.3.2 Relapse characteristics

The following relapse characteristics will be summarized at relapse level: severity (mild/moderate/severe), affecting daily activities (yes/no), steroid used (yes/no), hospitalization (yes/no), recovery status (none/partial/complete), and duration of relapse (days).

The following relapse characteristics will be summarized at patient level: severity, affecting daily activities, steroid used, hospitalization, and recovery status. For summaries at patient level, a patient is counted only in the most severe category for each variable.

Summaries of the above mentioned relapse characteristics will be performed for confirmed relapses and for all relapses (confirmed and unconfirmed) as well.

The proportion of subjects hospitalized for relapse and the proportion of subjects with severe relapses will be compared between treatment groups using a Chi-square test.

The severity of confirmed MS relapses will be derived according to the criteria in Table 4-3. For severity derivation, the EDSS data obtained on the date of relapse confirmation (as collected on the Relapse eCRF page) will be used.

Table 4-3 Severity of MS relapse

Mild relapse	Moderate relapse	Severe relapse
EDSS increase of 0.5 point	EDSS increase of 1 or 2 points	Exceeding Moderate criteria
Or	Or	Or
1 point FS change in one to three systems	2-point FS change in one or two systems	Exceeding Moderate criteria
	Or	Or
	1-point change in four or more systems	Exceeding Moderate criteria

[Panitch et al 2002]

4.2.2.3.3 ARR in the first and second year

The ARR will be summarized for the following periods: 0 to 12 months and 13 to 24 months. The number of days on study for a specified period is calculated as follows:

- For 0 to 12 month summaries: (12 months cutoff date – first core dose date + 1).
- For 13 to 24 month summaries: (24 months cutoff date – 12 months cutoff date).

The 24 months cutoff date is the date corresponding to the last day of time in the core phase defined in Section 4.1.2. If a subject completes the study or discontinues the study after Month 12 visit, the 12 months cutoff date will be the Month 12 visit date. If the Month 12

visit date (from the VIS data set) is not available, the 12 months cutoff date will be date corresponding to Day 360. If a subject discontinues or completes the study prior to Month 12 visit, he or she will only be included in the 0 to 12 month summaries and the 12 months cutoff date will be the date corresponding to the last day of time in study if time in study is less than 360 days or the date corresponding to Day 360 if otherwise.

4.2.2.3.4 Cumulative ARR by time

To further investigate the possible change in ARR over time, cumulative ARRs from Month 0 to the following time points will be estimated using the same NB model as for the primary analysis:

- Month 0 to Month 3 (including data up to 90 days)
- Month 0 to Month 6 (including data up to 180 days)
- Month 0 to Month 9 (including data up to 270 days)
- Month 0 to Month 12 (including data up to 360 days)
- Month 0 to Month 15 (including data up to 450 days)
- Month 0 to Month 18 (including data up to 540 days)
- Month 0 to Month 21 (including data up to 630 days)
- Month 0 to Month 24 (including data up to 720 days)

The estimated ARR (LSmean estimated with the OM option and 95% CI) by time and treatment will be plotted to allow for the detection of any trend of ARR over time within each treatment arm.

4.2.2.4 MRI

MRI will be assessed at screening, Months 6, 12, 18, and 24 (end of study or early discontinuation) visits. Parameters collected at screening include: number and volume of T2 lesions, volume of T1 hypointense lesions, number and volume of Gd-enhancing lesions, and normalized brain volume. The following parameters will be collected at each post baseline visit:

- Number of new/newly enlarging T2 lesions (on each scheduled scan compared to previous scheduled scan and compared to the Screening scan)
- [redacted]
- Number of Gd-enhancing lesions (on each scheduled scan; optional for M18 scan)
- Volume of Gd-enhancing lesions (on each scheduled scan; optional for M18 scan)
- [redacted]

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

The following is the complete list of the MRI endpoints, including the annualized rate of the number of new/newly enlarged T2 lesions (the designated key secondary endpoint) and the secondary endpoint of the number of T1 Gd-enhancing lesions per scan up to Month 24 (i.e., month 0 to EOS), which is also discussed in detail in Section 4.2.2.2, as well as the exploratory endpoints. The time point of EOS is the last MRI scan in the core study. :

T2 lesions:

- Annualized rate of the number of new/newly enlarged T2 lesions up to Month 24 (i.e., Month 0 to EOS) and by visit (i.e., Month 0 to 6, Month 0 to 12, Month 0 to 18, and Month 0 to 24).
- Proportion of subjects free of new/newly enlarging T2 lesions up to Month 24 (Month 0 to EOS) and by visit (i.e., Month 0 to 6, Month 0 to 12, Month 0 to 18, and Month 0 to 24).
- [redacted]

Gd-enhancing lesions:

- Number of Gd-enhancing lesions per scan up to Month 24 (Month 0 to EOS) and number of Gd-enhancing lesions by visit (i.e., Month 6, Month 12, Month 18, and Month 24).
- Proportion of subjects free of Gd-enhancing lesions at EOS and by visit (i.e., Month 6, Month 12, Month 18, and Month 24).
- Volume of Gd-enhancing lesions at EOS and by visit (i.e., Month 6, Month 12, Month 18, and Month 24).

[redacted]

[redacted]

All MRI endpoints specified above will be summarized using descriptive statistics by treatment group and visit.

The proportion type endpoints will be analyzed using the logistic regression model adjusted for treatment, region, pubertal status (the stratification factor in IVRS), and other covariates as applicable. The odds ratio and its 95% confidence intervals, and the p-value for treatment comparisons from the logistic regression model will be provided.

The count type endpoints (i.e., number of lesions) will be analyzed using the negative binomial regression model adjusted for treatment, region, pubertal status (the stratification factor in IVRS), and other covariates as applicable. The LSmean estimated with the OM option and their 95% confidence intervals, estimate of the rate ratio (i.e., relative treatment effect) and its 95% confidence intervals, and the p-value for treatment comparisons from the negative binomial regression model will be provided.

All other continuous endpoints will be analyzed using a rank ANCOVA model adjusted for treatment, region, and other covariates as applicable.

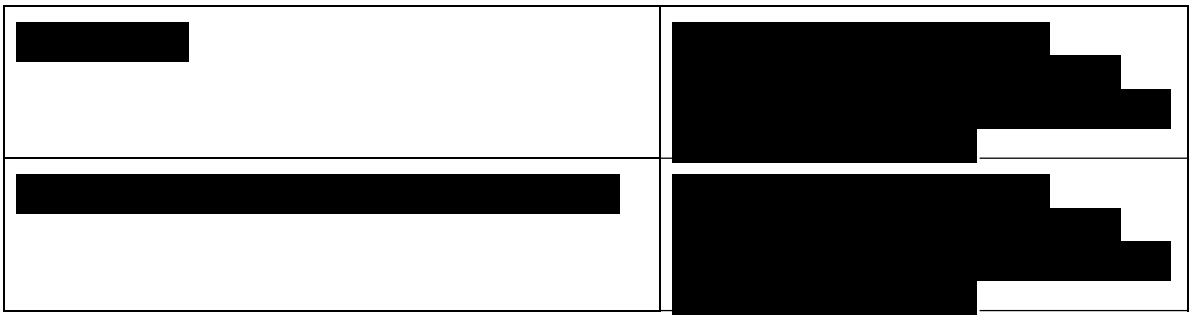
A parametric ANCOVA model with the same factors and covariates will be used for the endpoint of percent brain volume change and ARBA. The LSmean estimated with the OM option and its 95% confidence intervals, the estimate of between group difference and its 95% confidence interval, and the p-value for treatment comparisons from the ANCOVA model will be provided.

