STATISTICAL ANALYSIS PLAN FOR CLINICAL STUDY REPORT AC-058B301

Multicenter, randomized, double-blind, parallel-group, active-controlled, superiority study to compare the efficacy and safety of ponesimod to teriflunomide in subjects with relapsing multiple sclerosis

Purpose of Analysis	Clinical Study Report
Investigational Drug	Ponesimod (ACT-128800)
Protocol Number	AC-058B301 / OPTIMUM
Actelion Document Number	D-19.178
Document Status/Version Number	Final Version 2.0
Date	20 June 2019

Author	[redacted] Statistician
Author	[redacted] Statistician
Reviewer	[redacted] Project Statistician
Reviewer	[redacted] Clinical Project Physician
Reviewer	[redacted], Clinical Project Scientist
Reviewer	[redacted], Safety Project Physician
Reviewer	[redacted], Clinical Trial Physician
Reviewer	[redacted], Clinical Trial Scientist
Reviewer	[redacted], Medical Writer
Reviewer	[redacted], Statistical Programmer

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND ACRONYMS

%FEV1	FEV1 expressed as % of predicted normal value
%FVC	FVC expressed as % of predicted normal value
9 HPT	9-Hole Peg Test
AE	Adverse event
AESI	Adverse event of special interest
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ARR	Annualized relapse rate
AST	Aspartate aminotransferase
BMI	Body mass index
BP	Blood pressure
bpm	Beats per minute
CDA	Confirmed disability accumulation
CFR	(US) Code of Federal Regulations
[redacted]	[redacted]
CI	Confidence interval
CL	Confidence limit(s)
CRF	Case report form
CRO	Contract research organization
CSR	Clinical study report
CUAL	Cumulative number of combined unique active lesions
DBP	Diastolic blood pressure
DL_{CO}	Diffusing capacity of the lung for carbon monoxide
DMTs	Disease modifying therapies for MS
ECG	Electrocardiogram
eCRF	Electronic case report form
eC-SSRS	Columbia-Suicide Severity Rating Scale (electronic self-rated

	version)
EDSS	Expanded Disability Status Scale
EMA	European Medicines Agency
EOS	End-of-study
EOT	End-of-treatment
EU	European Union
FAS	Full analysis set
FDA	(US) Food and Drug Administration
FEV1	Forced expiratory volume in 1 second
FSFV	First subject first visit
FSIQ-RMS	Fatigue Symptom and Impact Questionnaire-Relapsing Multiple Sclerosis
FU	Follow-up
FVC	Forced vital capacity
GCP	Good Clinical Practice
Gd+	Gadolinium enhancing
ICF	Informed Consent Form
ICH	International Council on Harmonisation
ICU	Intensive care unit
IDMC	Independent Data Monitoring Committee
INR	International normalized ratio
IRT	Interactive response technology
KM	Kaplan-Meier
LOCF	Last observation carried forward
LSLV	Last subject last visit
MACE	Major adverse cardiovascular events
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MMRM	Mixed effect model repeated measures
MRI	Magnetic resonance imaging

MS	Multiple Sclerosis
MSFC	Multiple Sclerosis Functional Composite
NAP	Not applicable
NCI	(US) National Cancer Institute
OCT	Optical coherence tomography
PASAT-3	Paced Auditory Serial Addition Test
PBH	Persistent black hole
PD	Pharmacodynamic
PFT	Pulmonary function test
[redacted]	[redacted]
PK	Pharmacokinetic
PPS	Per-protocol analysis set
PTOP	Post treatment observation period
QOL	Quality of life
QT_C	Corrected QT interval
QT_CB	QT interval corrected for heart rate using Bazett's formula
QT_CF	QT interval corrected for heart rate using Fridericia's formula
RMS	Relapsing multiple sclerosis
RRMS	Relapsing-remitting multiple sclerosis
SAF	Safety analysis set
SAP	Statistical analysis plan
SAS	Statistical analysis system
SBP	Systolic blood pressure
SCR	Screened analysis set
SD	Standard deviation
SDMT	Symbol digit modalities test
SDTM	Study Data Tabulation Model
SE	Standard error
SI	Standard international

SOC	System organ class
SOP	Standard operating procedure
SPMS	Secondary progressive multiple sclerosis with superimposed relapses
T25W	Timed 25 Foot Walk
TBIL	Total bilirubin
ULN	Upper limit of the normal range
WHO	World Health Organization
WHODRUG	WHO drug dictionary
WPAI:MS	Work Productivity Activity Impairment: Multiple Sclerosis

1 INTRODUCTION

This Statistical Analysis Plan (SAP) describes the analyses and presentation of the primary, secondary and exploratory efficacy endpoints as well as all safety, quality of life (QOL), pharmacoeconomic and pharmacodynamic (PD) endpoints for the clinical study report (CSR) of the AC-058B301 study (OPTIMUM).

Separate SAPs (not described here) are developed for analyses on:

- Pharmacokinetic (PK) endpoints
- Patient preference questionnaire
- MRI sub-study endpoints

Source data for the analyses are provided as Statistical Analysis Software (SAS®) data sets according to Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM).

2 STUDY DESIGN AND FLOW

The study is a prospective, multicenter, randomized, double-blind, active-controlled, parallel-group, Phase 3, superiority study in adult subjects with relapsing forms of MS. 1133 subjects are randomized in two groups: ponesimod 20 mg or teriflunomide 14 mg, using a 1:1 ratio, stratified by use of multiple sclerosis (MS) disease-modifying treatment in the last two years prior to randomization (yes, no) and by baseline Expanded Disability Status Scale (EDSS) score (EDSS $\leq$ 3.5, EDSS > 3.5).

The study is conducted at approximately 161 sites (randomized subjects) in 28 countries in North America, Eastern and Western Europe, and Pacific region.

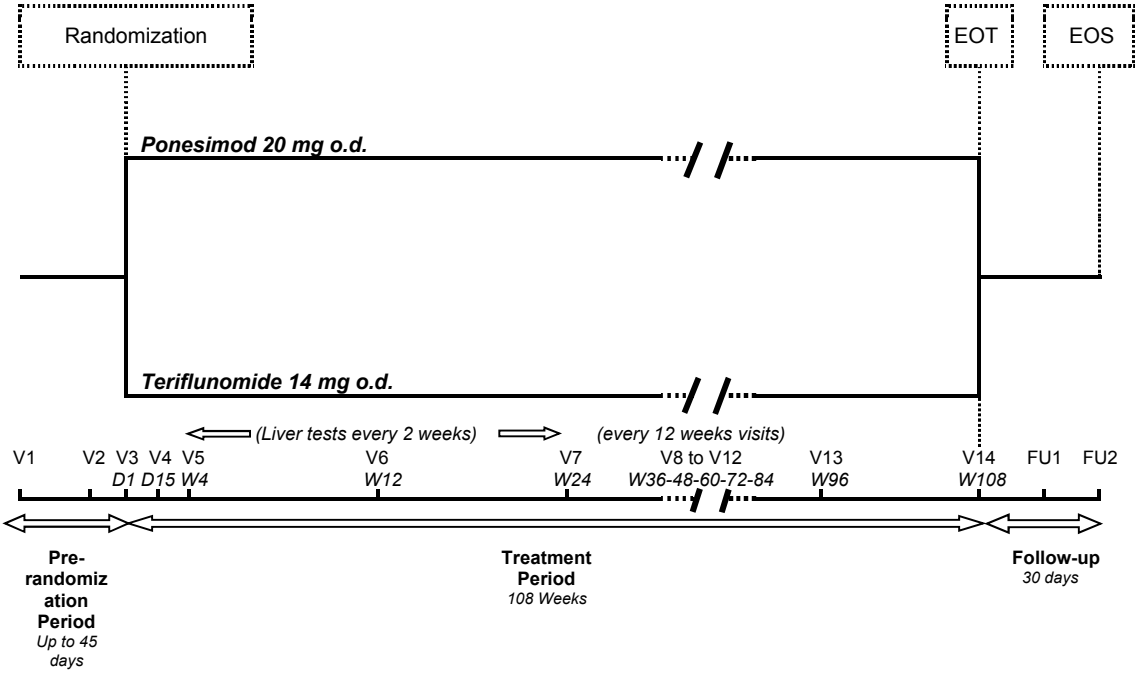
2.1 Study design

The overall design of the two-arm study is shown in Figure 1.

Subjects who have completed treatment up to Week 108 can be enrolled in a long-term open-label extension study with ponesimod, protocol AC-058B303 (OPTIMUM-LT). Data from the extension study will not be included in the analyses described in this SAP.

Subjects who prematurely discontinue treatment can remain in the study up to Week 108, in a post-treatment observation period (PTOP).

Figure 1 Study Design



D = day; EOS = End-of-Study; EOT = End-of-Treatment; FU = follow-up; M = month; o.d. = once daily; V = visit; W = week.

2.1.1 Periods

The study consists of the following protocol defined periods:

- **Pre-randomization period** starts with signature of the Informed Consent Form (ICF) and ends with subject's randomization and lasts up to 45 days (per screening attempt). Subjects can be re-screened once.
- **Treatment period** starts on the day of randomization immediately after the 1st dose of study drug intake (Visit 3–Day 1 of the study) and continues until the premature discontinuation of study drug or the scheduled End-of-Treatment (EOT) visit at Week 108. Subjects receive double-blind treatment throughout the treatment period. The treatment period consists of a titration and a maintenance phase.
  - A 2-week titration scheme is implemented on Day 1 (or at re-initiation following a treatment interruption of more than 3 days) to reduce the first-dose effect of ponesimod [see Table 1].
  - The titration phase is followed by the maintenance phase with one capsule of ponesimod 20 mg or teriflunomide 14 mg administered orally once daily in the morning (from Day 15 to End-of-Treatment [EOT]).

- **Post-treatment period** starts immediately after the last dose of study drug and ends with the End-of-Study visit. It comprises the post-treatment safety follow-up (FU) period and if applicable, the PTOP. During the post-treatment period, all subjects undergo the accelerated elimination procedure of teriflunomide lasting for 11 days with administration of cholestyramine (8 mg three times daily, or if not tolerated 4 mg three times daily) or activated charcoal (50 g two times daily), see protocol for details.

  – **Post-treatment safety follow-up period**
  For subjects who enter the extension study AC-058B303, the safety FU period ends with the FU1 visit (scheduled 14–22 days after last dose of study drug). If they have not complied with the teriflunomide accelerated elimination procedure at the time of FU1, it ends with the abbreviated FU2 visit (scheduled 23–37 days after the last dose of study drug).

  For subjects who do not enter the extension study, the safety FU period ends with FU2 visit (scheduled 30–37 days after last dose of study drug).

  – **Post-treatment observation period**
  Subjects who prematurely discontinue study treatment enter the PTOP period. It lasts from last dose of study drug until 108 weeks after randomization (i.e., end of planned treatment period) with 12-weekly visits as per the original visit schedule. Safety follow-up visits are conducted in addition, following the post-treatment safety follow-up period schedule.

Table 1 Dosing scheme

Treatment period	Duration	Dose regimen in the ponesimod group	Dose regimen in the teriflunomide group
Titration	Day 1 and 2	2 mg	14 mg
Titration	Day 3 and 4	3 mg	14 mg
Titration	Day 5 and 6	4 mg	14 mg
Titration	Day 7	5 mg	14 mg
Titration	Day 8	6 mg	14 mg
Titration	Day 9	7 mg	14 mg
Titration	Day 10	8 mg	14 mg
Titration	Day 11	9 mg	14 mg
Titration	Day 12 to 14	10 mg	14 mg
Maintenance	Day 15 until EOT*	20 mg	14 mg

*Visit 4 is to take place at Day 15 ± 1 day. The titration kit will therefore include an additional capsule and tablet (if applicable) for treatment on Day 15 (i.e., dose regimen in the ponesimod group = 10 mg). Dosing scheme during re-initiation following treatment interruption of > 3 days follows the same titration dosing scheme.
EOT = End-of-Treatment.

2.1.2 Blinding

This study is performed in a double-blind fashion. Refer to the protocol for aspects related to study drug material related blinding and functional blinding. To maintain functional blinding data are collected in two databases, i.e., a main database and a database with data considered to possibly lead to unblinding of treatment allocation (database of unblinding potential). During the conduct of the study, the Actelion trial team only has access to the main database. The main database will contain all data except data collected on the day of the first dose or on the day of re-initiation until discharge of the subject for: adverse events, hourly vital signs and electrocardiogram (ECG) measurements, and will exclude all post-baseline data for the lymphocytes and white-blood-cell laboratory parameters. Spirometry data will not be available to Actelion Biostatistics. Final analyses for the CSR described in this SAP will be presented from both main and first-dose databases combined.

2.2 Study visit and assessment schedule

Table 2–Table 3 show a schematic representation of the assessments during the study. Table 4 shows a schematic representation of the assessments during the PTOP.

Table 2 Visit and assessment schedule (Part 1)

Periods	Name	PRE-RANDOMIZATION		TREATMENT PERIOD						
	Duration	Up to 45 Days		108 Weeks						
Visits	Number	1	2	3	4	5	6	7	8–9	10
	Name	Screening	Baseline	Rand	W2	W4	W12	W24	W36 – 48	W60
	Time	Day –45 to –1		Day 1	Day 15	Week 4	Week 12	Week 24	Week 36 – 48	[redacted]
	Visit window				± 1 day	± 3 days	± 7 days	± 7 days	± 7 days	± 7 days
Informed consent*		X								
Inclusion/exclusion criteria*		X	X	X						
Demographics*, MS history & treatment*, McDonald criteria (revision 2010), Chest X-ray*, Tuberculosis test / Viral serology **		X								
Additional serum sample for viral serology			X							
Medical history / smoking status*		X	X							
EDSS/FS*		X	X				X	X	X	X
Relapse*			X	X <----------> X						
[redacted]*		X	X	X			X	X		X
FSIQ-RMS** (1), [redacted] SF-36v2**, WPAI:MS**			X				X	X		X
Health care resource utilization*					X	X	X	X	X	X
Patient preference questionnaire** (Sub-study)		X	X							
eC-SSRS**			X							X
MRI** (3)			X							X
Physical examination*		X	X			X	X	X		X
Dermatological examination*			X							X
Body weight and height* (4)		X								X
Body temperature*		X	X	X	X	X	X	X	X	X
SBP/DBP*		X	X	X(5)	X	X	X	X	X	X
12-lead ECG**		X	X	X(5)	X	X	X(6)	X	X	X
Ophthalmological examination* / OCT*		X					X	X		X
PFT incl. spirometry (All) and DL_{CO} (sub-study)**			X			X	X			X
Hematology/Chemistry** (fasted)		X	X		X	X (7) <----------> X (7)			X	X
Urinalysis		X				X	X	X		X
Pregnancy test*/**		X	X		X	X	X	X	X	X
PK sampling for ponesimod* (8)				X			X			X
Study drug dispensing & accountability */**				X	X	X	X	X	X	X
Study drug swallowing test (optional)			X							
AEs*/SAEs, Concomitant medications*		X	X	X	X	X	X	X	X	X

*Data collected in the eCRF, **Data electronically transferred to sponsor, Day 1 (date of randomization visit).
(1) FSIQ-RMS symptoms scale will be completed for 7 consecutive days. (2) No [redacted] assessment at Baseline. (3) Non-conventional [redacted]) at selected sites only. (4) Height only at Visit 1 (Screening). (5) Pre-dose and hourly for at least 4 h post-dose and up to 12 h. (6) Pre-dose and 3 h post-dose. (7) In addition, ALT, AST, INR, alkaline phosphatase and total bilirubin at Weeks 6, 8, 10, 14, 16, 18, 20, and, 22; total white blood cell and total lymphocyte counts at Weeks 8, 16, and 20.
[redacted]harmacokinetic sampling pre-dose at Weeks 12, 60 and 108, and 3 h post-dose on Day 1 and Week 12.

Table 3 Visit and assessment schedule (Part 2)

Periods	Name	TREATMENT PERIOD				FOLLOW-UP		UNSCHEDULED			
	Duration	108 Weeks				30 Days					
Visits	Number	11	12	13	14	15	16	R1, R2, …	U1, U2, …	I1, I2, …	
										d1	d15
	Name	W72	W84	W96	EOT	FU1	FU2	Relapse	Unscheduled (1)	Re-initiation	
										Day Re-1	Day Re-15
	Time	Week 72	Week 84	Week 96	Week 108 or at premature discontinuation	Last study drug intake +15 days	Last study drug intake +30 days	Any day between Day 1 and EOS			
	Visit window	± 7 days	± 7 days	± 7 days	± 7 days	−1 day, +7days	+7 days (22)	+7 days	NA	NA	± 1 day
EDSS/FS*		X	X	X	X		X	X	X		
Relapse*		X<----------X				X	X	X	X		
[redacted]			[redacted]		[redacted]						
FSIQ-RMS**[redacted]			X		X			X	X		
SF-36v2**			X		X			X			
Health care resource utilization*		X	X	X	X			X			
WPAI:MS**			X		X						
[redacted] (Sub-study)						X					
Chest X-ray*					X						
MRI** (2)					X				X		
eC-SSRS**					X						
Physical examination*			X		X			X	X		
Dermatological examination*, Body weight *					X				X		
SBP/DBP*, 12-lead ECG**		X	X	X	X	X	X		X	X(3)	X
Pulse rate*								X	X(4)		
Ophthalmological examination / OCT*					X				X		
PFT incl. spirometry (All) and DL_{CO} (sub-study)**					X	X	X		X		
Hematology/Chemistry**(fasted)		X	X	X	X	X	X		X		
Urinalysis			X		X	X			X		
Viral serology, Serum sample vaccination*									X		
Pregnancy test*/**		X	X	X	X	X	X		X		
PK sampling for ponesimod*					X				X (5)		
Teriflunomide plasma concentration									X (4)		
Accelerated elimination procedure					X	X (4)					
Accelerated elimination procedure compliance						X	X (4)				
Study drug dispensing/accountability		X	X	X	X			X	X	X	X
AEs*/SAEs, Concomitant medications*, Body temperature*		X	X	X	X	X	X	X	X	X	X

*Data collected in the eCRF, **Electronically transferred to sponsor, Day 1 = date of randomization visit,
(1) Unscheduled visits may include all or some of the indicated assessments. (2) Brain MRI to be performed at any time an opportunistic infection in the CNS is suspected. Non-conventional [redacted] at selected sites only. Note: MRI assessment at EOT does not need to be performed if the EOT visit occurs within < 4 weeks of the MRI assessment at Visit 10. (3) Pre-dose and hourly for at least 4 h post-dose and up to 12 h. (4) If needed (see protocol for details). (5) When possible, collect PK sample upon experiencing SAE.

Table 4 Visit and assessment schedule (Part 3)

Periods	Name	POST-TREATMENT OBSERVATION PERIOD (PTOP) (to be performed after EOT)		
	Duration	Up to 108 Weeks		
Visits	Number	6A, 8A, 9A, 11A,13A	7A, 12A	10A, 14A
	Name	W12A, W36A, W48A, W72A, W96A	W24A, W84A	W60A, W108A
	Time	Week 12-36-48-72-96	Week 24 and 84	Week 60 and 108
	Visit window	±7 days	±7 days	±7 days
EDSS/FS*		X	X	X
Relapse*		X<----------X		
MRI*				X
FSIQ-RMS** (1), [redacted] **, [redacted] *		X (only Week12)	X	X
Concomitant medications*		X	X	X
Physical examination*				X
Dermatological examination*				X
SBP/DBP*		X	X	X
12-lead ECG **				X
Body temperature*		X	X	X
PFT incl. spirometry (All) and DL_{CO} (sub-study)**				X (Week 108 only)
Hematology/Chemistry** (fasted)			X	X
Urinalysis*			X	X
Pregnancy test*/**		X	X	X
AEs/SAEs *		X	X	X

*Data collected in the eCRF, **Electronically transferred to sponsor, Day 1 = randomization visit. If the first PTOP visit windows overlaps with FU1 or FU2 visits, visits and assessments can be combined.
AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; [redacted]; CNS = central nervous system; d = day(s); DBP = diastolic blood pressure; DLco = Diffusing capacity for carbon monoxide; [redacted] ECG = electrocardiogram; eCRF = electronic case report form; eC-SSRS = Columbia Suicide Severity Rating Scale (electronic self-rated version); EDSS = Expanded Disability Status Scale; EOS = End of Study; EOT = End of Trial; FS = functional system; FSIQ-RMS = Fatigue Symptom and Impact Questionnaire-Relapsing Multiple Sclerosis; FU = Follow-up; h = hour(s); INR = international normalized ratio; MRI = magnetic resonance imaging; MS = multiple sclerosis; MTR = magnetization transfer ratio; OCT = Optical coherence tomography; PFT = pulmonary function test; [redacted] severity; PK = pharmacokinetic(s); SAE = serious adverse event; SBP = systolic blood pressure; [redacted] test; W = week(s); WPAI:MS = Work Productivity Activity Impairment: Multiple Sclerosis.

3 OBJECTIVES

3.1 Primary objective(s)

The primary objective of the study is to determine whether ponesimod is more efficacious than teriflunomide in terms of reducing relapses in subjects with Relapsing Multiple Sclerosis (RMS).

3.2 Secondary objectives

- To assess the effect of ponesimod on disability accumulation and on other aspects of MS disease control;
- To assess the safety and tolerability of ponesimod in subjects with RMS.

4 CHANGES OR CLARIFICATIONS TO ANALYSES PLANNED IN THE STUDY PROTOCOL

4.1 Changes to the analyses planned in the study protocol

Sensitivity analyses of the primary endpoints related to early study withdrawal as specified in section 11.3.2.2.2 of the protocol have been updated to account for recent advancements in statistical methodology by applying multiple imputation (MI) techniques for negative binomial count data [see Section 9.6.4.3].

The assumptions for the multiple imputation approaches for imputation of relapses between EOS and Week 108 following premature study withdrawal are generally in line with those specified in the protocol.

An overview of the newly introduced multiple imputation approaches is provided below:

Protocol approach	**SAP approach**
Impute with the observed individual treatment ARR	'Missing at Random' approach: ARR after withdrawal for each treatment arm is assumed to be the ARR in the respective arm before withdrawal. Equivalent to protocol specified approach.
Impute with the observed teriflunomide ARR for both treatment arms	Reference-based MI approaches: The rate in the reference (teriflunomide) arm is used for imputation in both the ponesimod and teriflunomide arms. Two different approaches, 'Copy reference' and 'Jump to reference', are applied, see Section 9.6.4.3 for details. Similar to protocol specified approach.

Impute with the observed ARR for ponesimod and impute zero relapses for teriflunomide	'Delta adjustment' MI approach: A range of deltas (deviation from the MAR assumption based on multiplicative factors δ for ARR) covering plausible to extreme and unlikely scenarios are considered for both treatment arms. The scenario with δ = 0.1 for teriflunomide (impute based on 0.1 × teriflunomide ARR under MAR) and δ = 1 for ponesimod (impute based on ponesimod ARR under MAR) represents a scenario that is similar, but slightly less extreme than the protocol specified approach. The equivalent delta adjustment scenario to the protocol approach (δ = 0 for teriflunomide) is not considered to avoid numerical instabilities of the multiple imputation algorithm.

For the imputation analyses after EOT + 7 days (see sensitivity analyses in Section 9.6.5.1) the period for imputing relapses is from EOT + 7 days to Week 108 in the SAP instead of from EOT + 7 days to EOS as specified in section 11.3.2.4 of the protocol. Rationale for the change is to align with the per protocol planned period for observing relapses at least up to Week 108 and to handle intercurrent events of premature treatment discontinuation irrespective of subsequent study withdrawal. Purpose of the analysis remains to assess the impact of treatment discontinuation and/or switching to alternative MS treatments as specified in the protocol.

4.2 Changes in the conduct of the study / data collection

A total of 6 protocol amendments were conducted. The main reasons are described below.

Protocol version 2, 29 April 2015: Addition of a sub-study assessing patient preferences for different outcomes in the treatment of multiple sclerosis.

Protocol version 3, 16 July 2015: Among other changes, the following updates were made to address comments received after Voluntary Harmonization Procedure review for this Clinical Trial Application in the EU: Introduction of the electronic self-rated version of the Columbia-Suicide Severity Rating Scale (eC-SSRS), total white blood cell and total lymphocyte count monitoring every four weeks up to Week 24; Reducing the time window for the EDSS assessments to ensure that EDSS assessments of two subsequent visits are at least 10 weeks and no more than 14 weeks apart.

Protocol version 4, 5 February 2016: Addressing a comment received from the US FDA regarding the assessment of relapses. For this purpose, a standardized step-wise procedure for the early detection, confirmation and reporting of relapses has been introduced. This included incorporating a relapse assessment questionnaire into the study (based on telephone calls and visits).

Protocol version 5, 14 November 2016 Main reason for the amendment was to modify the procedure for teriflunomide plasma concentration testing after the subject's discontinuation from study treatment. Further small clarifications and updates were made.

Protocol version 6, 30 August 2017 Main reason for the amendment was to allow for testing of teriflunomide plasma concentration in any subject who has discontinued study drug if deemed necessary for the subject's safety, at the discretion of the investigator. Timing of the testing of teriflunomide plasma concentration remains dependent on study drug discontinuation and compliance with the accelerated elimination procedure. Further small clarifications and updates were made.

Protocol version 7, 5 Dec 2018 Changes include updating the testing strategy for the secondary endpoints from a hierarchical testing strategy to the described fallback type strategy. [redacted] are demoted from secondary to exploratory endpoints. Time to 24-week disability has been added as a secondary endpoint following CHMP feedback.

4.3 Clarifications concerning endpoint definitions and related variables or statistical methods

The protocol endpoints 'Change from baseline to EOS vs change from baseline to EOT' for lymphocyte counts, pulmonary function tests (PFTs) and DL_{CO} parameters analyses for assessing reversibility are clarified as follows: 'Change from baseline to EOS' is clarified to not include assessments from the post-treatment observation period, but only from the safety follow-up period (up to EOT + 37 days). Consequently, in this SAP the terminology 'change from baseline to last follow-up assessment' instead of 'change from baseline to EOS' is used. The rationale for this is to allow the assessment of the reversibility of the effects on these endpoints prior to the initiation of other DMTs in the PTOP.

5 DEFINITIONS OF VARIABLES

This section provides the definitions and sources for all variables used in the analyses, including specifications for the derivations.

General recurrent definitions (e.g., study treatment start date, baseline, and EOS date) or unit conversions are described in Section 11.

5.1 Subject disposition

5.1.1 Screened subjects

A subject is considered screened if the subject has signed any informed consent (for the main study or any of the sub-studies) and is assigned a subject number by the interactive response technology (IRT) provider. A screened subject is considered re-screened if a second screening visit date is documented in the electronic Case Report Form (eCRF).

5.1.2 Screening failures

A subject is considered a screening failure, if screened but not randomized into the study (as per the recording in the IRT system). Subjects screened more than once and subsequently randomized are not counted as screening failures.

The reason for not being randomized is documented on the randomization eCRF:

- Failure to meet randomization criteria ("Subject eligible as per inclusion/exclusion criteria?" answered "No")

  OR

- "Subject withdrew consent" or "Other" (Reason provided following "Was the subject randomized?" answered with "No")

If a subject at a screening attempt is not randomized but no reason for not being randomized is reported, the reason is categorized as 'Unknown'. Two variables are derived to hold reason(s) for not being randomized:

- Reason for not being randomized at first attempt (not applicable [NAP], if subject randomized at first attempt; reason derived from first attempt, irrespective of existence and outcome of a second attempt)
- Reason for screening failure (NAP, if subject randomized; reason derived from first attempt if no second attempt, or from second attempt otherwise)

5.1.3 Subjects randomized

A subject is considered randomized if a randomization date and number are recorded in the IRT system.

5.1.4 Subjects treated

A subject is considered treated if they received as least one dose of study drug as documented in any study drug log eCRFs (see Section 5.3.1 for eCRFs involved).

5.1.5 Subject study treatment completion status

Subjects who complete study treatment as per protocol are those with a record in the eCRF 'Maintenance Ponesimod/Teriflunomide' form with reason for treatment end given as 'Completed as per protocol'.

A subject is considered to have prematurely discontinued from the study treatment if:

- at least one reason ('Study treatment stopped due to') is reported on the eCRF form 'Premature Discontinuation of Study Treatment' and/or
- the reason for treatment end is documented as 'Premature Permanent Discontinuation' on the eCRF 'Up-titration Ponesimod/Placebo', 'Up-titration Teriflunomide/Placebo' or 'Maintenance Ponesimod/Teriflunomide'

The date of treatment discontinuation is the last date study drug was taken [defined in Section 5.3.1]. In case of a (partially) missing discontinuation date, the date is imputed with the EOS date as defined in Section 5.1.6, or with the upper limit of the partial date if prior to EOS.

Reasons for premature discontinuation from study treatment are documented on the 'Premature Discontinuation of Study Treatment' eCRF with the following possible answers: 'Death', 'Lost to follow-up', 'Pre-specified study treatment discontinuation criteria', 'Subject decision' (further split into 'Tolerability related', 'Efficacy related', 'Other', or 'Not known'), 'Physician decision' (further split into 'Adverse event', 'Lack of efficacy / Treatment failure' or 'Other') or 'Sponsor decision' (further split into 'Study termination by sponsor' or 'Other'). Whenever 'Other' is selected, free text can be provided in addition. Both recorded levels for treatment discontinuation (e.g., 'Subject decision' as well as 'Tolerability related'), if pre-specified, are to be included in analyses, while any free text is not to be reported in tables, only listed. An additional category called 'Reason not provided' is defined for subjects where the reason is missing.

Subjects who have prematurely discontinued treatment are considered to have stayed in the study beyond safety follow-up if their EOS date is after EOT + 30 days. These subjects will be considered to have entered the PTOP beyond the safety follow-up period.

5.1.6 Subject study completion status

A subject is considered to have completed the study as per the corresponding entry on the 'End of Study' eCRF ("Did the subject complete the study?" answered with "Yes"). All subjects not completing the study are considered to have prematurely discontinued the study.

Note: A subject who prematurely discontinued study treatment but completed the PTOP is considered to have completed the study according to the protocol.

The End-of-Study (EOS) date is defined as the 'Date of End of Study' collected on the end of study eCRF, unless a subject is lost to follow-up or dies, in which case 'Date of Last Contact' from the end-of-study eCRF or 'Date of Death' from Death eCRF are used. If missing, the last recorded visit date (i.e., the latest assessment date within any visit) is considered as the EOS date.

For subjects completing the 108 weeks of treatment and entering the extension study, the EOS visit should occur at the FU1 visit (scheduled 14–22 days after last drug intake), or at the abbreviated FU2 visit if available (scheduled 23–37 days after last drug intake). For subjects completing the 108 weeks of treatment and not entering the extension study, the EOS visit corresponds to the FU2 visit (scheduled 30–37 days after last drug intake). For subjects prematurely discontinuing study treatment, EOS is scheduled to be the Week 108 visit (Visit 14A – W108A – Week 108).

For subjects prematurely discontinuing the study, reasons for premature discontinuation from study are: 'Death', 'Lost to follow-up', 'Subject decision' (further split into 'Withdrawal of consent', 'Tolerability related', 'Efficacy related', and 'Other'), 'Physician decision' (further split into 'Adverse event' and 'Other') or 'Sponsor decision' (further split into 'Study terminated' and 'Other'). Whenever 'Other' is selected, free text can be provided in addition. An additional category 'Reason not provided' is defined for subjects where the reason is missing.

5.1.7 Time in study and time on treatment

Time in study (years) is derived as (EOS date − randomization date + 1 day) / 365.25.

Time in study separated into prior and after protocol version 4 is derived as:

- Time in study after protocol version 4 (years) = (EOS date − maximum of (randomization date, start date under protocol version 4) + 1 day) / 365.25
- Time in study prior to protocol version 4 (years) = Time on study (years) − Time on study after protocol version 4 (years). It is zero for subjects with no data collected under protocol version < 4.

Time on treatment is derived as (EOT date − study treatment start date + 1 day) / 365.25.

5.1.8 Disposition by protocol version

As part of protocol version 4, systematic relapse assessment questionnaires at phone interviews and visits have been introduced, which may affect data collection of the primary endpoint. For each subject, it is therefore established whether they were randomized or had data collected under protocol version 1-3 or under protocol version 4 or higher.