Table 4-4 lists the MRI endpoints and the analysis to be performed. Covariates to be included in the above mentioned statistical models are also specified in this table.

Table 4-4 MRI endpoints & analyses

Endpoint	Analysis model (covariates)
Annualized rate of the number of new/newly-enlarged T2 lesions by visit	As in Table 4-2. Response variable is the number of ne/newly-enlarged T2 lesions compared to baseline at the visit of interest
Number of Gd-enhancing lesions by visit	As in Table 4-2. Response variable is the number of Gd-enhancing lesions at the

	visit of interest.
[redacted]	[redacted]
[redacted]	[redacted]
Proportion of subjects free of new/newly-enlarged T2 lesions compared to baseline by visit	Logistic regression (treatment, pubertal status (the stratification factor in IVRS), and region as factors, baseline no. of T2 lesions as covariate)
Proportion of subjects free of Gd-enhancing lesions by visit	Logistic regression (treatment, pubertal status (the stratification factor in IVRS), and region and factors, baseline no. of Gd-enhancing lesions as covariate)
[redacted]	[redacted]
Volume of Gd-enhancing lesions by visit	RANK ANCOVA (treatment, region, pubertal status (the stratification factor in IVRS), and baseline volume of Gd-enhancing lesions)
[redacted]	[redacted]
[redacted]	[redacted]



4.2.2.5 EDSS

The EDSS score and its change from baseline will be summarized by treatment and visit. Only the assessment not obtained during a relapse will be considered as a baseline value.

Patient count (%) based on EDSS change categories, i.e., improvement, stable, deterioration, will be presented. The analysis will consider only those patients who were evaluated up to the assessment time point in the analysis (e.g. only patients with ≥12 months of follow-up in the 12-month assessment of EDSS, etc.). Only scheduled assessments will be considered; intermediate missing values for a scheduled visit will be considered not free of the event.

The definition of the categories are as follows, depending on the baseline EDSS score:

- If baseline EDSS score ≤5.0:
  - improvement: reduction from baseline ≥1.0;
  - stable: change from baseline between (inclusive) -0.5 to 0.5;
  - deterioration: increase from baseline ≥1.0
- If baseline EDSS score >5.0:
  - improvement: reduction from baseline ≥0.5;
  - stable: no change from baseline;
  - deterioration: increase from baseline ≥0.5

4.2.3 Efficacy subgroup analyses

Key efficacy variables data (including primary and key secondary variables, i.e., ARR (time-based) and ARR (subject-based), and annualized rate of the number of new/newly enlarged T2 lesions) will be summarized separately for the following subgroups based on demographic and baseline characteristics.

- [redacted]
- age ≤12 yrs vs. >12 yrs subgroups
- Body weight at randomization visit <=40 kg vs. >40 kg
- Female vs. male subgroups
- Hispanic/Latino vs. non-Hispanic/Latino subgroups

5 Pharmacokinetic/pharmacodynamics evaluations

5.1 Population pharmacokinetic analysis

A population pharmacokinetic analysis will be used to characterize the key pharmacokinetic parameters in this patient population. Additionally, this analysis will also determine whether there are population covariates that affect the pharmacokinetics of fingolimod-P when administered to this patient population.

The effect of covariates on pharmacokinetic parameters will be explored. Covariates examined will include age, weight, gender, and ethnicity as well as comedications such as oral ketoconazole which has been shown to interact with fingolimod-P in adults (increase in fingolimod and fingolimod-P plasma concentrations). The choice of covariates to be included in the model will be guided by exploratory plots of random effects (inter-individual variability parameters) against covariates. Those that are judged to show evidence of a relationship with the random effects will be tested for entry into the model, using the likelihood-ratio test with $p < 0.05$. The final covariate model will be derived using a rigorous and acceptable model building procedure.

The above analysis will be performed by Pharmacometrics (PMX) group. The report analysis plan and results will be provided separately.

5.2 Population pharmacokinetic/pharmacodynamics analysis

Population PK/PD modeling approaches will be used to relate the individual fingolimod PK parameter estimates to key efficacy measurements (MRI-related, relapse-related). A modeling plan will be prepared before final clinical database lock providing details for the proposed PK/PD analysis. This modeling plan will be prepared and executed by PMX group. The report analysis plan and the results will be provided separately.

6 Safety evaluation

Safety assessments include adverse events, bradycardia (first dose) events, laboratory tests, vital signs, ECG, pulmonary function tests, ophthalmic examinations, Tanner stages (bone x-ray), skin assessments, and C-SSRS data.

All safety data will be summarized on the SAF set. Some safety data (AEs, vital signs, labs, PFTs and C-SSRS) will also be summarized on the follow-up set by time-window (days 1 to 45 after study drug discontinuation and more than 45 days after study drug discontinuation) to assess subjects' safety after discontinuation of the study drug.

Unless otherwise specified, safety assessments in core phase (core phase data defined in section 3.8.3) 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.

Adverse events starting before the first dose date will be excluded from the adverse event.

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

6.1 Adverse events

6.1.1 Adverse event start date imputation

To ensure that all reported AEs are summarized, a conservative approach will be taken to impute the AE start date when it is incomplete or missing.

1. If the year value is missing, it will be imputed as the treatment start date.

2. If the year value is less than the treatment start date year value, it will be imputed as the midmonth date (15MONYYYY) or as the midyear date (01JulYYYY) if the month is missing.
3. If the year value is greater than the treatment start date year value, it will be imputed as the month start date (01MONYYYY) or as the year start date (01JanYYYY) if the month is missing.
4. If the year value is equal to the treatment start date year value, it will be imputed as the midmonth date (15MONYYYY) if the month is less than the treatment start month or as the start month date (01MONYYYY) if the month is greater than the treatment start month. In case the month value is missing or equal to the treatment start month, it will be imputed as the date one day after treatment start date.

6.1.2 Adverse events

In this section, the term adverse event (AE) will refer to any entry on the adverse events eCRFs. All AEs will be coded using MedDRA.

AEs will be summarized on the SAF set by presenting, for each treatment group, the number and percentage of patients having any adverse event by primary system organ class (SOC) and preferred term (PT). AE summary will also be presented in the pre-pubertal (Tanner staging score <2) and young subjects (<=12 years) subgroups.