- **Randomized under protocol version (1–3, ≥ 4).** A subject is considered randomized under protocol version ≥ 4 if randomized on or after the start date under protocol version ≥ 4, otherwise they are considered randomized under protocol version 1–3. The **start date under protocol version 4** is the earliest date of informed consent for the main study on any protocol version > 3 (as documented on the 'Informed Consent' eCRF).
- **Data collected under protocol version 1–3** (yes/no). A subject has data collected under protocol version 1–3 if they have been randomized under protocol version 1–3.
- **Data collected under protocol version ≥ 4** (yes/no). A subject has data collected under protocol version ≥ 4 if their EOS date is after their start date under protocol version ≥ 4.

Note: Subjects can have data collected under multiple versions and can have data collected according to both definitions above.

5.2 Subject characteristics

5.2.1 Demographics

The following demographic variables are derived:

- Sex, from the 'Demographics' eCRF (Male, Female)
- Age (years), from the 'Demographics' eCRF; if available, 'Age at re-screening' overwrites the initially entered 'Age'
- Age categories derived from the above as follows: < 18, 18–30, 31–40, 41–55, and ≥ 56 years. Age high-level categories: < 40, and ≥ 40. Age categories as per EudraCT requirement: < 12, 12–17, 18–64, 65–84, and ≥ 85
- Height (cm), from the 'Height & Body Weight' eCRF
- Weight (kg), from the 'Height & Body Weight' or from the 'Body Weight' eCRF:

  The latest measurement available prior to study treatment start date is selected. If first intake is not documented, use the latest measurement available prior to randomization. If neither date is available (for screen failures), the latest weight assessment available is used.
- BMI (kg/m^2) derived as Weight (kg) / (Height (cm) / 100)^2
- BMI categories: < 18.5, ≥ 18.5 – < 25, ≥ 25 – < 30, and ≥ 30
- Race, from the 'Demographics' eCRF (Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, White, Other, and Not applicable)
- Ethnicity, from the 'Demographics' eCRF (Hispanic or Latino, Not Hispanic or Latino, and Unknown)
- Country based on site mapping performed on SDTM level
- Geographical region of site, countries are assigned to the following regions:
  - European Union (EU): Bulgaria, Croatia, Czech Republic, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Poland, Portugal, Romania, Spain, Sweden, United Kingdom
  - Europe Non-EU + Russia: Belarus, Bosnia and Herzegovina, Georgia, Russian Federation, Serbia, Ukraine
  - North America: Canada, United States
  - Rest of World: Israel, Mexico, Turkey

5.2.2 Baseline disease characteristics

Baseline characteristics variables are defined as follows:

Stratification variables from IRT

- **Baseline expanded disability status scale (EDSS) (from IRT)** (≤ 3.5, > 3.5) from the IRT system
- **MS disease modifying treatment within 2 years prior to randomization (from IRT)** (Yes/No), from IRT system

Baseline disease characteristics variables from eCRF / non-IRT external data

- **Baseline EDSS score,** defined as the latest available EDSS score as entered in the EDSS/FS eCRF (i.e., not re-derived based on sub-scores) prior to or on the date of randomization
- **Baseline EDSS (from eCRF),** categorized as (≤ 3.5, > 3.5), derived from the baseline EDSS score
- **MS disease modifying treatment within 2 years prior to randomization (from eCRF)** (Yes/No), each documented Disease modifying therapies for MS (DMTs), as specified in Section 5.2.5.4, will be flagged as follows:

  'Taken in previous 2 years', if:
  - Medication end date* is after/on randomization date – 730 days and prior to randomization date; or
  - Medication end date* is after/on randomization date – 730 days and start date* is not after date of randomization; or
  - Medication start date* is after/on randomization date – 730 days and prior to randomization date; or
  - Ongoing at end of study? is ticked "Yes" and start date* is available and not after date of randomization; or
  - Ongoing at start of Treatment? is ticked "Yes", unless end date* is prior to randomization date – 730 days.

  'Not taken in previous 2 years', if
  - Medication end date* is prior to randomization date – 730 days; or
  - Medication start date* of the medication is after the randomization date.

  'Undetermined whether taken in previous 2 years', if
  - Medication end date is partial and only part of its range^ is within the period between randomization date – 730 days and randomization date; or
  - Medication end date is missing but question "Ongoing at start of Treatment?" is not ticked "Yes" and either start date* is prior to randomization date – 730 days, or start date is missing.

  *Or in case of partial dates if the entire range^ of possible dates (both lower and upper limit) falls within the specified period.

^ Range of partial dates is the time interval defined by lower and upper limit of the partial date.

A subject is considered as having had **MS disease modifying treatment within 2 years prior to randomization (from eCRF)**, if they either have at least one DMT flagged 'Taken in previous 2 years', or if they have any DMT flagged as 'Undetermined whether taken in previous 2 years' and the site classified the subject for 'MS treatment within 2 years (from IRT)' as 'Yes'. All other subjects are not considered to have had DMT in the last two years prior to randomization.

Any prior MS disease modifying treatment (Yes/No), any DMT prior to randomization, i.e., classified as previous, or ongoing at study treatment start (derived as specified in Section 5.2.5, or ongoing at randomization for non-treated randomized subjects), unless the start date is between randomization date + 1 and study treatment start date.

- **Time since first symptoms (years)**, defined as [Date of randomization – Date of first MS symptoms + 1] in days / 365.25. Taken from 'MS History Disease Characteristics' eCRF at screening. Partial dates of first MS symptoms are imputed to the 1st day of the month (if the day is missing) and to the 1st of January (if the day and month are missing). No imputation is performed if the date is completely missing.
- **Time since initial diagnosis (years)**, defined as: [Date of randomization – Date of initial diagnosis + 1] in days / 365.25. Taken from 'MS History Disease Characteristics' eCRF at screening. Partial dates of initial diagnosis are imputed to the maximum of [Date of first MS symptoms, 1st day of the month] (if the day is missing) and to the maximum of [Date of first MS symptoms, 1st of January] (if the day and month are missing). No imputation is performed if the date is completely missing.
- **MS subtype** (RRMS / SPMS), relapsing-remitting multiple sclerosis [RRMS], secondary progressive multiple sclerosis [SPMS] with superimposed relapses). Taken from 'MS History Disease Characteristics' eCRF at screening; subtype at re-screening overwrites the initially entered subtype.
- **Time since most recent relapse (months) at screening**, defined as [Date of screening – Onset date of latest MS relapse prior to screening + 1] in days / 30.4375. From eCRF form 'MS History - Relapse' at screening. Take the most recent "Onset date of previous relapse" among all previous relapses entered on the 'MS History - Relapse' eCRF. Partial onset dates are imputed to the maximum of [Date of first MS symptoms, 1st day of the month] (if the day is missing) and to the maximum of [Date of first MS symptoms, 1st of January] (if the day and month are missing). Data collected at re-screening overwrites the initially entered data.

- **Number of relapses in the year prior to study entry**, as reported in the corresponding question on the 'MS History - Relapse' eCRF (i.e., not based on recorded onset dates of previous relapses). Data collected at re-screening overwrites the initially entered data.
- **Number of relapses in the 2 years prior to study entry** as reported in the corresponding question on the 'MS History - Relapse' eCRF (i.e., not based on recorded onset dates of previous relapses). Data collected at re-screening overwrites the initially entered data.
- **Presence of Gd+ T1 lesions (yes/no)** on baseline MRI scan as provided by the MRI central reading**.**
- **Number of documented Gd+ T1 lesions** on baseline MRI scan as provided by the MRI central reading**.**
- **Volume of T2 lesions** on baseline MRI scan as provided by the MRI central reading.
- **Number of T2 lesions** on baseline MRI scan as provided by the MRI central reading, categorized into < 9 and ≥ 9.
- **Highly active disease (yes, no).** A subject is considered to have highly active disease if one or both of the following conditions are met:
  - Any disease modifying therapy for MS as specified in Section 5.2.5.4 received during the year* before randomization and one or both of the following conditions are met:
    - ≥ 1 relapse within 1 year prior to study entry and the baseline MRI read centrally shows either ≥ 1 Gd+ T1 lesion and/or ≥ 9 T2 lesions
    - Number of relapses within 1 year prior to study entry $\geq$ number of relapses between 2 and 1 year prior to study entry, for subjects with at least one relapse within 2 years prior to study entry.
  - ≥ 2 relapses within the 1 year prior to study entry and baseline EDSS score > 2 and baseline MRI read centrally shows ≥ 1 Gd+ T1 lesion.

  * i.e., any intake documented between randomization date – 365 days and randomization date; (partially) missing medication start and stop dates are imputed as described in Section 12.1.

  A subject is considered not to have a highly active disease if either condition stated above is not met unless either condition cannot be assessed due to missing data.
- **Smoking status** (Current smoker, Former smoker, Never smoked), taken from 'Smoking status' eCRF.

Note: *For re-screened subjects, information for baseline characteristics 'at screening' is taken from the last Re-screening Visit, if available, instead of the Screening Visit.*

Handling of missing/partial dates: In case of partial dates for required dates (MS symptoms, initial diagnosis, latest MS relapse prior to Screening), the lower limit is used in derivations. Missing dates are not replaced; the corresponding variable is considered missing.

5.2.3 Other baseline characteristics

Other baseline characteristics are not derived but reported as recorded on the eCRF in listings. It includes information collected on:

- 'Smoking status' eCRF (Number of cigarettes a day, Number of years smoked, Year quit, in addition to smoking status baseline characteristics)

5.2.4 Medical history

Medical history includes previous and/or concomitant diseases or diagnoses recorded in the Medical History eCRF. Reported terms are coded using MedDRA (version 21.0). Previous medical history are those diseases or diagnoses with an end date before or equal to the study treatment start date. In the event of a partial end date that overlaps with first dose (e.g., May 2013 and first dose occurred on 23 May 2013) or if an end date is missing, the medical history is considered as previous if ongoing is ticked 'No'. All other medical history terms are considered concomitant. For subjects randomized but not treated the randomization date is used instead of the study treatment start date.

Data originating from the eCRF 'MS History - Complication or Symptoms Associated with MS' are flagged to be summarized separately. No start and stop dates are collected on this page, however, an ongoing (no/yes) at (re-)screening flag is available. If this flag is missing, the complication/symptoms is considered concomitant.

5.2.5 Previous and concomitant therapies

Therapies are collected in the following CRF pages: 'Previous Medications', 'Concomitant Medications', 'MS Specific Treatment History' (List of pre-specified MS specific treatments if ticked as 'Yes') automatically leading to a corresponding record on 'MS Specific Treatment History Log', 'Study-treatment-concomitant Therapy - Disease-modifying Treatment for MS', 'Corticosteroids for Treatment of Relapse'.

Terms are coded using the WHO drug code dictionary and the anatomic therapeutic chemical (ATC) class code (version dated 1 March 2018 or later).

Handling of partial or missing start and end dates is detailed in Section 12.1.

5.2.5.1 Previous therapies

Previous therapies are therapies that were started and stopped prior to study treatment start date.

5.2.5.2 Study concomitant therapies: taken between study treatment start and EOS

Includes all therapies that have been taken after study treatment start date up to EOS. This includes therapies ongoing at study start, as well as therapies starting after study treatment start date, or after study treatment end date.

5.2.5.3 Treatment concomitant therapies: taken between study treatment start and EOT

Includes all therapies that have been taken after study treatment start date up to EOT (last drug intake). This includes therapies ongoing at study start, as well as therapies starting after study treatment start date.

In addition, the following are flagged in the datasets:

- *Therapies starting during study treatment administration:* Derived as treatment concomitant therapies starting on or after study treatment start.
- *Therapies ongoing at study treatment start:* Includes all treatment concomitant therapies that started prior to study treatment start, i.e. have not been started during study treatment administration.
- *Therapies taken between study treatment start and EOT + 15 days*: Includes all study concomitant therapies with start date prior to EOT + 15 days. It reflects the therapies taken during the period defining treatment-emergent used for safety reporting.

Handling of partial or missing start and end dates is detailed in Section 12.1.

5.2.5.4 Therapies starting after EOT

Includes all study concomitant therapies that have been started after EOT (last drug intake). This includes therapies with start date from EOT + 1 day onwards.

Handling of partial or missing start and end dates is detailed in Section 12.1. Therapies with missing start date will be considered to have started on treatment.

5.2.5.5 Disease modifying therapies for MS (DMTs)

Disease modifying therapies are defined as therapies which can favorably alter the course of the disease by reducing the rate and severity of relapses or delaying disease progression by preventing accumulation of disability. For thisstatistical analyses, DMTs will be identified on an ingredient level if containing any of the text indicated in Table 5 below. Medications in ATC class S (sensory organs, i.e., ophthalmologicals and/or otologicals) will not be considered.

In addition, any other potential DMT, i.e. medications with preferred drug name “INVESTIGATIONAL DRUG” or “BLINDED THERAPY” for Multiple sclerosis (recorded either on the ‘Study-treatment-concomitant therapy - Disease-modifying’ or ‘MS Specific Treatment History Log’ eCRF) will also be taken into account.

Table 5 Disease modifying therapies for MS

ALEMTUZUMAB	MITOXANTRONE
AZATHIOPRINE	MYCOPHENOLIC ACID
CERALIFIMOD	NATALIZUMAB
CICLOSPORIN	OCRELIZUMAB
CLADRIBINE	OFATUMUMAB
CYCLOPHOSPHAMIDE	OZANIMOD
DACLIZUMAB	PEGINTERFERON BETA-1A
DIMETHYL FUMARATE	PLOVAMER
FINGOLIMOD	PLOVAMER ACETATE
GLATIRAMER ACETATE	RITUXIMAB
INTERFERON BETA-1A	SECUKINUMAB
INTERFERON BETA-1B	TERIFLUNOMIDE
LAQUINIMOD	SIPONIMOD
METHOTREXATE	IMMUNOGLOBULINS NOS
INTERFERON ALFA	ETRASIMOD

A text search in ingredients will be conducted: Include all medications containing any of the text above to ensure salts (e.g., "Fingolimod Hydrochloride") and combination therapies are included.

Start of alternative DMTs for MS (required for efficacy analysis):

For each subject receiving study concomitant DMTs the start date is defined as the first start date of any study concomitant DMT.

Imputation of partial start date is detailed in Section 12.1.1. Missing start dates will not be imputed, consequently therapies with missing start date will not be considered in efficacy analyses up to start of alternative DMTs for MS.

Note on previous DMTs: Occurrence of previous DMTs (stopped prior to ICF consent) is assessed systematically (Yes/No tick boxes) in the 'MS Specific Treatment History' eCRF for selected medications. However, for reporting purposes, medications are identified via coding as specified above. On the 'MS Specific Treatment History Log' the 'reason for discontinuation' (free text) is collected and will be included in listings of MS specific therapies.

5.2.5.6 Corticosteroids for Treatment of Relapse

Corticosteroids for treatment of relapses are medications with preferred drug name from Table 6* recorded on the corresponding eCRF.

Doses of corticosteroids will be converted to prednisone equivalent doses in mg/day and are calculated as described in Table 6 with cumulative dose derived.

Accumulated steroid dose (prednisone equivalent doses in mg) for treatment of relapse from study treatment start (i) up to EOS, and (ii) up to EOT + 7 days are derived for each subject. It is the sum of cumulative steroid dose (mg) across all medications taken in the respective period. Cumulative steroid dose mg per medication is derived as dose mg/day × duration (days). Subjects without steroid use for treatment of relapses have an accumulated dose of 0 mg.

Table 6 Conversion into prednisone-equivalent dose

Corticosteroid	Dose equivalent to 1 mg prednisone	Conversion factor into prednisone equivalent
Prednisone	1 mg	1/1
Cortisone	5 mg	1/5
Hydrocortisone	4 mg	1/4
Prednisolone	1 mg	1/1
Triamcinolone	0.8 mg	1/0.8
Methylprednisolone	0.8 mg	1/0.8
Dexamethasone	0.15 mg	1/0.15
Betamethasone	0.12 mg	1/0.12

*Corresponding salts, i.e., preferred drug name starting with the listed terms are considered as well and corresponding dose is converted in the same way.
Source: GlobalRPh.

5.2.5.7 Beta blocking agents

Beta blocking agents are defined as medications with WHO Drug ATC level 2 code C07: "BETA BLOCKING AGENTS". It presents a sub-class of ATC class C 'cardiovascular system'.

5.2.5.8 Therapies used for accelerated elimination

Defined as medications to be used for the protocol mandated accelerated elimination procedure for teriflunomide following treatment discontinuation, i.e., preferred drug names "COLESTYRAMINE" or "CHARCOAL, ACTIVATED".