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. Severe AEs, serious AEs, drug related AEs, most common AEs, AEs leading to premature discontinuation of study drug, AEs requiring study drug interruption, and AEs requiring additional therapy will also be summarized and presented in a similar format as AEs. Summaries on the pre-pubertal subgroup will also be presented.

Since the patients will have various exposure duration to the study treatment, incidence rates (IRs) of AEs, i.e., incidence of AEs per 100 patient-years, will be summarized by primary organ class and preferred term. The incidence rates of AEs that fulfill the risk search terms with RMP risk will also be summarized. The IR is calculated as the number of patients experiencing at least one event in a particular category over the total patient-years “at risk” for that event multiplied by 100. For patients who have that adverse event, the time at risk is the time from first study dose to the first onset date of that adverse event. For patients who did not have that adverse event, the time at risk is the time from first study dose to

- the minimum between i) the last dose date of study medication in the core phase +45 days and ii) the final study phase visit date for those who do not enter the extension phase
- the final study phase visit date on study phase completion eCRF (i.e., the day before the first dose date in extension) for those who enter the extension phase.

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 rates of AEs that fulfill the risk search terms with RMP risk will also 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.

AEs will also be summarized on the follow-up set for the following time periods defined by days relative to study drug discontinuation: days 1 to 45 after study drug discontinuation and more than 45 days after study drug discontinuation.

6.1.3 SAEs and DAEs

SAEs include death and non-fatal SAEs. All SAEs starting on or 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.

DAEs refer to AEs leading to permanent study drug discontinuation.

6.1.4 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 ≤ 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 ≤ 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.

6.1.5 AEs that fulfill the risk search terms by maximum severity

The Adverse events that fulfill the risk search terms (level 1 and level 2 only) for Fingolimod are summarized by the maximum severity to support RMP report.

6.1.6 Data summaries

Summaries on the SAF set by treatment group include:

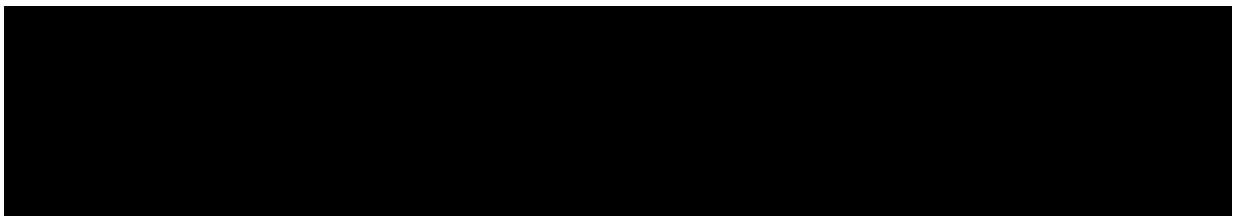
- Incidence of AEs by SOC and PT
- Incidence of AEs that fulfill the risk search terms with RMP risk by risk name and level terms
- Incidence of AEs by SOC and PT and maximum severity
- Incidence of AEs related to study drug by SOC and PT
- Incidence rate of AEs per 100 patient years by SOC and PT
- Incidence rate per 100 patient years of AEs that fulfill the risk search terms with RMP risk by risk name and level terms
- Occurrence of AEs per 100 patient-years by SOC and PT
- Occurrence per 100 patient years of AEs that fulfill the risk search terms with RMP risk by risk name and level terms
- Incidence of AEs by PT
- Incidence of deaths by SOC and PT
- Incidence of non-fatal 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 requiring study drug interruption by SOC and PT
- Incidence of AEs requiring additional therapy by SOC and PT
- Incidence of AEs leading to study drug discontinuation by SOC and PT
- Occurrence of SAEs by SOC and PT
- Occurrence of NSAEs by SOC and PT
- Incidence of AEs that fulfill the risk search terms with RMP risk by risk name, level terms (only for level 1 and level 2), and severity

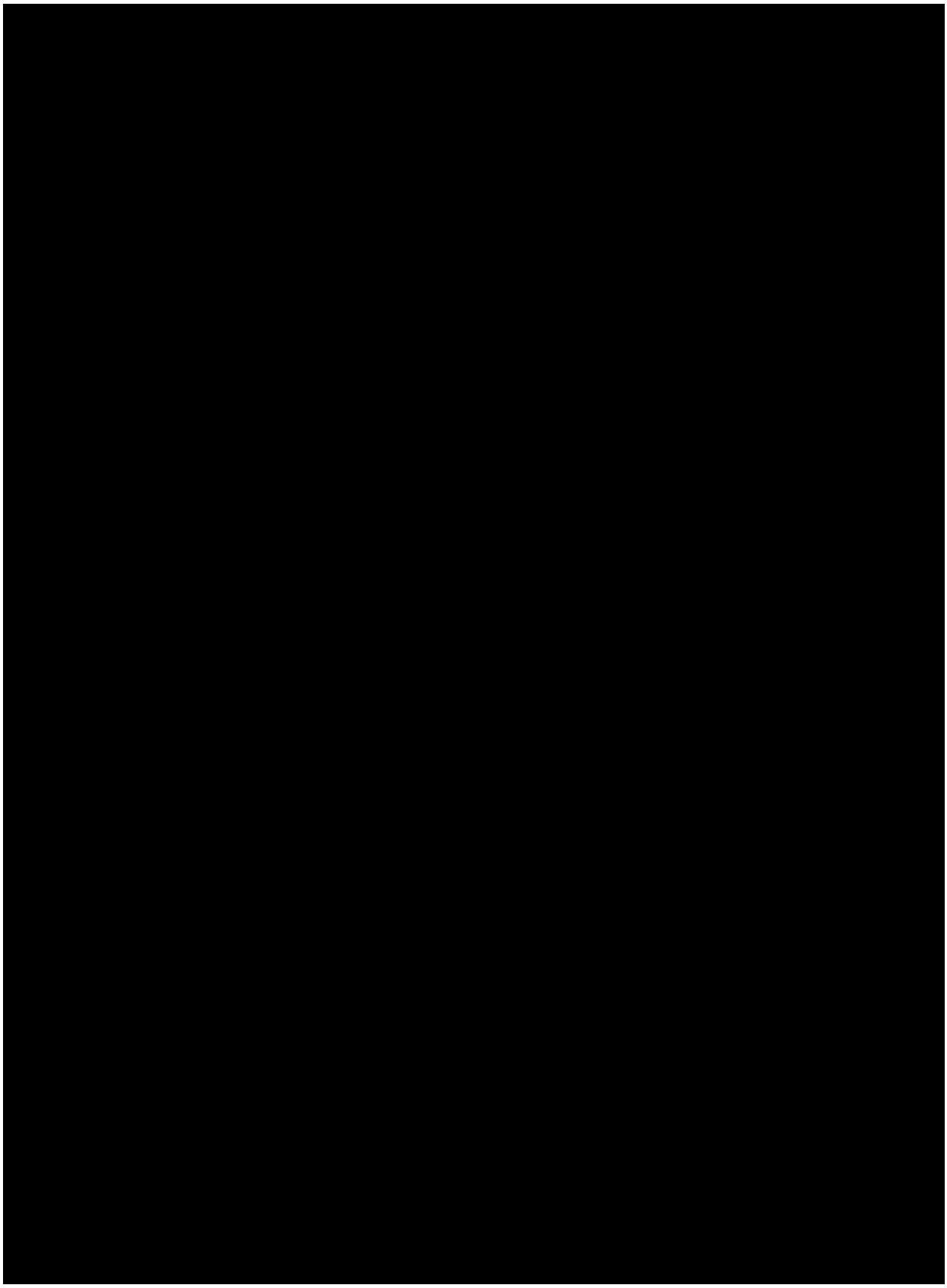
Summaries on the follow-up set by treatment group include:

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

6.2 Physical examinations

According to Novartis standard eCRF policy, physical examination data will not be entered in the database. Any clinical significant findings should be reported as AEs and summarized in the AE summaries and any clinical significant change from baseline will be recorded in the vital sign eCRF and summarized in the vital sign data summaries.





6.4 Vital signs

6.4.1 Vital signs measurements

According to the protocol, three sitting measurements of blood pressure (SBP and DBP) will be taken at each blood pressure assessment. At first dose monitoring, pulse, systolic and diastolic blood pressures all will be measured 3 times at the scheduled pre-dose time point.

For post-baseline assessments (not including Day 1), the blood pressure value will be the average of the non-missing values of the 3 measurements. If more than one blood pressure assessment (scheduled or unscheduled, except for the first dose or second dose or restart dose monitoring data) exists in a particular visit-window, the blood pressure value will be the average of all assessments.

6.4.2 Vital signs notable criteria

Clinical notable criteria in vital signs are summarized in Table 6-1.

6.4.3 Data summaries

The following summaries will exclude hourly vital sign data from the first dose or second dose or restart dose monitoring (first dose data discussed in Section 6.7.1).

Summaries on the SAF set by treatment group include:

- Summary statistics and changes from baseline in vital sign parameters by visit
- Frequency (%) distributions of highest SBP, DBP, and lowest pulse in pre-defined categories respectively (as specified in the table shell);
- Incidence of notable vital sign abnormalities based on the notable criteria in Table 6-1

Summaries on the FUS by treatment group include:

- Incidence of notable vital sign abnormalities based on the notable criteria in Table 6-1 will be summarized by time-window (days 1 to 45 after study drug discontinuation, and more than 45 days after study drug discontinuation)

Plots on the SAF set by treatment group include:

- Box and whisker plot of vital signs by visit

Table 6-1 Criteria for clinically notable vital signs

Vital Sign	Age group	Notable criteria
Heart Rate	<12 years	>130 bpm or increase of ≥15 bpm from baseline Or
		<70 bpm or decrease of ≥15 bpm from baseline
	≥12 years	>120 bpm or increase of ≥15 bpm from baseline Or
		<50 bpm or decrease of ≥15 bpm from baseline
Systolic Blood Pressure	<12 years	≥125 mmHg or increase of ≥20 mmHg from baseline Or
		≤70 mmHg or decrease of ≥20 mmHg from baseline
	≥12 to <18 years	≥160 mmHg or increase of ≥20 mmHg from baseline Or
		≤90 mmHg or decrease of ≥20 mmHg from baseline
	≥18 years	≥180 mmHg or increase of ≥20 mmHg from baseline or
		≤90 mmHg or decrease of ≥20 mmHg from baseline
Diastolic Blood Pressure	<12 years	≥85 mmHg or increase of ≥15 mmHg from baseline Or
		≤50 mmHg or decrease of ≥15 mmHg from baseline
	≥12 to <18 years	≥95 mmHg or increase of ≥15 mmHg from baseline Or
		≤50 mmHg or decrease of ≥15 mmHg from baseline
	≥18 years	≥105 mmHg or increase of ≥15 mmHg from baseline Or
		≤50 mmHg or decrease of ≥15 mmHg from baseline
Temperature		>38.3 °C or an absolute change >= 1.1 °C from baseline

6.5 Laboratory data

The summary of laboratory evaluations will be presented for three groups of laboratory tests: Hematology, Chemistry, and Urinalysis. On presenting summary statistics, laboratory data will be grouped and displayed in an alphabetical order within the test groups and subgroups if applicable.

6.5.1 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.

6.5.2 Clinically notable criteria

Clinically notable abnormalities criteria are summarized in Table 6-2.

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, subjects may have a character value of "<x.x" or "<x" instead of the actual numeric value which is below the lower limit of detection (LLOD). The 50% of the LLOD rule will be applied to transform a character value into a numerical value. This rule will be applied to all applicable laboratory values prior to checking against the notable criteria as well as computation of any statistics.

Table 6-2 Criteria for clinically notable abnormalities

Hematology	**Less Than**	**Greater Than**
Hemoglobin	10 g/dL	20 g/dL
Hematocrit	0.3	0.6
RBCs	3,300,000/mm^3	6,800,000/mm^3
WBCs	3000/mm^3	15,000/mm^3
Granulocytes (Poly, neutrophils)	1000 mm^3	12000/mm^3
Lymphocytes	200 mm^3	8000 mm^3
Platelets	100,000/mm^3	600,000/mm^3
Chemistry		
Glucose	70 mg/dL (fasting and random)	120 mg/dL (fasting) 200 mg/dL (random)
Calcium	7.5 mg/dL	11.6 mg/dL
Sodium	130 mEq/L	150 mEq/L
Potassium	3.0 mEq/L	5.2 mEq/L
Chloride	85 mEq/L	119 mEq/L
BUN	2 mg/dL	30 mg/dL
Creatinine	0.2 mg/dL	1.6 mg/dL
eGFR	>19% decrease from baseline	
Total bilirubin	0 mg/dL	1.2 mg/dL
SGOT (AST)	0 U/L	100 U/L
SGPT (ALT)	0 U/L	110 U/L
GGT	0 U/L	120 U/L
LDH	0 U/L	500 U/L
Alkaline Phosphatase*		
10 -<13 years	42 U/L (m) / 51U/L (f)	362 U/L (m) / 332 U/L (f)
13 -< 16 years	74U/L (m) / 50 U/L (f)	390 U/L (m) / 162 U/L (f)
16 -<18 years	52 U/L (m) / 47 U/L (f)	171 U/L (m) / 119 U/L (f)
18 years and above	37 U/L (m) / (f)	116 U/L (m) / (f)
Total Protein	4.0 g/dL	9.5 g/dL
Albumin	2.5 g/dL	6.0 g/dL
Uric Acid	1.5 mg/dL	10.0 mg/dL
*Ranges according to NIH Clinical Centre [http://cclnprod.cc.nih.gov/dlm/testguide.nsf; 10th May 2012]		

Urinalysis clinically notable values Parameter	Abnormality
WBC	>5
RBC	>5
Protein	+ or greater*
Glucose	+ or greater*
* Trace should proceed +, otherwise ++ or greater	

6.5.3 Renal function

For each subject, the estimated creatinine clearance values (without 24 hr urine collection) will be calculated using the Cockcroft-Gault and the abbreviated MDRD formulas separately (as specified in Table 6-3). In these calculations, the body weight is the last measurement collected on or before the day when the subject takes the laboratory test and age should also be calculated based on the time when the subject takes the laboratory test.

If the creatinine value is collected in the unit *µmol/L* (SI unit), it will be converted to *mg/dL* in order to use the formulas. The conversion is via the equation below:

- mg/dL = 88.4 µmol/L (e.g., creatinine = 2.0 mg/dL = 176.8 µmol/L).

Table 6-3 Creatinine clearance calculation

Variable	Formula
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 constant: *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.

The estimated creatinine clearance will be included as one of the laboratory parameters.

6.5.4 Data Summaries

All laboratory parameters (excluding all hematology % values except % for white cell differential values) and the estimated creatinine clearance calculated using the Cockcroft-Gault formula and the abbreviated MDRD formula will be summarized.

Summaries on the SAF set by treatment group include:

- Summary statistics and changes from baseline in lab results (hematology and biochemistry) by visit
- Summary statistics of percent of baseline in hematology results by visit
- Incidence of notable laboratory abnormalities (based on criteria in Table 6-2)
- Incidence of notable laboratory abnormalities by visit
- Frequency (%) distributions of liver function test results and hematology results in pre-defined categories (as specified in the table shells); For liver function tests, the frequencies and percentages of subjects with elevations of certain times upper limit normal will be summarized by treatment group.
- Shift table (low, normal, or high as defined by the normal ranges) from baseline to post-baseline extreme values in hematology and biochemistry results

- Shift table (positive and negative) from baseline to post baseline in Urinalysis results.
- Neutralizing antibodies (Nabs) to IFN at Month 24 by treatment

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.

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

- Frequency (%) distributions of liver function test results and hematology results in pre-defined categories (as defined in the table shell) by time-window (days 1 to 45 after study drug discontinuation and more than 45 days after study drug discontinuation)

Plots on the SAF set by treatment group include:

- Box and whisker plot of liver function tests results by visit
- Line plot of liver function test results for patients meeting liver function notable criteria
- Box and whisker plot of hematology results by visit
- Line plot of hematology results for patients meeting hematology notable criteria

6.6 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).

6.6.1 Corrected QT interval

The QT interval data will be corrected according to 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). Maximum increase in corrected QT interval from baseline will be used to summarize the frequencies of subjects who fulfill the abnormality criteria based on the corrected QT interval will be calculated.

6.6.2 Abnormality criteria for corrected QT interval

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

Table 6-4 Abnormality criteria for corrected QT interval

	Male Patients	Female Patients
1	>450 msec	>460 msec
2	>500 msec	>500 msec
3	30 - 60 msec increase from Baseline	
4	>60 msec increase from Baseline	

6.6.3 ECG findings

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

6.6.4 Data summaries

Summaries on the SAF set by treatment group include:

- Summary statistics and 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 6-4
- Frequency (%) distribution of ECG findings by ECG evaluation type by visit

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.

6.7 First dose (second dose/restart dose/dose increase) monitoring

6.7.1 Vital signs

Hourly vital signs including pulse, SBP and DBP are collected for the first dose, second dose (if necessary), dose increase (if applicable), and restarting of dose after interruption (if applicable). The bradycardia events, bradycardia symptoms, and bradycardia treatment during first and second dose monitoring will be collected and summarized as well.

6.7.2 ECG

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

6.7.3 Dose monitoring experience

The overall dose monitoring experience refers to the information collected on the dose administration monitoring eCRFs (e.g., 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 discontinue the study drug permanently, and whether SAE is reported, etc.). All data with yes/no responses will be summarized.

6.7.4 Data summaries

Summaries on the safety set by treatment group include:

- Summary statistics and changes from pre-dose in vital signs by hour (1^{st} dose, 2^{nd} dose, dose increase, and restarts)
- Frequency (%) distribution of categorized sitting pulse or change (or percent change) from pre-dose sitting pulse during first dose administration (note that categories are given in table shells)

- Summary of the overall dose monitoring experience (1st dose, 2nd dose, dose increase, and restarts)
- Incidence of bradycardia events (1st dose and 2nd dose)
- Incidence of medications used for bradycardia (1st dose and 2nd dose)
- Incidence of bradycardia symptoms by SOC, PT, and severity (1st and 2nd dose)
- Incidence of notable vital sign abnormalities based on the notable criteria in Table 6-1 (1st dose and 2nd dose)
- Summary statistics and changes from pre-dose in ECG parameters by time point during 1st and 2nd dose administration
- Incidence of abnormal QTc interval (Bazett and Fridericia) as defined by the abnormality criteria in Table 6-4 during 1st and 2nd dose administration
- Frequency (%) distribution of ECG findings by ECG evaluation type by time point during 1st and 2nd dose administration

Plots on the SAF set by treatment group include:

- Box and whisker plot of vital signs during 6 hours post first dose administration

6.8 Pulmonary function tests (PFTs)

Pulmonary function test (PFT) data will consist of the following parameters: FEV1, FVC, and D_LCO. FEV1/FVC ratio will be derived based on the FEV1 and FVC recorded in the database.

6.8.1 Predicted values and units of PFTs

For each subject, the predicted values of the PFT parameters (FEV1, FVC, FEV1/FVC, and D_LCO) will be calculated based on the formulas below.

Reference equations (Koopman M, Zanen P, et al (2010) Reference values for paediatric pulmonary function testing: the Utrecht dataset. Respiratory Medicine; 105: 15-23) to calculate the predicted values of FEV1, FVC, FEV1/FVC

- **Forced expiratory volume in one second (FEV_1) (Unit: Liter/sec):**

  Patients <18 years old:
  male: $e^{(-1.74+0.016 \cdot \text{height (cm)}+ 0.0017 \cdot \text{age (months)} + 0.036)}$
  female: $e^{(-1.74+0.016 \cdot \text{height (cm)}+ 0.0017 \cdot \text{age (months)})}$

  Patients >=18 years old:
  male: $(4.3 \cdot \text{height (m)}) – (0.029 \cdot \text{age (yr)}) – 2.49$
  female: $(3.95 \cdot \text{height (m)}) - (0.025 \cdot \text{age (yr)}) – 2.6$