5.2.6 Other subject characteristics

5.2.6.1 Chest X-Ray

Collected at screening and EOT visit in the Chest X-Ray eCRF. No definitions and derivations (other than treatment day of assessment) necessary. Any resulting findings are recorded as adverse event (or medical history).

5.2.6.2 Reproductive system findings

Data will be taken from the eCRF. No definitions and derivations (other than treatment day of assessment) necessary.

Includes all information mapped to RP SDTM domain (reproductive system findings) including Childbearing potential / Fertility from Demographics eCRF.

5.3 Study treatment exposure and compliance

Exposure and compliance information will be taken from the Study Drug Log eCRFs:

- **Up-titration Summary eCRF** (collects information on up-titrations conducted as planned as per protocol, recording start and end date of up-titration);
- **Up-titration Ponesimod/Placebo eCRF** (collects information on up-titration if not conducted as planned as per protocol, separately for up-titration tablets on daily basis);
- **Up-titration Teriflunomide/Placebo eCRF** (collects information on up-titration if not conducted as planned as per protocol, separately for up-titration capsules on daily basis);
- **Maintenance Ponesimod/Teriflunomide eCRF** (collects information on maintenance study drug intake based on maintenance capsules).

To derive the study treatment start date and time, in addition information collected on Study drug administration eCRFs ('Study Drug Administration - Randomization/Day 1', 'Study Drug Administration', 'Study Drug Administration - Re- Initiation' eCRF) is considered.

In addition, subjects filled an electronic study treatment diary (eDiary). Sites were instructed to review the entries in the subject study treatment diary versus the protocol-mandated drug intake regimen and versus the number of tablets and capsules dispensed and returned. Any discrepancy was to be clarified during the site visits, and actual drug intake was to be recorded in the eCRF, hence the information from the eCRF is considered to be the most reliable, and all exposure derivations including derivation of EOT date are only based on eCRF data.

5.3.1 Exposure

Study treatment start date is defined as the earliest study drug start date documented in the study drug administration or study drug log eCRF pages, as listed above.

The time of study treatment start (time of first exposure to treatment) is taken from the study drug administration eCRF, where the date is equal to the study treatment start date, if available.

The EOT date is defined as the latest study treatment end date, as recorded on the Study Drug Log (SDL) eCRFs listed above.

The following exposure variables are derived:

Duration of Study Treatment (days) = EOT date - Study treatment start date + 1 day. Duration of Study Treatment is also derived in weeks (days divided by 7), months (days divided by 30.4375), and years (days divided by 365.25).

Duration of Study Treatment is categorized as follows: ≤ 24 weeks (168 days), > 24 weeks to ≤ 60 weeks (420 days), > 60 weeks to ≤ 84 weeks (588 days), > 84 weeks to ≤ 108 weeks (756 days), > 108 weeks (756 days).

'Study treatment exposure, interruptions excluded' equals the 'Number of days with intake documented' for a subject overall as defined as part of the compliance derivations described in Section 5.3.2.1.

5.3.2 Compliance with study treatment

5.3.2.1 Compliance as percentage of days on study drug

Compliance is assessed as the percentage of days from the date of first intake until EOT with study drug intake documented in the eCRF:

$$\frac{\text{no. of days with intake documented}}{(\text{date of EOT-study treatment start date}) + 1\text{ day}} \times 100$$

Note: Compliance below 100% must not be indicative of a deviation from the protocol since the protocol mandates dose interruptions for certain very specific safety scenarios as well as based on investigator's judgment.

For reporting, the calculated compliance is categorized as follows: 100%, 90% – < 100%, 80% –< 90%, 50% – < 80%, > 0% – < 50%, 0%.

The number of days with study drug intake documented is derived as follows:

During up-titration period

The number of days with intake documented is based on entries on the eCRF 'Up-titration Summary' and/or 'Up-titration Ponesimod/Placebo' or 'Up-titration Teriflunomide/ Placebo', with separate entries for tablets (Ponesimod/Placebo) and capsules (Teriflunomide/ Placebo).

Only the intake of the dose form with active drug is considered.

For subjects where the investigator confirmed the up-titration has been conducted as planned, the 'Up-titration Summary' eCRF is filled, collecting only start and end date of the entire titration period without collecting 'number of capsules taken' or 'number of tablets taken'. In that case the 'Number of days with intake documented' is derived and considered to be 'titration end date − titration start date + 1 day'.

For all other subjects, information is collected on the 'Up-titration Ponesimod/Placebo' or 'Up-titration Teriflunomide/Placebo' eCRF, 'Treatment start date' resulting in potentially

multiple records of up-titration drug intake (collecting in addition 'number of capsules taken' or 'number of tablets taken'). In that case 'Number of days with intake documented' for each record is derived in the same way as during the maintenance period but based on the dose form with active drug only.

During maintenance period

As per protocol, one capsule of study drug is to be taken per day. On the 'Maintenance Ponesimod/Teriflunomide' eCRF log form, each record collects 'Treatment start date', 'Treatment end date', and 'Number of capsules taken'. If, for a record, the 'Number of capsules taken' is ≤ 'Treatment end date' − 'Treatment start date' + 1, it is assumed that the 'Number of capsules taken' represents the 'Number of days with intake documented' during that record's period of study drug intake. If, for a record, the 'Number of capsules taken' is > 'Treatment end date' − 'Treatment start date' + 1, 'Number of days with intake documented' is set to 'Treatment end date' − 'Treatment start date' + 1 for the compliance calculation. These subjects are flagged as having taken > 1 capsule per day on at least one occasion.

The 'Number of days with intake documented' for a subject overall is derived as the sum across all study drug log records (sum across up-titration, or maintenance records; multiple periods of study drug intake during up-titration or maintenance are possible, e.g., due to interruptions, or detailed up-titration eCRF filled).

5.3.2.2 Other compliance variables

Compliance to up-titration regimen based on eCRF

For each episode of up-titration, a subject is considered as **not having complied with the up-titration scheme** if either of the following is documented in the study drug log eCRFs for the kit containing active drug:

- Premature permanent study drug discontinuation (documented as reason for treatment end in at least one instance in the up-titration eCRF drug log);
- Capsule or Tablet (whatever corresponds to the active drug) is taken in an incorrect sequence during the up-titration (documented as reason for treatment end 'Up-titration scheme not followed (other than interruption)' in at least one instance in the up-titration eCRF drug log);
- 1 or more tablets or capsules are missed during the up-titration ('temporarily interrupted' documented as reason for treatment end at least one instance in the up-titration eCRF drug log);
- More than one tablet or capsule taken on any day during the up-titration;
- Up-titration lasted < 13 days based on duration of up-titration period (see definition below);
- Up-titration lasted > 15 days based on duration of up-titration period (see definition below).

Duration of up-titration is derived as last titration end date - first titration start date for each up-titration period.

Other compliance variables during maintenance period

The following deviations from the protocol-intended study drug maintenance are identified from the 'Maintenance Ponesimod/Teriflunomide' eCRF log form and flagged on a subject basis:

- More than 1 capsule of study drug taken per day on at least one occasion (derivation see above);
- At least one occasion of treatment interruption of more than 3 days documented, but maintenance treatment resumed or started without up-titration. This includes occasions when maintenance treatment started more than 4 days after the day up-titrations stopped.

5.3.3 Study treatment adjustments or interruptions

Study-specific criteria for interruption and discontinuation of study treatment were pre-specified in the protocol [section 5.1.13]. The protocol did not allow any study treatment dose adjustments.

Study treatment interruptions are assessed via the compliance variables described in Section 5.3.2 and the assessment adverse events leading to changes in study treatment as described in Section 5.5.5.

Subjects with at least one interruption: Derived as subjects with 'Study treatment exposure, interruptions excluded' < 'Duration of Study Treatment', based on the definitions in Section 5.3.1.

Total duration of interruption(s) (days) is derived as 'Duration of Study Treatment'−'Study treatment exposure, interruptions excluded'.

Interruptions with reported reason are those where the reported reason for treatment end in the study drug log is 'Temporarily interrupted due to an AE', or 'Temporarily interrupted not due to an AE'.

5.4 Efficacy variables

5.4.1 Primary efficacy variable(s): ARR up to EOS

The primary endpoint is the Annualized Relapse Rate (ARR) up to EOS, based on confirmed relapses according to the treating neurologist / principal investigator. ARR is defined as the number of confirmed relapses per subject-year.

For subjects completing the study as planned the EOS is scheduled to occur at least 108 weeks (± 7 days) after randomization.

More precisely, for subjects prematurely discontinuing treatment the EOS is scheduled to occur 108 weeks (± 7 days) after randomization. For subjects completing the 108-week treatment period the EOS is scheduled to occur prior to entry into the open-label extension (14–22 days after EOT, or 23–37 days after EOT in exceptional cases) or after the safety follow-up period for subjects not entering the open-label extension (30–37 days after EOT).

A confirmed relapse is defined as a record with 'Relapse meeting the criteria for a confirmed relapse?' answered 'Yes' on the 'Relapse Summary' eCRF. All confirmed relapses from randomization up to EOS will be included in the analysis.

For the statistical analysis of ARR, the following data will be used:

- Number of confirmed relapses from date of randomization up to EOS date;
- Length of observation expressed in years, defined as: [EOS date − date of randomization + 1] in days, divided by 365.25.

The logarithm of the length of observation up to EOS will be used as an offset variable in the primary efficacy and sensitivity analyses.

Confirmed relapses 'from date of randomization up to EOS date' includes all confirmed relapses unless starting prior to randomization.

5.4.1.1 General derivation details for relapses

A relapse (confirmed or unconfirmed) is identified from the 'Relapse Summary' eCRF with 'Did the subject experience a relapse' ticked as 'Yes'. For each relapse the corresponding start and end date/time, as well as outcome, treatment with corticosteroids, action taken with study drug and need for hospitalization, and whether an EDSS was completed are recorded. For relapses with completed EDSS it is collected whether it qualifies as a confirmed relapse (see definition details for primary efficacy variable above). A relapse which is not a confirmed relapse is considered an 'unconfirmed relapse'.

Relapses 'from date of randomization up to EOS date' include all relapses unless starting prior to randomization. Additional flags will be added for relapses from date of randomization up to EOT + 7 days (relapse start date in respective period), relapses after EOT + 7 days (start date from EOT + 8 days).

Note: Relapses with missing start date are considered to have started on randomization date for analysis. Relapses with partial start date are considered to have started at the lower limit (but earliest on randomization date).

5.4.1.2 Variables for supplementary analyses of primary endpoint

According to the protocol, several supportive/sensitivity analyses are planned for the primary efficacy endpoint, including the ARR analysis from randomization up to EOT + 7 days, with EOT date defined as date of last drug intake [see Section 11.1]. Those and other additional variables are defined below. Analyses that do not require additional variable derivations will not be mentioned in this section but will be described in the 'Statistical Analysis' section, e.g., the ARR up to EOS analysis with treatment as the only covariate uses the same variable as the primary efficacy analysis.

5.4.1.2.1 ARR up to EOT+7 days for confirmed relapses

The variable considers confirmed relapses up to EOT + 7:

- The subject's number of confirmed relapses from date of randomization up to date of EOT + 7 days;
- Length of observation up to EOT + 7 expressed in years, defined as: [Min (EOT date + 7 days, EOS date) – date of randomization + 1] in days, divided by 365.25.

5.4.1.2.2 ARR up to EOS for all relapses

The variable considers all relapses (confirmed and unconfirmed) up to EOS:

- The subject's number of all relapses from date of randomization up to date of EOS;
- Length of observation expressed in years, defined as: [EOS date − date of randomization + 1] in days, divided by 365.25.

5.4.1.2.3 ARR up to EOT+7 days for all relapses

The variable considers all relapses up to EOT + 7:

- The subject's number of all relapses from date of randomization up to date of EOT + 7 days
- Length of observation up to EOT + 7 days, see above.

5.4.1.2.4 ARR up to EOS or per protocol exclusion for confirmed relapses

ARR up to EOS for the per protocol analysis is derived in the same way as the variable for the main analysis but including only confirmed relapses prior to first occurrence of a deviation leading to exclusion from per protocol [see Section 7.1.3]. The considered observation time is derived as for the main analysis but replacing EOS date with the per protocol exclusion date, if available.

Similar variables are derived for ARR up to EOT + 7 days for the per protocol analysis.

5.4.1.2.5 ARR up to the start of alternative DMTs for MS

This variable considers confirmed relapses up to the point where an alternative disease modifying treatments (DMT) for MS is started as follows:

- The subject's number of confirmed relapses from date of randomization up to the day before the start of new DMT for MS or up to EOS date for subjects not starting a new DMT;
- Length of observation expressed in years, defined as: [Min (Day before start date of new DMT for MS, date of EOS) − date of randomization + 1] in days, divided by 365.25.

Start of alternative DMTs for MS is defined in Section 5.2.5.4 and based on study concomitant MS specific therapies. Variables will be derived separately for confirmed and all relapses.

5.4.1.2.6 ARR up to EOS for imputed confirmed relapses

Defined in the same way as number of confirmed relapses for the primary analysis unless the confirmation status is missing (e.g., due to the EDSS not being completed). Relapses with missing confirmation status are imputed as confirmed relapses. The length of observation is defined as for the primary analysis.

5.4.1.2.7 ARR post EOT + 7 days following premature treatment discontinuation

Defined for subjects prematurely discontinuing treatment with EOS date after EOT + 7 days:

- The subject's number of confirmed relapses from EOT + 8 days to EOS;
- Length of observation expressed in years, defined as: [(EOS date − EOT date + 8) + 1] in days, divided by 365.25.

5.4.1.2.8 ARR by period

ARR is derived per period. Variables are defined separately for confirmed relapses and for all relapses. The following periods are considered:

ARR up to Week 60

Derived in the same way as for the ARR up to EOS but including only confirmed/all relapses with onset date from randomization date up to Day 420. Observation time is the time from randomization until earliest of Day 420 or EOS.

ARR after Week 60

Derived in the same way as for the ARR up to EOS but excluding confirmed/all relapses with onset date prior to or on Day 420.

ARR by 12-week periods

Similarly, ARR for approximately 3-month intervals are defined: Baseline to Week 12 (Day 1–84), Week 13 to Week 24 (Day 85–168), Week 25 to Week 36 (Day 169–252), Week 37 to Week 48 (Day 253–336), Week 49 to Week 60 (Day 337–420), Week 61 to

Week 72 (Day 421–504), Week 73 to Week 84 (Day 505–588), Week 85 to Week 96 (Day 589–672), after Week 96 (≥ Day 673).

Note: Relapses with missing start date are considered to have started on randomization date for period assignation. Relapses with partial start date are considered to have started at the lower limit (but earliest on randomization date).

5.4.1.2.9 Number of new or worsening neurological symptoms up to EOS

The variable is defined as the number of new or worsening neurological symptoms reported on the relapse assessment questionnaire ('Is/was the subject having new neurological symptom(s) or an acute worsening of pre-existing neurological symptom(s)?' answered with 'Yes') from randomization up to EOS.

Each new or worsening neurological symptom is further sub-classified into the following categories: confirmed relapse, unconfirmed relapse, and symptoms not considered a relapse.

Derivation details

Information is taken from the relapse assessment questionnaire eCRF pages ('Phone interview', or 'Relapse Assessment Questionnaire' either 'Visit Interview' or 'Unscheduled Visit Interview'), which have been introduced with protocol version 4.

A new or worsening neurological symptom is linked to a relapse via the onset date ('Date of start of symptoms' corresponding to the new or worsening neurological symptom is identical to the relapse onset date of a relapse reported on the Relapse Summary eCRF). If a symptom cannot be linked to a relapse it is classified as 'Symptoms not considered a relapse'. Symptoms linked to a relapse are further categorized into confirmed and unconfirmed relapses, as per Section 5.4.1.1.

Note that for subjects randomized under protocol versions 1–3 neurological symptoms are only expected to be recorded from the start date of protocol version 4 onwards.

5.4.1.2.10 Variables for further analyses related to relapses

Duration of relapse (days) is derived as relapse end date − relapse start date + 1 days.

Relapse symptoms linked to a relapse. The following categories of symptoms are considered: Visual (optic) functions, Brainstem functions, Pyramidal functions, Cerebellar functions, Sensory functions, Bowel and Bladder functions, Cerebral functions, Ambulation, Other relapse symptom taken from the relapse symptoms eCRF.

The **number of functional systems affected** (number of categories, see above) per relapse is derived.

5.4.2 Secondary efficacy variables

5.4.2.1 Change from baseline to Week 108 in fatigue-related symptoms as measured by the FSIQ–RMS symptoms score

The Fatigue Symptom and Impact Questionnaire-Relapsing Multiple Sclerosis (FSIQ-RMS) is a validated patient reported outcome instrument from which a fatigue symptoms score and sub-domain impact scores can be obtained, where higher scores indicate greater fatigue or impact [Protocol appendix 9]. The questionnaire consists of 2 parts: section 1 is a daily diary and is answered during 7 consecutive days. Section 2 is a weekly diary and is answered on the 7th day of the 7 consecutive days where the daily diaries were filled in. The questionnaire is completed using an electronic device that automatically tracks the completion date. Data are transferred electronically to the sponsor.

FSIQ-RMS is completed during the pre-randomization period (i.e., at Visit 2: 7 consecutive days preceding the randomization [Baseline visit]), and during the post-randomization period, i.e., at scheduled Visits 6, 7, 10, and 12 (Weeks 12, 24, 60, 84, respectively), 14 (Week 108), as well at premature EOT visit and at unscheduled visits (due to relapses or other unscheduled visits). If applicable, it is also completed at the corresponding visits in the PTOP. For more details refer to Table 2–Table 4.

FSIQ-RMS **symptoms domain (FSIQ-RMS-S)** consists of 7-items assessing fatigue related symptoms collected as a daily diary (section 1) on an 11-point numeric rating scale (0–10) with a 24-hour recall period.

The **FSIQ-RMS-S score** may be computed as a daily or weekly 0 to 100 score (higher score implies greater fatigue). To be able to compute a daily symptoms score, at least 4 items of the symptoms diary must be non-missing; otherwise the score is considered "missing". For each 7-day weekly score, at least 4 reported diaries with at least 4 items completed on each diary day are needed to calculate the FSIQ-RMS-S weekly score. If fewer than 4 diaries with data on at least 4 items are available within the 7-day period, then the weekly score is considered as "missing". Details of the scoring rules to calculate FSIQ-RMS-S daily and weekly scores as well as rules to define baseline and post baseline FSIQ-RMS-S weekly scores are given in the Appendix C.2.

Change from baseline in FSIQ-RMS symptoms score is defined as the absolute change from baseline to each time point.

5.4.2.1.1 Variables for supplementary analysis

For supplementary analyses change from baseline for the following weekly scores will be considered in addition separately:

- **Up to EOT + 7 days**
- **Per protocol** - for subjects in the PPS data up to the first protocol deviation leading to exclusion from the per protocol analysis

- **LOCF by visit** weekly scores by visit based on last observation carried forward (LOCF) will be derived. Baseline data will not be carried forward.

5.4.2.2 Cumulative number of combined unique active lesions (CUAL) from baseline to Week 108

CUAL from baseline to Week 108 is calculated as the sum of new T1 Gadolinium-enhanced (Gd+) lesions and new or enlarging T2 lesions without Gd enhancement on T1 at all post-baseline MRI visits up to the Week 108 MRI visit (if the Week 108 visit is missing, consider all post-baseline scans up to EOS).

MRI scans are scheduled at Baseline, at the Week 60 Visit, and the Week 108 Visit. For subjects who prematurely discontinue treatment, an additional premature EOT visit MRI is scheduled. Unscheduled MRI scans may be conducted in addition.

MRI data are read centrally and provided to the sponsor.

To account for varying observation time, e.g. in case of premature study withdrawal or missing MRI assessments, the analysis will be adjusted by: Time up to the last MRI (years) = (Date of last MRI considered in derivation of CUAL − Randomization date + 1) / 365.25 days.

Rationale: Due to the fact that yearly MRI scans are scheduled, the CUAL count is expected to be mainly driven by new T2 lesions. The number of T2 lesions is expected to increase with observation time. In contrast, in studies where MRI scans are scheduled more regularly, the CUAL count is mainly driven by new T1 Gd+ lesions which are expected to increase with the 'Number of MRI scans' conducted. Also, in this study, there is an additional MRI following premature treatment discontinuation which can occur at any time, therefore the 'number of MRI scans' is not necessarily expected to be a good approximation of the observation time.

5.4.2.2.1 Derivations details

The central reader provides the 'Number of new T1 Gd+ enhancing lesions' (variable T1GdNew_R) and 'Number of new T2 lesions without Gd enhancement' (variable T2New_R) at a visit as compared to the last scheduled MRI scan (with premature EOT visit being considered a scheduled MRI). At the first scheduled post-baseline visit comparison is made to baseline (with baseline defined as the Baseline Visit 2, or Visit 2 - 2nd attempt if available [see Section 5.4.3] for details and a rationale).

CUAL from baseline to Week 108 is derived as the sum of 'Number of new T1 Gd+ enhancing lesions' and 'Number of new T2 lesions without Gd enhancement on T1' across all post-baseline MRI visits up to the Week 108 MRI visit (if missing up to EOS) irrespectively if a visit is conducted during PTOP or not.

For subjects without unscheduled MRI visits this includes MRIs from the Week 60 Visit, Week 108 Visit, and if applicable premature EOT visit.

As only the number of new lesions as compared to the last scheduled visit is recorded by the central reader, lesions are not expected to be double counted when summing over scheduled visits.

In the rare event that the Baseline MRI visit is conducted after randomization date, the variable is derived in the same way, i.e., based on the nominal visit. Note that effectively the derived variable then covers a period starting slightly after randomization. If the baseline MRI is missing, new or enlarging T2 lesions as compared to baseline cannot be assessed and consequently the endpoint is considered missing. The baseline MRI is considered missing if presence of T1 Gd+ lesions at baseline is missing.

For subjects with at least one unscheduled post-baseline MRI scan, care needs to be taken in the derivation to avoid double counting of lesions and the following algorithm applies:

- If a subject has an unscheduled MRI which is the last available MRI in the time period considered (i.e., if Week 108 MRI visit is missing, last MRI result up to EOS): Include that last unscheduled MRI when deriving CUAL in the same way as for scheduled visits.
- If a subject has an unscheduled MRI which is not the last available MRI for the subject: Do not include that unscheduled MRI. Rationale: New / enlarging lesions at the unscheduled MRI as well as at the later scheduled MRI are reported in reference to the previous scheduled visit. Including both may result in double counting.

5.4.2.2.2 Variables for supplementary Analysis

CUAL up to EOT + 7 days

Defined as the cumulative number of combined unique active lesions from baseline to EOT + 7 days. Includes only MRI scans conducted after randomization and prior to EOT + 7 days (last drug intake + 7 days). All other derivations follow the same approach as for CUAL from baseline to Week 108, in Section 5.4.2.2. Note that unscheduled MRI assessments previously not included may now be included in the derivation, if they become the last MRI assessment in the newly considered period. The time up to the last MRI scan used in the derivation is calculated.

Subjects completing treatment with complete MRI information will be flagged (completing treatment, with baseline MRI scan available, and all post-baseline scheduled MRI scan results [Week 60 and Week 108] available and conducted up to EOT + 7 days). This will allow for a complete-case analysis, if needed.

CUAL up to EOS or per protocol exclusion

Defined for subjects in the PPS as the cumulative number of combined unique active lesions from baseline to first deviation leading to exclusion from the per protocol analysis. Includes only MRI scans conducted after randomization and prior to per protocol exclusion date. All other derivations follow the same approach as for CUAL

from baseline to Week 108, in Section 5.4.2.2. The time up to the last MRI scan used in the derivation is calculated.

5.4.2.3 Time to 12-week Confirmed Disability Accumulation (CDA) up to EOS

5.4.2.3.1 Main analysis of secondary endpoint

Time to first 12-week CDA is defined as the time from start date of the first onset of 12-week CDA minus date of randomization + 1 in days. Subjects without 12-week CDA are censored, following the rules below. It is assessed from baseline up to EOS.

Assessing disability by EDSS

Disability as measured by EDSS is assessed in all subjects at screening, baseline, and thereafter at scheduled visits every 12 weeks until the end of study. Additional EDSS assessments for individual patients may be conducted between scheduled visits (i.e., during an MS relapse). The EDSS is a disability scale that ranges from 0 (normal) to 10.0 (death) in 0.5-point steps (1-point step from 0 to 1). It is based on standard neurological examination in conjunction with observations concerning ambulation. For analysis the overall EDSS score as reported by the investigator from the eCRF is considered (Field: 'Expanded disability status scale').

EDSS increase criteria for disability accumulation

The following criteria for an increase in EDSS in derivation of disability accumulation apply:

- Increase of at least 1.5 in EDSS for subjects with a baseline EDSS score of 0;
- Increase of at least 1.0 in EDSS for subjects with a baseline EDSS score of 1.0 to 5.0;
- Increase of at least 0.5 in EDSS for subjects with a baseline EDSS score $\geq$ 5.5;

where baseline EDSS is defined as the last available EDSS score prior to or on the date of randomization.

12-week CDA

A 12-week CDA is an increase in EDSS as compared to baseline according to the criteria above which is confirmed at a scheduled visit after 12 weeks. The increase needs to be persistent at all EDSS assessments (scheduled or unscheduled) between onset (first EDSS increase in the considered period) and confirmation. 12-week CDA is assessed sequentially up to the last available EDSS assessment for a subject in the study following the algorithm detailed below.

Disability progression can only be confirmed at a scheduled visit, where the EOT (Visit Week 108 or premature EOT visit), and FU visits count as scheduled visits, outside of an ongoing relapse. In this context, relapse duration is defined as period between start date (inclusive) and end date (exclusive) if available and limited to 90 days from onset if end date is not available or duration is longer than 90 days.

Algorithm for deriving 12-week CDA

1. For all post-baseline EDSS assessments (scheduled or unscheduled) assess if the absolute change from baseline meets the EDSS increase criteria for disability accumulation, described above. Start with the first EDSS assessment that meets the criteria for EDSS increase. Onset date of the corresponding potential 12-week CDA is the EDSS assessment date.

2. Confirmation of EDSS increase:

   a. An EDSS increase of a potential 12-week CDA has to be confirmed by an EDSS increase at a scheduled visit, at least 70 days*, from the onset date of the potential 12-week CDA.

      * Note: As visits are scheduled every 12 weeks with a time window of ± 7 days for conducting the EDSS assessment, the minimum per-protocol allowed time difference between two EDSS assessments scheduled 12 weeks apart is 70 days (12 weeks = 84 days minus a 7-day visit time window for each of the 2 visits, i.e., 84 days − 14 days = 70 days).

      EDSS assessments conducted during a relapse, i.e., from relapse start date (inclusive) to minimum of relapse end date and relapse start date + 90 days (exclusive) are not considered for confirmation.

      If an EDSS increase cannot be confirmed due to no available EDSS assessment, but the subjects dies due to MS (experiences a relapse with outcome death, or death with reason 'Multiple sclerosis') the CDA is also considered confirmed, with confirmation date being the death date and onset date the initial EDSS increase.

   b. A confirmed EDSS increase as per (a) is considered persistent if every EDSS score available (scheduled or unscheduled) between the onset and the confirmation date of a potential 12-week CDA meets the EDSS increase criteria. The initial EDSS increase is considered a 12-week CDA.

      If any EDSS from onset to confirmation does not meet the EDSS increase criteria, the increase is not persistent and the initial EDSS increase is not considered a 12-week CDA.

3. Repeat step 2 for all EDSS assessments with an EDSS increase as per the criteria above for disability accumulation, starting with the first possible onset, by date, up to the last possible onset. Once a 12-week CDA is identified the algorithm can stop.

4. If at none of the assessments a 12-week CDA is identified, but the subject dies due to MS (experiences a relapse with outcome death) the subject is considered to have a 12-week CDA with onset date being the death date.

5. If no 12-week CDA is identified, the subject is considered censored with the censoring date derived according to the rules below.

Censoring
Censoring date is defined as: Date of last EDSS assessment without an EDSS increase. Time to censoring is defined as censoring date minus date of randomization + 1.

Subjects without post-baseline assessment are censored on randomization date. Similar, subjects without CDA who have only EDSS assessments with an EDSS increase that cannot be confirmed, are censored on randomization date. For subjects without baseline EDSS score, missing values are assigned.

Handling of missing dates
Missing dates for EDSS assessments will be imputed as follows:

- Partial date: Maximum of 'lower limit', 'previous scheduled EDSS assessment according to the visit label' + 1 day (if available), and 'randomization date'
- Missing date: do not impute and exclude from analysis

Missing dates for relapse start and end dates will be handled as follows:

- Partial or missing relapse end date with available relapse start date: use relapse start date + 90 days or upper limit of partial end date if earlier.
- Partial relapse start date: If only day is missing, consider the relapse has started on the last day of the month (or on relapse end date − 1 day if earlier) and ended 60 days later or at (upper limit of a partial) relapse end date, if earlier. If the month or year are missing the relapse does not lead to non-consideration of an EDSS assessment.
- Missing relapse start date: Corresponding relapse does not lead to non-consideration of an EDSS assessment. Do not consider imputed relapse start dates.

5.4.2.3.2 Variables for supplementary analyses

5.4.2.3.2.1 Time to 12-week CDA up to EOS or per protocol exclusion

Time to 12-week CDA for the per protocol analysis is derived in the same way as for the variable for the main analysis, but only based on EDDS assessments (for onset and confirmation) conducted prior to first occurrence of a deviation leading to exclusion from per protocol [see Section 7.1.3]. Also, deaths due to MS are only considered up to that date. Subjects not experiencing a 12-week CDA event based on their per protocol data are censored at the last EDSS prior to per protocol exclusion without possible onset as for the main analysis.

5.4.2.3.2.2 Time to 12-week CDA up to EOT + 7

The variable for the main analysis described above will be repeated for 12-week CDAs up to EOT + 7 days. The onset date of the 12-week CDA must be up to EOT + 7 days; the confirmation of the 12-week CDA can be on-treatment or based on post-treatment

data. Otherwise the same rules as described in Section 5.4.2.3.1 apply. Subjects that are considered not to have 12-week on-treatment CDA are censored on the last EDSS assessment up to EOT + 7 days, following the same rules as described in the previous section.

5.4.2.3.2.3 Time to potential 12-week CDA up to EOS

12-week CDA requires confirmation of an EDSS increase by a scheduled EDSS assessment 12 weeks after progression onset. A potential onset at the end of the study may not be confirmed due to end of follow-up within the study. For time to potential 12-week CDA up to EOS all subjects censored in the main analysis with an EDSS increase at the last EDSS assessment in the study are considered as having a potential 12-week CDA, although the confirmation is missing. The onset date of the event for these subjects is the first EDSS increase after the last EDSS assessment without increase. Otherwise the same approach for derivation as described in Section 5.4.2.3.1 is followed.

5.4.2.3.2.4 Time to potential 12-week CDA up to EOS or per protocol exclusion

The time to potential 12-week CDA analysis is repeated for time to 12-week CDA up to EOS or per protocol exclusion.

5.4.2.3.2.5 Time to 12-week CDA up to EOS – alternative censoring

12-week CDA events and corresponding time to event are derived as above for the main analysis [see Section 5.4.2.3.1].

Subjects without 12-week CDA event are censored at their last scheduled EDSS assessment − 12 weeks (84 days).

This alternative censoring algorithm is applied (a) to ensure independence of censoring from the EDSS outcome and consequently avoid informative censoring and (b) to reflect that onset of 12-week CDA events can only occur 12 weeks prior to the last scheduled assessment.

5.4.2.4 Time to 24-week CDA

5.4.2.4.1 Main analysis of secondary endpoint

The derivation of the variables for Time to 24-week CDA up to EOS follows the same approach as described in Section 5.4.2.3.1, but onset of progression is to be confirmed after 24 weeks (≥ 154 days, i.e., 168 days − two times 7-day visit time window) instead of after 12 weeks. The progression has to be persistent between onset and confirmation.

5.4.2.4.2 Variables for supplementary analyses

5.4.2.4.2.1 Time to 24-week CDA up to EOS or per protocol exclusion

Time to 24-week CDA for the per protocol analysis is derived in the same way as for the variable for the main analysis, but only based on EDDS assessments conducted prior to first occurrence of a deviation leading to exclusion from per protocol [see Section 7.1.3].