- **Forced vital capacity (FVC) (Unit: Liter):**

  Patients <18 years old:
  male: $e^{(-11.10+2.37 \cdot \text{Ln (height (cm))} + 0.0016 \cdot \text{(age (months))} – 0.61+0.13 \cdot \text{Ln (height (cm))})}$
  female: $e^{(-11.10+2.37 \cdot \text{Ln(height (cm))} + 0.0016 \cdot \text{(age (months))})}$

  Patients >=18 years old:
  male: $(5.76 \cdot \text{height (m)}) – (0.026 \cdot \text{age (yr)}) – 4.34$

female: (4.43 · height (m)) – (0.026 · age (yr))-2.89

- **FEV_1/FVC (Unit: %):**
  Patients <18 years old:
  male: 982.77-179.27·Ln (height (cm)) – 177.09 · Ln(age (months)) -2.30
  + 35.57·Ln (height (cm)) · Ln(age (months))
  female: 982.77-179.27·Ln (height (cm)) – 177.09 · Ln(age (months))
  + 35.57·Ln (height (cm)) · Ln(age (months))
  Patients >=18 years old:
  male: - 0.18 · age (yr) + 87.21
  female: 0.19 · age (yr) + 89.10

Reference equations to obtain predicted values of D_LCO

- **Single breath diffusion capacity (D_LCO):**
  Patients <18 years old (Unit: mmol/min/kPa):
  Male: $e^{(34.80\text{-}6.89 \cdot \text{Ln(height (cm))} - 8.66 \cdot \text{Ln(age (months))} + 0.10 + 1.79 \cdot \text{Ln(height (cm))} \cdot \text{Ln(age (months))})}$
  Female: $e^{(34.80\text{-}6.89 \cdot \text{Ln(height (cm))} - 8.66 \cdot \text{Ln(age (months))} + 1.79 \cdot \text{Ln(height (cm))} \cdot \text{Ln(age (months))})}$
  Patients >=18 years old (Unit: mlCO/min/mmHg):
  Male: (0.3319 · height (cm)) – (0.1971 · age (yr)) – 18.006
  Female: (0.2441 · height (cm)) – (0.1436 · age (yr)) – 8.20

In the above calculations, the height of a subject is from the same visit (if not available, the most recent height measurement prior to or after the visit) when the PFT is conducted and the age should be calculated based on the date when he or she performs the PFTs.

For D_LCO, three units may be collected in the eCRFs, which are mlCO/min/torr, mlCO/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 mlCO/min/torr = 1 mlCO/min/mmHg;
- 1 mmol/min/kpa = 2.986 mlCO/min/mmHg.

6.8.2 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).

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.

6.8.3 Data summaries

Summaries on the SAF set by treatment group include:

- Summary statistics and changes from baseline in percent predicted values of PFT parameters by visit
- Number (%) of subjects who had PFT measurements below 80% of predicted by visit
- Shift tables from baseline to lowest post-baseline presenting the frequency of subjects with normal and abnormal (categories are specified in the shell)
- Frequency (%) distribution of relevant pulmonary function test changes (<60% or <80% for any visit or 2 consecutive visits) for percent predicted FEV_1, FVC, and D_LCO

Summaries on the follow-up set include:

- Number (%) of subjects who had PFT measurements below 80% of predicted by time-window (days 1 to 45 after study drug discontinuation and more than 45 days after study drug discontinuation)

The following plots will be provided on the SAF set:

- Box and whisker plot for PFT parameters (percent predicted values) by visit

6.9 Ophthalmic evaluations

The ophthalmic evaluation includes assessment of visual acuity, assessment of optical coherence tomography (OCT), [redacted] 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).

The central foveal thickness (CFT) and macular volume from OCT will be summarized by visit where nominal visits are used as they are. Other ophthalmic data will be summarized by visit where visits are defined by the visit-windows. Last assessment on study drug and unscheduled assessment may also be summarized as specified in the table shells. If multiple unscheduled values exist for a subject, the worst (i.e., highest) value will be used in the CFT and macular volume summaries and the largest (positive) percent change from baseline value will be used in the CFT and macular volume summaries. For the CFT summaries on the FUS, the same rules apply when multiple observations exist for a subject in a predefined follow-up visit-window.

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. The same rules apply for the Sloan letter chart testing data 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). The worst value for the other parameters in the Sloan letter chart testing (i.e., 100% chart total correct and 2.5% chart total correct) refers to the smallest value.

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

6.9.1 Data summaries

Summaries on the SAF set by treatment group include:

- Summary statistics of visual acuity and changes from baseline by visit
- Summary statistics of CFT and percent changes from baseline by visit
- Frequency (%) distribution of categorized visual acuity and change from baseline by visit
- Frequency (%) distribution of categorized CFT percent change from baseline by visit
- Summary of diagnosis of macular edema
- Summary statistics of macular volume and percent change from baseline by visit
- Summary statistics of Sloan letter chart testing data and changes from baseline by chart type and visit.

6.10 Chest x-ray and MRI of chest

Chest x-ray was originally scheduled for screening and Month 24 visits but was removed according to protocol Amendment 3. A listing of the chest x-ray (or chest MRI) data will be provided.

6.11 Skin assessment

A complete skin examination will be performed at screening and EOT. Results of the skin exams are categorical (normal or abnormal with abnormal types) and will be summarized by treatment group.

6.12 C-SSRS

The C-SSRS questionnaire will be administered electronically (eC-SSRS), via an interactive voice response (IVR) system. However, as per internal Novartis guidelines, a CRF page 'Supplemental Data for Suicidal Ideation and Behavior Categories' is also used for those cases when the subject did not conduct the phone interview because, e.g. the phone system was inoperable, the subject was hospitalized and unable to conduct the interview, or the subject refused to conduct the interview/withdrew from the study and external information on suicidal ideation and behavior is required. For these cases the information will be entered on the supplemental suicide CRF page by the patient or caregiver as appropriate. In addition, cases may require additional information or additional information may be available. The information will be collected from external parties (spouse, caregiver, nurse, investigator, etc.) through means of the supplemental suicide CRF page.

In the analysis, data from both sources (IVR eC-SSRS and Supplemental CRF) will be used in a pooled manner and no distinction will be made between the two. There are cases when incomplete call was made via IVR and any data from incomplete call will not be included in analysis. When data from both sources are available on the same date, both data will be analyzed.

The Columbia Suicide Severity Rating Scale (C-SSRS) data will be mapped to Columbia Classification Algorithm for Suicide assessment (C-CASA) as per FDA guidance on suicidality. Following is a mapping algorithm (code and category) that should be applied for the C-SSRS:

Table 6-5 Mapping between C-CASA and C-SSRS scales for suicidal ideation and behavior

Event code	C-CASA event	C-SSRS event
1	Suicide completed	Suicide completed
2	Suicide attempt	Actual attempt
3	Preparatory Actions Toward Imminent Suicidal Behavior	Interrupted attempt Aborted attempt Preparatory acts or behavior
4	Suicidal Ideation	*Subscale code and event* 1. Wish to be dead 2. Active suicidal thoughts (non-specific) 3. Active suicidal thoughts with method (no plan) 4. Active suicidal thoughts with intent (no plan) 5. Active suicidal thoughts with plan and intent
7	Self-Injurious Behavior Without Suicidal Intent	Self-injurious behavior, no suicidal intent

C-SSRS data used here are “yes” or “no”.
Source: [Gassmann-Mayer et al (2011)]

The proportion of subjects with completed suicide, suicide attempt, preparatory actions toward imminent suicidal behavior, suicidal ideation, and self-injurious behavior without suicidal intent as per the C-CASA scale during the study will be summarized by treatment group.