Also, deaths due to MS are only considered up to that date. Subjects not experiencing an event based on their per protocol data are censored at the last EDSS prior to per protocol exclusion without an EDSS increase as for the main analysis.

5.4.2.4.2.2 Time to 24-week CDA up to EOT + 7

The variable for the main analysis described above will be repeated for 24-week CDAs up to EOT + 7. The onset date of the 24-week CDA must be up to EOT + 7 days; the confirmation of the 24-week CDA can be on-treatment or based on post-treatment data. Otherwise exactly the same rules as described in Section 5.4.2.3.1 apply. Subjects that are considered not to have 24-week CDA up to EOT + 7 are censored on the last EDSS assessment without an EDSS increase up to EOT + 7 days, following the same rules as described in the previous section.

5.4.2.4.2.3 Time to potential 24-week CDA

24-week CDA requires confirmation of an EDSS increase by a scheduled EDSS assessment 24 weeks after progression onset. A potential onset at the end of the study may not be confirmed due to end of follow-up within the study. For time to potential 24-week CDA up to EOS all subjects censored in the main analysis with an EDSS increase at the last EDSS assessment in the study are considered as having a potential 24-week CDA, although the confirmation is missing. The onset date of the event for these subjects is the first EDSS increase after the last EDSS assessment without increase. Otherwise the same approach for derivation as described in Section 5.4.2.3.1 is followed.

5.4.2.4.2.4 Time to potential 24-week CDA up to EOS or per protocol exclusion

The time to potential 24-week CDA analysis is repeated for time to 24-week CDA up to EOS or per protocol exclusion.

5.4.2.4.2.5 Time to 24-week CDA up to EOS – alternative censoring

24-week CDA events and corresponding time to event are derived as above for the main analysis [see Section 5.4.2.3.1].

Subjects without 24-week CDA event are censored at their last scheduled EDSS assessment − 24 weeks (168 days).

This alternative censoring algorithm is applied (a) to ensure independence of censoring from the EDSS outcome and consequently avoid informative censoring and (b) to reflect that onset of 24-week CDA events can only occur 24 weeks prior to the last scheduled assessment.

5.4.3 Exploratory MRI efficacy variables

MRI data are provided from a central reader.

Baseline

The central reader defines the result from 'Visit 2 - Baseline' as baseline (or 'Visit 2a - Baseline 2nd attempt' if available) and uses that Baseline visit as a reference in derivations of various variables (e.g. new or enlarging T2 lesions).

For analysis, baseline for MRI variables is defined in the same way, i.e., Visit 2 unless Visit 2a is available. In the unexpected case a Visit 2 scan is conducted after randomization, it is still considered as baseline.

MRI at [redacted] or at [redacted] visit windowing

MRI variables at [redacted] or [redacted] are based on the result from the corresponding nominal visit (or from a re-mapped premature EOT visit [see Section 11.3]). In case of missing results, MRI results from an unscheduled visit can be re-mapped to a scheduled visit if conducted during the time window specified in Section 11.3. Care needs to be taken however in derivation of cumulative variables where occasionally unscheduled MRIs are not included, as per the respective derivation details.

Handling of missing MRI results at [redacted] at [redacted]

If MRI at [redacted] or [redacted] is missing, a last observation carried forward (LOCF) approach is implemented. The last MRI scan up to the visit window upper limit as specified in Section 11.3 is considered. Baseline information is not carried forward.

5.4.3.1 Number of [redacted] at [redacted] and [redacted]

The [redacted] **per visit** is derived as the sum of [redacted] at that visit.

Handling of missing MRI results at [redacted] or [redacted] follows the general rules for MRI data described in Section 5.4.3 above. [redacted] LOCF and [redacted] LOCF variables are also derived.

The [redacted] **to** [redacted] is derived as the sum of [redacted] This includes [redacted] scans if available. It may include [redacted] if it is the [redacted], following selection of [redacted] To account for a potentially varying number of scans contributing to the variable, the analysis will be adjusted by the 'Number of MRI scans' included in the derivation. Similarly, the [redacted] to [redacted] is derived.

The [redacted] All other derivations follow the same approach as

above for lesions up to [redacted] To account for a potentially varying number of scans contributing to the variable, the analysis will be adjusted by the 'Number of [redacted] is derived.

Note: The number of [redacted] At baseline [redacted] are not available, thus no corresponding record is expected.

Note: [redacted] are assumed to not enhance for more than 12 weeks. This requires handling of data as follows:

- *If at a visit, number of [redacted] are not recorded, the number of [redacted] Rationale: As per [redacted] t and [redacted]*
- *If at a [redacted] is recorded, these [redacted] are considered implausible and therefore not to be considered in derivations of [redacted].*

5.4.3.2 Cumulative number of [redacted]

Cumulative number of [redacted] is calculated as the sum of [redacted]

[redacted] as compared to the previous scheduled visit are recorded by the [redacted] in two distinct categories: [redacted] and [redacted] at a visit as compared to the previous scheduled visit. The sum of both is the [redacted]' at that visit compared to the previous scheduled visit. For subjects with missing baseline MRI, the number of [redacted] from baseline cannot be assessed and the endpoint is considered missing.

The cumulative number of [redacted] is derived as the sum of the '[redacted]' at all scheduled post-baseline MRI visits: [redacted] and premature EOT visit if applicable (irrespectively if a visit is conducted during PTOP or not), for subjects with only scheduled post-baseline MRI scans.

For subjects with at least one unscheduled post-baseline MRI scan, care needs to be taken in the derivation to avoid double counting of [redacted] and the following algorithm applies:

- If a subject has an unscheduled MRI which is the last available MRI in the time period considered: Include that last unscheduled MRI when summing over 'number of new or [redacted]' in the same way as the scheduled visits.
- If a subject has an unscheduled MRI which is not the last available MRI for the subject in the time period considered: Do not include that unscheduled MRI. Rationale: New / [redacted] at the unscheduled MRI as well as at the later scheduled MRI are reported in reference to the previous scheduled visit. Including both would result in double counting.

To account for varying observation time the analysis will be adjusted by: Time up to the last MRI (years) = Date of last MRI considered in derivation − Randomization date + 1 divided by 365.25 days.

Rationale: The number of [redacted] is expected to increase with observation time, not solely by the '[redacted]'. Therefore, adjustment by observation time is conducted. As in this study an additional MRI following premature treatment discontinuation is scheduled, the 'number of MRI scans' is not necessarily expected to be a good approximation of the observation time.

Similarly, the cumulative number of [redacted] is derived. All post-baseline MRI visits up to the [redacted] visit (or if missing [redacted]) are considered.

Cumulative number of [redacted]
Cumulative number of [redacted] on treatment is derived in the same way as above but only considering MRI scans recorded between study treatment start and [redacted]. The time up to last MRI scans included in the derivation of the variable is calculated and used for adjustment in the analysis.

5.4.3.3 Cumulative number of [redacted]

Derived in the same way as for [redacted] [see Section 5.4.2.2], but only including MRI scans up to the [redacted] (if missing, up to [redacted]

Note that if an unscheduled MRI becomes the last available MRI in the considered time period it is to be included in deriving the sum of lesions.

5.4.3.4 Percent Change of [redacted]

Percent change of [redacted] is not manually derived, but taken directly from the [redacted] at the corresponding scheduled [redacted] visit (irrespectively if during PTOP or not). Percent change from baseline to [redacted] is defined in a similar way.

Handling of missing MRI results follows the general rules for MRI data described in Section 5.4.3 above, with [redacted] and [redacted] LOCF variables also derived.

Variables for supplementary analysis

For analysis using a repeated measurements model all post-baseline MRI scans will be included ([redacted]) with time (days from randomization to each MRI scan) as continuous covariate. Percent change of [redacted] from baseline will be taken from the corresponding central reader variable, and the time up to each MRI scan (in days) from randomization is derived as Date of MRI scan − Date of randomization + 1.

5.4.3.5 Change from baseline in the [redacted] and to [redacted]

[redacted] is recorded by the central reader at each visit ([redacted]).

Change from baseline in the [redacted] at a visit is calculated as the value at a post-baseline visit minus the value at Baseline visit. It is derived for [redacted] and [redacted] from the corresponding visit (irrespectively if conducted during PTOP or not).

Handling of missing MRI results follows the general rules for MRI data described in Section 5.4.3 above, with [redacted] LOCF and [redacted] LOCF variables also derived.

5.4.3.6 Change from baseline in the [redacted] to [redacted] and [redacted]

Handling of missing MRI results follows the general rules for MRI data described in Section 5.4.3 above, with [redacted] and [redacted] variables also derived.

5.4.3.7 Absence of [redacted] and [redacted]

[redacted] at [redacted] and [redacted] are considered 'absent', if the number of [redacted] at [redacted] and at [redacted] is zero. If > 0 they are considered 'present'.

See Section 5.4.3.1 for derivation of '[redacted] at [redacted] and [redacted] at [redacted] If no MRI result at the visit is available, the information is considered missing.

Similarly, '[redacted] to [redacted] (including all unscheduled scans)' (required for derivation of [redacted] in Sections 5.4.4.4– 5.4.4.5) is defined 'absent' if the 'number of new Gd+ T1 lesions' at all post-baseline [redacted] up to the [redacted] MRI visit (if missing, up to EOS) is zero. This includes all scheduled and all unscheduled visits. If > 0 at any visit it is considered 'present'. The corresponding 'Number of MRI scans' and time to last MRI scan is derived. If no MRI result at the visit

is available (based on visit window), the information is considered missing (for lesion free subjects).

5.4.3.8 Absence of [redacted] at [redacted] and [redacted]

[redacted] as compared to baseline are considered 'absent', if the number of [redacted] is zero at all scheduled and unscheduled visits up to [redacted] If it is > 0 at any visit it is considered 'present'.

[redacted] at [redacted] as compared to baseline are considered 'absent', if the number of [redacted] is zero at all scheduled and unscheduled visit up to [redacted] (or if the [redacted] visit is missing but the subject has absence of lesions up to [redacted] If it is > 0 at any visit up to [redacted] (or if the [redacted] visit is missing up to [redacted] it is considered 'present'.

If no MRI result at the respective visit is available (based on visit window) or no baseline MRI is available, the information is considered missing (for lesion free subjects).

5.4.3.9 Proportion of [redacted]

Only defined for subjects with [redacted] present at baseline.

It is the proportion (%) of [redacted] Results from the [redacted] visit are considered irrespectively if conducted during PTOP or not. In addition, it is derived at the [redacted] visit. Number of [redacted] by visit (at [redacted] and at [redacted] is considered as separate variable.

The number of [redacted]) by visit is taken from the central reader (variable [redacted]. For [redacted]' [see Section 5.4.3.1].

Handling of missing MRI results follows the general rules for MRI data described in Section 5.4.3 above. In addition, a variable with LOCF imputation at [redacted] and [redacted] is derived. For subjects with missing results (either at baseline or post-baseline), the proportion is considered missing.

5.4.4 Other clinical efficacy variables

5.4.4.1 Time to [redacted]

The time to first confirmed relapse (in days) is defined as [Date of [redacted] − Date of randomization + 1] in days.

Subjects without any [redacted] will be censored at the EOS date, and the time is defined as [Date of EOS − Date of randomization + 1] in days.

Derivation of [redacted] from randomization up to EOS is detailed in Section 5.4.1. Among these [redacted], the earliest [redacted] start date is the 'Date of [redacted]. For analysis the variable is displayed in weeks: variables in days divided by 7.

5.4.4.2 Absence of [redacted] to [redacted] and [redacted] ([redacted]

For the derivation of this variable [redacted] and their corresponding onset dates, as defined in Section 5.4.1, are considered.

Absence of [redacted] to [redacted] outcomes are derived as:

'Yes' at [redacted] Subjects are considered not [redacted] at [redacted] if any confirmed [redacted] starting between randomization date and [redacted] is recorded.

'No' at [redacted] Subjects are considered [redacted] at [redacted] if no [redacted] starting between randomization date and [redacted] is recorded, and they were still on study on [redacted] (EOS date on or after [redacted]).

'Missing' at [redacted] Subjects without [redacted], but no longer on study on Day 420, are considered missing.

Absence of [redacted] from baseline to EOS ([redacted] is derived in the same way as above but replacing [redacted] with EOS day. Relapse free subjects with premature study withdrawal prior to [redacted] ([redacted] will be considered to have 'Missing' outcome.

5.4.4.3 Change in [redacted] to [redacted]

Absolute changes [redacted] by visit are derived, based on the overall [redacted] as reported by the investigator in the eCRF [see Section 5.4.2.3.1]. For subjects prematurely discontinuing treatment, results collected during PTOP are considered at the corresponding nominal visit in the same way as for subjects completing treatment at [redacted] as planned (e.g. by visit analysis at [redacted] considers both '[redacted] and 'PTOP [redacted] visits). Baseline [redacted] is defined as in Section 5.4.2.3.1.

On-treatment [redacted] will be flagged, i.e., [redacted] [redacted] performed after study treatment start (date/time) and prior to EOT + 7 days.

5.4.4.4 [redacted]

[redacted]. If at least one of the criteria is not fulfilled or the subject discontinues treatment prematurely, the subject is not considered to have achieved [redacted]

Derivation details

A subject is considered as having [redacted] activity if at least one of the following applies:

- [redacted] derived as detailed in Section 5.4.1.
- [redacted] present [see Section 5.4.3.7].
- [redacted] present [see Section 5.4.3.8].
- [redacted]; see Section 5.4.2.3.1 for event derivation and event date.

Subjects that discontinue treatment prematurely are considered as showing [redacted], irrespective of what the reason for treatment discontinuation is.

Subjects [redacted] as defined above, who have missing outcome for any component (absence of [redacted]) are considered to have missing outcome for [redacted]

[redacted] *to* [redacted]

Derived in same way as above but replacing always EOS (or [redacted] with [redacted] Subjects that discontinue treatment prematurely but after [redacted] ([redacted]) are considered to have achieved [redacted] at [redacted] unless any other condition is met. Subjects [redacted] with missing outcome for any component at [redacted] are considered missing.

5.4.4.5 [redacted]

No [redacted] is defined by the [redacted] If at least one of the criteria is not fulfilled or the subject discontinues treatment prematurely, the subject is not considered to have achieved [redacted]

Derivation is as for [redacted], but in addition a subject is considered as having [redacted] if the following applies:

- [redacted] [see Section 5.4.3.4].

[redacted]

Derived in same way as above but replacing always EOS (or [redacted] with [redacted] Subjects who discontinue treatment prematurely but after [redacted] are

considered to have achieved [redacted] at [redacted] unless any other condition is met. Threshold for percent change of brain volume at [redacted] is ≤ −0.4%.

5.4.4.6 Change in [redacted]

The [redacted]

[redacted] will be assessed at pre-randomization [redacted] and at [redacted] and at unscheduled visits due to [redacted] or other unscheduled visits ([redacted] etc.) and if applicable, at the corresponding visits in the PTOP. The outcome is recorded in the eCRF.

For each assessment, the [redacted] score will be calculated as the mean of the Z-scores of the [redacted]

[redacted]

Details for generating the [redacted] Z-score and how to handle missing data are given in Appendix C.1.

For analysis, baseline and nominal post-baseline visits are considered. See Section 11.3 for remapping of premature end of treatment or unscheduled visits (includes unscheduled visits due to relapse).

Baseline is defined as the last available [redacted] score prior to or on randomization date. While this should usually present the subjects [redacted] the [redacted] may be used. A flag is added indicating the [redacted] conducted up to the [redacted] inclusive.

Furthermore, to address premature study discontinuation an additional variable based on last observation carried forward (LOCF) is derived.

5.4.4.7 Change in the

The It is administered along with the at pre-randomization (3 tests, where the first 2 are considered practice tests) and at , and if applicable, at the corresponding visits in the PTOP and at unscheduled visits (due to or other unscheduled visits) and its outcome the score is recorded in the eCRF.

For analysis, baseline and nominal post-baseline visits are considered. See Section 11.3 for remapping of premature end of treatment or unscheduled visits.

Baseline is defined as the last available score prior to or on randomization date. While this should usually present the subjects A flag is added indicating the conducted up to the .

Furthermore, in order to address premature study discontinuation an additional variable based on last observation carried forward (LOCF) is derived.