The following variables will also be summarized:

1. The proportion of subjects with any suicidal behavior engaged in during the study (answered “yes” for any of the C-SSRS suicidal behaviors during the study)
2. The proportion of subjects with suicidality (answered “yes” for any of the C-SSRS suicidal behaviors during the study and/or answered “yes” for any of the C-SSRS ideation during the study)

Similar summary will also be provided for baseline.

6.12.1 Data summaries

Summaries on the SAF set by treatment group include:

- The proportion of subjects who have completed suicide, suicide attempt, preparatory actions toward imminent suicidal behavior, suicidal ideation, and self-injurious behavior without suicidal intent during the study
- The proportion of subjects with any suicidal behavior engaged in during the study (answered "yes" for any of the C-SSRS suicidal behaviors during the study)
- The proportion of subjects with suicidality (answered "yes" for any of the C-SSRS suicidal behaviors during the study and/or answered "yes" for any of the C-SSRS ideation during the study)

Summaries on the FUS by treatment group will also be presented similarly.

7 Interim analyses

Regular bi-yearly and yearly (starting from the year of 2016) interim analysis for the Data Monitoring Committee (DMC) will be performed by Novartis designated CRO independent statistician and independent statistical programmer. Safety and, if requested by the DMC, efficacy data will be provided. The analysis plan for these safety interim analyses will be provided in a separate document.

Blinded data size re-estimation is also conducted to check the assumptions on ARR and dispersion parameter for power calculation. The interim analysis plan is provided in a separate document.

[redacted]

9 Determination of sample size

9.1 Sample size calculations

The sample size calculations are based on the primary efficacy variable, the annualized relapse rate (ARR). The sample size calculations and power considerations follow the method outlined in Keene et al, 2007 with a constant dispersion parameter k and were programmed manually using the SAS software.

Different scenarios of the anticipated treatment effect between fingolimod versus IFN β-1a, have been evaluated based on the available data on fingolimod and on IFN β-1a in the Phase III program in adults. In the TRANSFORMS study (FTY720D2302), a double-blind, one-year study, fingolimod 0.5 mg was directly compared to interferon beta-1a i.m. The ARR (0.21) in patients treated with fingolimod 0.5 mg was reduced by 52% compared to the ARR (0.43) in patient treated with INF beta-1a.

The initial assumptions for this study (based on adult patient data) was as follows: assuming a 50% relative reduction (from 0.72 to 0.36) in the relapse rate over 24 months (ARR: λ_1=0.18 for fingolimod, λ_2=0.36 for interferon beta-1a), a sample size of 95 patients per treatment group (or 190 total) would be needed for a study with a fixed follow-up of 24 months in order to provide 80% statistical power at a two-sided alpha level of 0.05 (with dispersion parameter of k=0.82). This corresponds to a statistical information I= 16.36.

The core phase of the study was amended so that when the information for the blinded data review exceeds the required value of 16.36 (i.e. I>16.36) which corresponds to the targeted power for the primary statistical test, all subjects who are ongoing in the core phase will be scheduled for the end of treatment/study visit. Any data collected during follow up but after the core phase data base lock will be reported in the extension phase CSR.

9.2 Power considerations for analysis of key secondary variable

For a 24-month study and assuming a fingolimod rate of n/neT2 lesions over 24 months of 2.3 a sample size of 95 per treatment group (or 190 total) will provide approximately 88% statistical power to detect a relative reduction of 50% (i.e., from 4.6 to 2.3) in the fingolimod group compared to the interferon beta-1a group. This sample size calculation assumes that the number of n/neT2 lesions over 24 months follows a negative binomial distribution. A negative binomial dispersion estimate of 2.0 was assigned which is within the 95% confidence interval. Using the same assumptions as above with a dispersion estimate of 2.67 will provide approximately 79% statistical power to detect a relative reduction of 50% (i.e., from 4.6 to 2.3) in the fingolimod group. These estimates were based on the FREEDOMS study (FTY720D2301).

As illustrated by the above sample size calculations, sample size requirements for the n/ne T2 endpoint to obtain 80% power are lower than those for the primary endpoint. It is therefore assumed that the power for the key secondary endpoint can be maintained at ≥80% even if the follow-up duration can be revised based on the results from the blinded data review.

10 Clinical Study Report - Appendix 16.1.9 Documentation of statistical methods

10.1 Statistical methods and analysis outputs

10.1.1 Statistical methods

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

10.1.1.1 Negative binomial regression model

The primary efficacy variable (aggregate ARR) will be analyzed using a negative binomial regression model adjusted for treatment, region, pubertal status (the stratification factor in IVRS), and number of relapses within the previous 2 years before randomization. Log (time in study in years) is the offset variable. This analysis will be performed on the full analysis set and the per-protocol set. In addition, the negative binomial regression model will be used for several supportive/sensitivity analyses for aggregate ARR excluding patients in the IFN-beta arm who are Nabs-positive at the end of the study, number of new or newly enlarged T2 lesions, number of Gd-T1 lesions. The corresponding negative binomial regression models are specified in Tables 4-1, 4-2 and 4-4.

Below is an example of the SAS codes to perform the negative binomial regression model.

```
**************************************************************************
* SAS Codes: Negative binomial regression model
* Variables in the model:
* relnum = number of relapses
* trt = treatment group code
* reg1a = region code
* pub=pubertal status
* n2 = number of relapses within the previous 2 years before randomization
* sid1a = subject ID
* lnday = log(days in study/365.25)
**************************************************************************;
proc genmod data=data1;
    class trt reg1a pub;
    model relnum=trt reg1a pub n2 / dist=nb
                                    link=log
                                    offset=lnday maxiter=500;
    lsmeans trt /cl exp om;
    estimate 'A – B' trt  1  -1 /exp;
     ods output estimates=est;
```

```
    ods output lsmeans=lst;
run;
```

10.1.1.2 Analysis of Covariance (ANCOVA)

The continuous efficacy variable (such as brain volume percent change from baseline) will be analyzed using an ANCOVA model adjusted for treatment, region, pubertal status (the stratification factor in IVRS), and baseline disease characters as appropriate.