5.4.4.8

as a binary variable and defined as:

- he threshold has been derived based on a blinded interim analysis anchored on the

5.4.4.9 Change from baseline by visit up to

The

For each of these (higher score implies greater) are calculated from the scores of the items in the

respective sub-domains. A total score will not be computed for the [redacted] as the research interest, currently, is on the [redacted]. See Appendix C.3 for details on the scoring.

It is derived for the following time points: Baseline, [redacted] Change from baseline is defined as absolute change from baseline.

5.4.4.10 Change from baseline in [redacted]

Th [redacted] The questionnaire will be administered in an electronic format.

PGI-S of Fatigue is completed during the pre-randomization period at [redacted] and at unscheduled visits (due to [redacted] or other unscheduled visits). If applicable, [redacted] will also be performed at the corresponding visits in the PTOP. For analysis, visits are assigned based time windowing as detailed in Section 11.3.

Absolute change from baseline is derived with baseline defined as last available score up to and including randomization date.

5.4.4.11 Change from baseline in [redacted]

The [redacted] The questionnaire will be completed on a paper form and subsequently transcribed in the eCRF.

[redacted] will be completed at [redacted] and at unscheduled visits (due to [redacted] or other unscheduled visits). If applicable, CGI-C will also be performed at the corresponding visits in the PTOP. For analysis, nominal visits are considered. See Section 11.3 for remapping of premature end of treatment or unscheduled visits.

5.5 Safety variables

5.5.1 Definition of treatment-emergency, baseline and change-from-baseline, last on treatment, follow-up assessments for safety variables

For all safety data analyses described in this SAP, baseline is considered to be the last valid assessment prior to first study drug intake.

For each individual safety variable, baseline is defined to be the last valid measurement available before the date and time of first dose intake of study treatment, unless otherwise specified. On days where study drug is initiated or re-initiated, the pre-dose assessment for that day is defined as the last non-missing assessment prior to the study drug intake on that day. The pre-dose assessment on Day 1 (if available) is identical to baseline; however, missing pre-dose data on Day 1 is not imputed with baseline. In case of (partially) missing assessment dates/times, or for assessments reported on Day 1 (or day of re-initiation) with recorded assessment time contradicting the time point label (e.g., pre-dose, 2 hours post-dose), the nominal visit and time point labels will be used to determine whether an assessment is considered for baseline (for example a blood pressure measurement with an assessment date on Day 1 but with a missing assessment time or an assessment time prior to the reported time of first study drug intake, will not be considered for baseline if reported under the time-point "2 hours post-dose" in the eCRF, however it will be considered if reported as "Pre-dose" in the eCRF).

Absolute changes from baseline are defined as post-baseline value minus baseline value, i.e., a positive sign indicates an increase compared to baseline.

A percentage change from baseline is defined as the absolute change from baseline divided by the baseline value (if the baseline value is > 0) and then multiplied by 100.

Absolute and percent change from baseline is calculated and stored for all continuous safety data.

In addition, the following assessments are flagged or derived per parameter:

- Last on-treatment assessment: Last assessment prior to or on EOT date +1 day)
- Day-15 follow-up, between EOT + 8 days and EOT + 22 days [see Section 11.4]
- Day-30 follow-up, between EOT + 23 days and EOT + 37 days [see Section 11.4]
- Last follow-up assessment: Last assessment between EOT + 8 and EOT + 37 days

A safety assessment (ECG, vital signs, laboratory) is considered treatment-emergent, if the assessment date is on or after the study treatment start date and prior to or on the study treatment end date + 15 days (inclusive). If both, the assessment date/time and the date/time of study treatment start are available, only events with date/time on or after the date/time of study treatment start are considered to be treatment emergent. In case of (partially) missing assessment dates/times, or for assessments on treatment Day 1 with reported assessment time contradicting the time point label, the nominal visit and time point labels will be used to determine whether an assessment is considered as treatment-emergent or not. For example, a blood pressure measurement with an assessment date on Day 1 but a missing assessment time or an assessment time after the time of first study drug intake, will not be considered treatment emergent if reported under the time-point "Pre-dose" in the eCRF, however it will be considered if reported as "2 hours post-dose" in the eCRF.

A safety event (AE, SAE, Death) is considered treatment-emergent if the onset date/date of occurrence is on or after the study treatment start date and prior to or on the EOT date + 15 days (inclusive). Missing or partially missing onset or occurrence dates are imputed as described in Section 12.3. If both the onset/occurrence date/time and the date/time of study treatment start are available, only events with date/time on or after the date/time of study treatment start are considered to be treatment emergent.

5.5.2 Adverse events

An adverse event (AE) is any event reported by the investigator on the Adverse Event eCRF of the main database or the database of unblinding potential. All AEs are coded using MedDRA dictionary (version 21.0).

5.5.2.1 Frequency and prevalence of adverse events

AEs are summarized according to both frequency and prevalence and based on various grouping terms (for example MedDRA preferred term, or MedDRA primary system organ class).

For frequency of subjects experiencing an AE, AEs reported more than once for a subject (based on grouping term) are counted only once per subject.

For total cumulative number of events, multiple records of the same MedDRA preferred term in the AE dataset for the same subject count as individual events (episodes of the same type of event) unless they have the same start date and – if available – time.

The observation time (in days) per subject for summaries of treatment-emergent AEs is calculated as: minimum (EOT + 15, EOS) − study treatment start date + 1.

5.5.2.2 Intensity of adverse events

For AEs reported more than once for a subject, only the worst outcome is considered. AEs with missing severity assessment are imputed to be of ‘severe’ intensity.

5.5.2.3 Relationship of adverse events

Relationship to study treatment is entered into the database as ‘related’ or ‘not related’. For AEs reported more than once for a subject, only the worst relationship is considered. Adverse events with missing relationship are considered to be related.

5.5.2.4 Fatal adverse events

Fatal adverse events are those with ‘Death’ reported as outcome.

5.5.3 Deaths

Death information (date of death and primary cause) is taken from the Death eCRF.

5.5.4 Serious adverse events

An AE is considered serious if the question 'Serious?' on the Adverse Event eCRF is answered with 'Yes'. AEs with seriousness criteria missing are considered to be serious AEs.

5.5.5 Adverse events leading to discontinuation of study treatment

Adverse events leading to discontinuation of study treatment are those with 'Action taken with study drug' reported as 'Permanently discontinued'.

5.5.6 Adverse events leading to temporary interruption of study treatment

Adverse events leading to temporary interruption of study treatment are those with 'Action taken with study drug' reported as 'Temporarily interrupted'.

5.5.7 Adverse events leading to hospitalization

AEs leading to hospitalization are those where the eCRF question 'Did the Adverse Event require subject hospitalization?' is recorded as 'Yes' on the Adverse Event eCRF.

5.5.8 Adverse events on Day 1 and Day 1 of re-initiation of study treatment

Treatment-emergent AEs on Day 1 of study treatment are those AEs that start on or after study treatment start (on Day 1), by time and date, and before the following calendar date.

Treatment-emergent AEs on Day 1 of re-initiation of study treatment are those AEs that start on the date of a study treatment re-initiation.

Note: AEs in the first dose database with a different onset date than the first study drug intake or Day 1 of re-initiation of study treatment are not considered as 'Day 1' / 'Day 1 of re-initiation of study treatment' events.

5.5.9 Other significant adverse events

5.5.9.1 Adverse events of special interest

Adverse events of special interest (AESI) include the anticipated risks of treatment with study drug and events that may be related to MS comorbidities. The following safety areas are addressed by the pre-defined adverse events of special interest, detailed definition is given in Appendix A:

- Effect on heart rate and rhythm AESI (including hypotension)
- Hypertension AESI
- Hepatobiliary disorders / Liver enzyme abnormality AESI
- Pulmonary AESI
- Macular edema AESI
- Infection AESI
- Herpetic infection AESI

- Skin malignancy AESI
- Non-skin malignancy AESI
- Seizure AESI

5.5.9.2 Major adverse cardiovascular events (MACE)

Based on a pre-defined list of preferred terms belonging to relevant Standardized Medical Dictionary for Regulatory Activities Queries (SMQs), AEs are selected for the MACE adjudication board evaluation. For each case sent for MACE adjudication, the board members individually assess whether the case is a myocardial infarction, a stroke, or another adverse event. For fatal cases, each member determines whether the death is considered of cardiovascular, non-cardiovascular, or undetermined cause. If not all individual assessments concur, the case is classified into the above listed categories based on a consensus meeting. For data analysis, each case is assigned to one of the following categories:

- cardiovascular death (if a death case is classified as cardiovascular)
- non-fatal myocardial infarction (if the case is classified as myocardial infarction but not as cardiovascular death)
- non-fatal stroke (if the case is classified as stroke but not as cardiovascular death)
- no MACE (if the case is classified as other adverse event, but not as cardiovascular death)

The onset date and treatment-emergent status of a MACE is determined by the onset date and treatment-emergent status of the corresponding AE. In case more than one AE is linked to the same MACE case, the earliest treatment-emergent AE onset date determines the MACE onset date; if none of the linked AE is considered treatment-emergent, the earliest AE onset date determines the MACE onset date. This may lead to cardiovascular death MACE with an onset date prior to the date of death.

5.5.10 Vital signs and body weight

Blood pressure data are collected in the eCRFs for both the main database and the database of unblinding potential. Weight and height are collected in the eCRF of the main database only.

5.5.10.1 Blood pressure

Systolic (SBP) and diastolic blood pressure (DBP) are measured twice (i.e., two SBP measurements and two DBP measurements) at all assessments (except the hourly post-dose assessments after first dose on Day 1 and on study drug re-initiation). For all data analyses, for derivation of baseline, last on-treatment, Day-15 follow-up, Day-30 follow-up, last follow-up assessment, and for flagging of high or low values, the average of the first and second measurements are considered. The average calculated in the eCRF irrespective of position of the subject (supine, standing, sitting) or arm (right, left) is

used. If only one measurement is available, this is used for all further derivations and summary statistics.

Baseline is flagged and absolute and percent changes from baseline are calculated for each post-baseline value, as described in Section 5.5.1. In the same manner, pre-dose values are flagged and absolute and percent change from pre-dose are calculated for the hourly post-dose assessments after first dose on Day 1 or at re-initiation.

Assessments meeting the following conditions are flagged in the ADaM dataset:

- SBP ≤ 90 mmHg or ≥ 20 mmHg decrease from baseline
- SBP ≤ 90 mmHg
- ≥ 20 mmHg decrease from baseline in SBP

- SBP ≥ 160 mmHg or ≥20 mmHg increase from baseline
- SBP ≥ 160 mmHg
- SBP ≥ 140 mmHg
- ≥ 20 mmHg increase from baseline SBP

- DBP ≤ 50 mmHg or ≥ 15 mmHg decrease from baseline
- DBP ≤ 50 mmHg
- ≥ 15 mmHg decrease from baseline in DBP

- DBP ≥ 100 mmHg or ≥ 15 mmHg increase from baseline
- DBP ≥ 100 mmHg
- DBP ≥ 90 mmHg
- ≥ 15 mmHg increase from baseline in DBP

For all hourly post-dose assessments after first dose on Day 1 or at re-initiation flags are set for the following conditions:

- SBP ≤ 90 mmHg or ≥ 20 mmHg decrease from pre-dose
- ≥ 20 mmHg decrease from pre-dose in SBP
- DBP ≤ 50 mmHg or ≥ 15 mmHg decrease from pre-dose
- ≥ 15 mmHg decrease from pre-dose in DBP

5.5.10.2 Height and Weight

Height (cm) will only be summarized in the baseline characteristics and thus no derivations (other than treatment day of assessment) are required. Absolute and percent

change from baseline is calculated for each post-baseline weight (kg) value, as described in Section 5.5.1.

5.5.10.3 Body Temperature

Collected at all visits in the Body Temperature eCRF as 'normal' or 'abnormal'. No definitions and derivations (other than treatment day of assessment) necessary.

5.5.10.4 Pulse rate

Collected at unscheduled visits in the Pulse Rate eCRF as 'normal' or 'abnormal'. No definitions and derivations (other than treatment day of assessment) necessary.

5.5.11 12-Lead Electrocardiogram (ECG)

Only the visit, date and time information of the ECG data is collected on the eCRF for both the main database and the database of unblinding potential. ECG measurement data is collected using vendor machines from ERT. Data is automatically submitted to ERT by the machine, centrally read, and transferred from ERT to Actelion.

5.5.11.1 Quantitative ECG variables

Central reader ECG data contains the following quantitative measurements and derived variables: Heart rate (HR) (bpm), PR interval (ms), QRS duration (ms), RR interval (ms), QT interval (ms), QT_CB (ms) and QT_CF (ms). Heart rate is labeled "Mean heart rate" in the data transferred, however for consistency with earlier studies and to avoid misinterpretation as value coming from Holter monitoring, the label used in the CSR is "Heart rate". Baseline is flagged and absolute and percent changes from baseline are calculated for each post-baseline value as described in Section 5.5.1. In the same manner, pre-dose values are flagged and absolute and percent changes from pre-dose are calculated for the hourly post-dose assessments after first dose on Day 1 or at re-initiation.

For by-visit and by-hour tables presenting summary statistics of quantitative ECG results, the latest (by date-time) transmitted measurement per nominal subject-visit/timepoint is used in case of multiple data available for a visit or visit/timepoint.

All assessments meeting the following conditions are flagged in the ADaM dataset:

- Heart rate ≤ 50 bpm
- Heart rate ≤ 45 bpm
- Heart rate ≤ 40 bpm
- PR interval > 200 ms and increase of > 20 ms compared to baseline assessment
- QTcF/QTcB prolongations of > 500 ms
- QTcF/QTcB prolongations of > 480 ms
- QTcF/QTcB prolongations of > 450 ms
- QTcF/QTcB increase from baseline > 30 ms

- QTcF/QTcB increase from baseline > 60 ms
- QTcF/QTcB prolongations of > 500 ms and increase from baseline > 30 ms
- QTcF/QTcB prolongations of > 500 ms and increase from baseline > 60 ms
- QTcF/QTcB prolongations of > 450 ms and increase from baseline > 30 ms
- QTcF/QTcB prolongations of > 450 ms and increase from baseline > 60 ms

In addition, similar notable abnormality flags are derived in relation to the pre-dose measurements for the post-dose assessments on Day 1 and start of study drug re-initiation, and for the 3-hours post-dose measurement at Week 12 visit:

- PR interval > 200 ms and increase of > 20 ms compared to pre-dose assessment
- QTcF/QTcB increase from pre-dose assessment of > 30 ms
- QTcF/QTcB increase from pre-dose assessment of > 60 ms

For presentation of these outliers in by-visit and by-hour tables, the most extreme (i.e., lowest heart rate and highest interval) is considered in case of multiple data available at a visit or visit/timepoint.

5.5.11.2 Qualitative ECG variables

Morphological ECG findings are reported by the central reader, and mapped to CDISC standard (codelist C71150, with high level categories from codelist C71152) in the SDTM.

For analyses in tables the following categories are anticipated:

- Atrioventricular Conduction
- Axis and Voltage
- Chamber Hypertrophy or Enlargement
- Conduction
- Ectopy
- Intraventricular-Intraatrial Conduction
- Rhythm Not Otherwise Specified
- ST Segment, T wave, and U wave
- Sinus Node Rhythms and Arrhythmias
- Supraventricular Arrhythmias
- Supraventricular Tachyarrhythmias
- Ventricular Arrhythmias

Findings related to interpretation or technical issues are only included in listings.

Morphological ECG findings are flagged as “New” if not present at any pre-treatment assessment (i.e. at any ECG assessed prior to the study treatment start) or “Pre-existing” if present at any pre-treatment assessment. In case of missing or non-evaluable pre-

treatment ECG assessment, it is conservatively assumed that any treatment-emergent morphological ECG finding is "New".

For presentation of these findings in by-visit and by-hour tables, all reported abnormalities are considered in case of multiple assessments available at a visit or visit/timepoint.

5.5.12 Laboratory

Safety laboratory samples are centrally analyzed by ACM and the results are electronically transferred to Actelion. In exceptional cases, the protocol allows the utilization of local laboratories. Local laboratory analysis results are entered in the eCRF with some exceptions for blinding reasons (e.g., lymphocytes are not entered in the eCRF, see protocol). Quantitative results from local laboratories are not summarized (i.e. not included in summary statistics or graphical representations). Qualitative results like (marked) abnormality categorization, liver test elevation categories, etc. are derived from local laboratory data and summarized together with qualitative results derived from central laboratory data.