Below is an example of the SAS codes to perform the ANCOVA model.

```
**************************************************************************
* SAS Codes: ANCOVA model
* Variables in the model:
* PBVC = percent brain volume change from baseline
* trt = treatment group code
* reg1a = region code
* Pbt = puberty status
* bl = baseline  character
* NBV = baseline Normalized brain volume at baseline
**************************************************************************;
proc mixed data=data1;
    class trt reg1a pbt;
    model PBVC = trt reg1a pbt NBV;
    lsmeans  trt /pdiff cl om;
    ods output lsmeans=lst;
run;
```

10.1.1.3 Rank Analysis of Covariance (ANCOVA) and Rank Analysis of Variance (ANOVA)

Rank Analysis of Covariance (ANCOVA)

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) to compute the ranks of the response variable and covariate in the combined group of treatment using PROC RANK; 2) to calculate the residuals from the linear regression of the response variable ranks vs. ranks of the covariates (without treatment) using PROC REG; 3) the CMH mean score statistics is used to compare the mean values of the residuals in treatment using TABLE scores (default scores) in PROC FREQ. Below is an example of the SAS codes used to perform this rank ANCOVA.

```
**************************************************************************
* SAS Codes: Rank ANCOVA model
* Variables in the model:
* VolGd =volume of Gd-enhancing lesions
* trt = treatment group code
```

```
* reg1a = region code
* Pbt = puberty status
* blvGd = baseline volume of Gd-enhancing lesions
* sid1a = subject ID
**************************************************************************;
Data eff;
    Set eff;
    If reg1a=' ' or blvGd =. or Pbt =. or VolGd =.then delete;
run;
proc rank data=eff nplus1 ties=mean out=ranks;
    by reg1a pbt;
    var blvGd VolGd;
run;
proc reg data=ranks noprint;
    by reg1a Pbt;
    model VolGd=blvGd;
    output out=residual r=resid;
run;
proc freq data=residual;
    tables reg1a*pbt*trt*resid / noprint cmh2;
          where trt in ('A','B');
    output out=rancova cmh2;
    *** SELECT P_CMHRMS;
run;
```

10.1.1.4 Logistic regression for proportion variables

The other secondary/other efficacy variables (proportion of patients free of relapses, proportion of patients free of new or newly enlarging T2 lesions, proportion of patients free of Gd-enhanced T1 lesions, proportion of patients free of new MRI activity) will be analyzed using the logistic regression model adjusted for treatment, region, and other corresponding covariates, on the full analysis set.

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

$Y_i \sim B(n_i, p_i)$, for i=1,…,n.

The logits of the unknown binomial probabilities (i.e., the logarithms of the odds) are modeled as a linear function of the Xi. $\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_k x_{k,i}$

Below is an example of the SAS codes to perform the analysis of the LOGISTIC regression.

```
**************************************************************************
* SAS Codes: LOGISTIC model
* Variables in the model:
* mrifree = proportion of patients free of Gd-enhanced T1 lesions
* reg1a = region code
```

```
* pbs=pubertal status at baseline
* blmri = baseline number of Gd-enhanced T1 lesions
*****************************************************************************;
proc logistic data=eff descending;
    class reg1a pbs trt /param=ref;
    model mrifree = reg1a pbs trt blmri;
    contrast ' A – B '  trt  1 /estimate=exp;

          ods output Contrastestimate=est;
run;
```

10.1.1.5 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, for example time to first confirmed relapse.The analyses were performed on the full analysis set.

The log-rank test statistic compares estimates of the hazard functions of the two groups at each observed event time. It is constructed by computing the observed and expected number of events in one of the groups at each observed event time and then adding these to obtain an overall summary across all time points where there is an event.

Cox proportional hazards model is to estimate the effect parameter(s) without any consideration of the hazard function assuming the proportional hazards assumption holds. The proportional hazards assumption is the assumption that effect parameters multiply hazard: for example, if taking drug X halves your hazard at time 0, it also halves your hazard at time 1, or time 0.5, or time t for any value of t. The effect parameter(s) estimated by any proportional hazards model can be reported as hazard ratios. Below is an example of the SAS codes to perform the Log-rank test with Kaplan-Meier Method and the Cox proportional hazards model.

```
**************************************************************************
* SAS Codes: LOG-RANK test with Kaplan-Meier Method
* Variables in the model:
* reltm = time to first confirmed relapse
* trt = treatment group code
* relcens = censor flag (1=censored, 0=not censored)
*****************************************************************************;
proc lifetest data=eff method=km;
    where trt in ('A','B');
    time reltm*relcens(1);
    strata trt;
run;
**************************************************************************
* SAS Codes: Cox's regression model
* Variables in the model:
* reltm = time to first confirmed relapse
* trt = treatment group code
```

```
* reg1a = region code
* n2 = number of relapses within the previous 2 years before randomization
* bledss = baseline EDSS
* relcens = censor flag (1=censored, 0=not censored)
**********************************************************************;
proc phreg data=datain nosummary;
  class trt reg1a/ order=internal parm=GLM;
  model reltm*relcens(1)=trt n2 bledss reg1a/rl ties=exact;
  contrast "A vs B" trt 1  -1/estimate=exp;

  ods output parameterestimates=pest;
  ods output contrastestimate=est;
run;
```

**Note that data set pest contains Hazard ratio estimates from model statement If parameter='trt', then SELECT probchisq hazardratio hrlowercl hruppercl for results;

**However, data set est contains Hazard ratio estimates from contrast statement (result variable names: estimate lowerlimit upperlimit probchisq);

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 patients experiencing a specific event, (2) at selected time-points, the between-treatment group difference in Kaplan-Meier estimates of the proportion of patients experiencing a specific event. Details of such confidence intervals are displayed in Table 2-1.

Table 2-1 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, $c_x = \Phi^{-1}(½ + x/200)$, where Φ 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_i^2 = estimated standard error of KM_i, i = 1,2; estimation of standard error uses Greenwood's formula, $c_x = \Phi^{-1}(½ + x/200)$, where Φ is the normal distribution function.		
Examples: for x = 95, $c_x = \Phi^{-1}(0.975) = 1.95996$; for x = 97.5, $c_x = \Phi^{-1}(0.9875) = 2.24140$			

10.1.1.6 Between-treatment group comparisons on baseline variables

For baseline comparability analysis, categorical variables will be analyzed by Cochran-Mantel-Haenszel (CMH) test stratified by region and pubertal status; continuous variables will be analyzed using ANOVA with covariates of treatment, region, and pubertal status.

```
**********************************************************************
* SAS Codes: CMH
```

```
* Variables in the model:
* trt = treatment group code
* reg1a = region code
* Pst =Pubertal status
*************************************************************************;
proc freq data=temp;
  tables reg1c*trt*(list of baseline categorical variables)/cmh;

run;


************************************************************************
* SAS Codes: ANOVA
* Variables in the model:
* trt = treatment group code
* reg1a = region code
* Pst =Pubertal status
*************************************************************************;
proc glm data=temp;
  class tgp1a reg1c pst;
  model list of baseline continuous variables =trt reg1c Pst;

run;
```