Safety laboratory hematology, clinical chemistry and urinalysis comprise the following parameters planned as per protocol:

> **Hematology:** Red blood cell count, Total and differential WBC counts (basophils, eosinophils, lymphocytes, monocytes, neutrophils, band forms), Platelet count, Hemoglobin, Hematocrit.
>
> **Liver function tests and coagulation**: INR, ALT, AST, AP, total bilirubin.
>
> **Clinical chemistry (excl. liver tests)**: Lactate dehydrogenase, Creatinine, creatinine clearance (calculated by the central laboratory using Cockroft-Gault), Blood Urea Nitrogen, Urate, Glucose, Total cholesterol, Triglycerides, Sodium, potassium, chloride, calcium, Total protein, albumin, C-reactive protein
>
> **Urinalysis (dipstick provided by central laboratory):** pH (5.0, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5), Glucose (Negative, Trace, +, ++, +++, ++++), Proteins (Negative, Trace, +, ++, +++, ++++), Occult Blood (Negative, Trace, +, ++, +++), Leukocytes (Negative, Trace, +, ++, +++), Bilirubin (Negative, +, ++, +++), Urobilinogen (3.2, 16, 33, 66, ≥131). Also, for each of those parameters it is collected whether the result is normal or abnormal, and - if abnormal - whether clinically significant (no/yes).

For the above protocol planned parameters for hematology, liver tests, and clinical chemistry, the derivations detailed in Section 5.5.12.1 will be performed.

Derivations (other than treatment day of assessment) are not performed for any other laboratory parameters, and only the raw information and treatment day of assessment are

listed. This includes any laboratory parameters that were not planned to be collected per the protocol, and the following protocol planned parameters:

Hepatitis B Virus Surface Antigen, Varicella Zoster Virus IgG Antibody, HIV-1/2 Antibody, Hepatitis C Virus Antibody, M. tuberculosis IFN Gamma Response, pregnancy tests (Choriogonadotropin Beta, serum / urine), Teriflunomide plasma concentration.

5.5.12.1 Derivations

Numerical results are converted into both conventional and SI units as per QS document OTH-000005 (Definition of Marked Abnormalities in Laboratory Data). Results reported as below the lower limit of quantification (LLOQ) are set to the LLOQ value. Results reported as > XX are set to XX for calculation of summary statistics. For all quantitative safety laboratory data, baseline, last on-treatment, Day-15 follow-up, Day-30 follow-up, last follow-up assessment is flagged and absolute change from baseline and percent changes from baseline are calculated for each post-baseline value, as described in Section 5.5.1.

Flags are derived according to project specific ranges for marked laboratory abnormalities as documented in the protocol [see Appendix D]. Marked laboratory abnormalities are labeled to indicate the increasing severity of abnormally low ("LL", "LLL"), or high values ("HH", and "HHH") for each of the laboratory parameter listed. For INR, the general range specified in the protocol is applied to all subjects irrespective of concomitant treatment with anticoagulants.

The following flags will be derived for Alanine aminotransferase (ALT), Aspartate aminotransferase (AST), Total bilirubin (TBIL), and Alkaline phosphatase (ALP):

- ALT: $\geq 1 \times$ ULN, $\geq 3 \times$ ULN, $\geq 5 \times$ ULN, $\geq 8 \times$ ULN, $\geq 10 \times$ ULN, $\geq 20 \times$ ULN
- AST: $\geq 1 \times$ ULN, $\geq 3 \times$ ULN, $\geq 5 \times$ ULN, $\geq 8 \times$ ULN, $\geq 10 \times$ ULN, $\geq 20 \times$ ULN
- ALT or AST: $\geq 3 \times$ ULN, $\geq 5 \times$ ULN, $\geq 8 \times$ ULN, $\geq 10 \times$ ULN, $\geq 20 \times$ ULN
- TBIL $\geq 2 \times$ ULN
- ALT or AST $\geq 3 \times$ ULN and TBIL $\geq 2 \times$ ULN (at the same sample date)
- ALT or AST $\geq 3 \times$ ULN and TBIL $\geq 2 \times$ ULN + AP $< 2 \times$ ULN (at the same sample date)
- INR > 1.5 combined with ALT or AST ≥ 3 ULN (at the same sample date).

5.5.13 Other safety variables

5.5.13.1 Pulmonary function tests - Spirometry

Spirometry measurement data are collected using vendor machines from ERT. Data is automatically submitted to ERT by the machine. Prior to database lock and unblinding, Spirometry data are not transferred to any member of Actelion biostatistics as these data

are considered to carry a moderate risk of potential unblinding. The following parameters are provided:

- Forced Vital Capacity [FEV1] (L),
- Forced Expiratory Volume in 1 Second [FVC] (L),
- FEV1/FVC (%),
- Predicted FEV1 (L),
- Percent Predicted FEV1 [%FEV1] (%),
- Predicted FVC (L),
- Percent Predicted Forced Vital Capacity [%FVC] (%),
- Peak Expiratory Flow (L/s),
- Forced Expiratory Flow 25–75% (L/s),
- Forced Expiratory Flow 25% (L/s),
- Forced Expiratory Flow 50% (L/s),
- Forced Expiratory Flow 75% (L/s)

Several efforts per visit per subject may be recorded in the database (usually 8 efforts), and potentially even two or more sets of efforts per subject per visit (e.g., 2 × 8 efforts). For each set of efforts the best effort at a visit per parameter (e.g., FEV1, FVC) is identified by the spirometry over read at ERT (variable XPSPID = "0"). The central reader grades the best effort as 'acceptable', 'borderline', or 'unacceptable' (variable 'BTR Grade Code').

Analyses are based on best efforts. If two or more sets of efforts are available at the same visit the following algorithm applies for the selection of the effort at that visit. Only one best effort per subject per visit will be selected. FEV1/FVC (%) is rederived from the FEV1 best effort and FVC best efforts. The selection of the records is defined as follows:

Step 1: Only select 'best effort' records.

Step 2:

a. If at least one 'acceptable' or 'borderline' rated test is available at a visit, chose the highest parameter value amongst all available 'acceptable' or 'borderline' rated test at that visit (per parameter).

b. If no 'acceptable' or 'borderline' rated test is available, but at least one 'unacceptable' rated test is available, chose the highest parameter value amongst all available 'unacceptable' rated tests at that visit (per parameter).

Note: if only one value is present at a visit then this is the highest value and it is selected.

Step 3:

For baseline choose the highest value prior to study treatment start.

Note: For deriving baseline apply the algorithm from above on all available records that could qualify for baseline (e.g., if two sets of efforts are available on two different days prior to study treatment start, derive BL as the highest best effort).

Note: To favor selection of re-screening records for re-screened subjects, records qualifying for baseline are selected in a two-step approach: (i) Derive baseline based on all records collected in a 45-day window prior to study treatment start. If no such record is available, then (ii) derive baseline based on all records prior to study treatment start.

Baseline is flagged and absolute and percent changes from baseline are calculated for each post-baseline value, and last on-treatment and follow-up assessments are derived as described in Section 5.5.1.

Assessments meeting the following conditions are flagged in the ADaM dataset:

- Percent change from baseline in FEV1: < −20 %, < −30%
- Percent change from baseline in FVC: < −20 %, < −30%
- Absolute change from baseline in %FEV1: < −20 %, < −30%
- Absolute change from baseline in %FVC: < −20 %, < −30%
- FEV1/FVC < 70%
- Absolute change from baseline in FEV1 ≤ −200 mL or percent change from baseline in FEV1 ≤ −12 %
- Absolute change from baseline in FVC ≤ −200 mL or percent change from baseline in FVC ≤ −12 %
- Absolute change from baseline in FEV1 > −200 mL and percent change from baseline in FEV1 > −12 %
- Absolute change from baseline in FVC > −200 mL and percent change from baseline in FVC > −12 %

In the subset of subjects experiencing a decrease of ≥ 200 mL or ≥ 12% in FEV1 (or FVC) from baseline to the last assessment on treatment, the condition is considered **Reversible**, if no decrease of ≥ 200 mL and ≥ 12% in FEV1 (or FVC) from baseline is present at the last follow-up assessment. If the last follow-up assessment is not fulfilling the above described reversibility criteria, the condition is considered **Not reversible**. If no follow-up assessment is available, reversibility cannot be assessed due to missing data.

5.5.13.2 Pulmonary function tests - DL_{CO}

DL_{CO} will be assessed in a subset of 333 subjects at selected sites with appropriate and established PFT expertise. DL_{CO} test results are reviewed by a centralized independent reader.

An assessment usually consists of multiple maneuvers. Among multiple maneuvers (reported within the same visit and assessment date), the mean value of the two highest DL_{CO} measurements are derived to represent the DL_{CO} at that assessment. If only one

DL_{CO} measurements is available at an assessment, that measurement is used. Within the eCRF, DL_{CO} values are converted into 'mmol/min/kPA' which is the unit used in reporting.

In addition, DLco based on acceptable results only is derived, for sensitivity analyses. A maneuver is considered acceptable if ERT's overall statement is entered as "OK", i.e., only assessments reported as "acceptable DL_{CO} data received" or "borderline acceptable DL_{CO} data received" are considered for analysis. The mean value of the two highest acceptable DL_{CO} measurements is derived. If only one acceptable DL_{CO} measurements is available at an assessment, that measurement is used.

Only these derived assessment-level values are used for any further derivation, the term 'DL_{CO} at visit' will refer to this value derived from the individual maneuver measurements.

The following derivations are performed (for DL_{CO} from all measurements and DL_{CO} from acceptable measurements separately):

- DL_{CO} adjusted for Hemoglobin (Hb) is derived based on the method of Cotes [Macintyre 2005] using the Hb reported by the central laboratory (or from local laboratory if missing) at the same nominal visit based on below formulae:
  - $DL_{CO\ Hb\ corrected} = DL_{CO} \times (10.22 + Hb\ [g/dL]) / (1.7 \times Hb\ [g/dL])$ for male subjects older than 14 years;
  - $DL_{CO\ Hb\ corrected} = DL_{CO} \times (9.38 + Hb\ [g/dL]) / (1.7 \times Hb\ [g/dL])$ for female subject subjects and all subjects $\leq$ 14 years;

  If no Hb value is available at the corresponding visit, this variable is missing.
- Predicted DL_{CO} corresponding to each assessment is derived based on Cotes [Cotes 1993] as follows:
  - $DL_{CO\ predicted} = 11.11 \times Height\ [m] - 0.066 \times Age\ [years] - 6.03$ for male subjects, age is set to 25 years if age below 25 years;
  - $DL_{CO\ predicted} = 8.18 \times Height\ [m] - 0.049 \times Age\ [years] - 2.74$ for female subjects, age is set to 25 years if age below 25 years;

  The age used in above formula is derived as follows: Age at baseline + 0.5 + (Date of assessment − Date of (re-)screening) / 365.25;
- Percent Predicted DL_{CO} is derived as:
  $100 \times DL_{CO\ Hb\ corrected} / DL_{CO\ predicted}$

Baseline is flagged and absolute and percent changes from baseline are calculated for each post-baseline value, and last on-treatment and follow-up assessments are derived, as described in Section 5.5.1.

Assessments meeting the following conditions are flagged in the ADaM dataset:

- Percent change from baseline in DL_{CO}: < −20 %
- Percent change from baseline in hemoglobin corrected DL_{CO} : < −20 %
- Absolute change from baseline in percent predicted DL_{CO}: < −20 %

5.5.13.3 Dermatological Examination

Collected at baseline, Week 60, and EOT visit in the Dermatological Examination eCRF. No definitions and derivations necessary. Any resulting findings are recorded as adverse event (or medical history).

5.5.13.4 Optical coherence tomography

Collected at screening, Week 12, Week 24, Week 60, and EOT visit in the Optical coherence tomography (OCT) eCRF. No definitions and derivations necessary. Any resulting findings are recorded as adverse event (or medical history or relapse).

5.5.13.5 Ophthalmological Examination

Collected at screening, Week 12, Week 24, Week 60, and EOT visit. Data will be taken from both Ophthalmological Examination eCRFs. No definitions and derivations necessary. Any resulting findings are recorded as adverse event (or medical history or relapse).

5.5.13.6 Physical examination

Collected at screening, baseline, Week 4, Week 12, Week 24, Week 60, Week 84, EOT and FUP visit eCRFs. No definitions and derivations necessary. Any resulting findings are recorded as adverse event (or medical history or relapse).

5.5.13.7 Suicidal ideation

The electronic self-rated version of the Columbia-Suicide Severity Rating Scale (eC-SSRS©) is an assessment instrument that evaluates suicidal ideation and behavior. During an initial assessment it assesses lifetime as well as the recent history suicidality (scheduled at Visit 2-Baseline), and then prospectively monitors ideations and behaviors at subsequent follow-up assessments since the last call (scheduled at Visit 10-Week 60 and Visit 14-Week 108).

The eC-SSRS outcome categories are provided below. Each category has a binary response (yes/no) and are numbered and ordered below for convenience.

1 – Wish to be Dead
2 – Non-specific Active Suicidal Thoughts

3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
5 – Active Suicidal Ideation with Specific Plan and Intent
6 – Preparatory Acts or Behavior
7 – Aborted Attempt
8 – Interrupted Attempt
9 – Actual Attempt (non-fatal)
10 – Completed Suicide

Furthermore, **self-injurious behavior without suicidal intent** is also an eC-SSRS outcome (although not suicide-related) and has a binary response (yes/no). (Q07a)

The initial assessment has a lifetime and a recent history assessment. The recent history covers the last 1 month for suicidal ideation, and the last 3 months for suicidal behavior. Recent history questionnaire questions are only assessed for outcome categories with corresponding lifetime outcome question answered as yes. If the lifetime assessment for an outcome category is 'No', also the corresponding recent history outcome is considered 'No' for analysis.

Scoring: Scores are created at each assessment as follows:

- **Suicidal Ideation Score**: The maximum suicidal ideation category (1–5 on the eC-SSRS) present at the assessment. Assign a score of 0 if no ideation is present.
- **Suicidal Behavior Score:** The maximum suicidal behavior category (6–10 on the eC-SSRS) present at the assessment. Assign a score of 0 if no behavior is present.

A subject with multiple reported outcomes for suicidal ideation and suicidal behavior will be summarized under the worst reported outcome (i.e., max[1–5] and max[6–10], respectively).

The following definitions will be used:

- **Suicidal ideation**: A "yes" answer to any one of the five suicidal ideation questions (Categories 1-5) on the eC-SSRS. Taken from the 'Reported most severe ideation level'.
- **Suicidal behavior**: A "yes" answer to any one of the five suicidal behavior questions (Categories 6-10) on the eC-SSRS.
- **Suicidal ideation or suicidal behavior**
- **Serious suicidal ideation (score ≥ 4) or suicidal behavior**

The definitions above will be defined at any time in the period from study treatment start up to EOT + 15 days*, for pre-treatment lifetime, and for pre-treatment recent history.

Shifts from baseline (worst outcome from pre-treatment recent history) to the worst outcome up to EOT + 15 days will be derived using the following categories: No ideation

(score 0), non-serious suicidal ideation (score 1–3), serious suicidal ideation (score 4–5), suicidal behavior (score 5–10).

* Only post-baseline assessments after study treatment start date from a 'since last call' (i.e., not from a lifetime) questionnaire are considered for the analysis. Note: The questionnaire was introduced while the study was ongoing. Approximately 15% subjects are therefore expected to have their lifetime questionnaire completed during a post-baseline visit.

Baseline is defined as last available recent history result (maximum score) up to study treatment start date, i.e., pre-treatment recent history result.

Lifetime (pre-treatment) is defined as any event (worst outcome) in recent history and lifetime evaluation(s) up to study treatment start date.

5.6 Quality of life variables

5.6.1 Change from baseline by visit up to Week 108 in SF-36v2™ Health Survey domain and component scores

The SF-36v2™ instrument is a generic health related quality of life assessment provided by the Medical Outcomes Trust, Boston, USA.

It contains 36 items measuring health across eight areas or domains: Physical Functioning (PF, 10 items: #3a – #3j); Social Functioning (SF, 2 items: #6 & #10); Role Limitations due to physical problems (RP, 4 items: #4a – #4d); Role Limitations due to emotional problems (RE, 3 items: #5a – #5c); Mental Health (MH, 5 items: #9b – # 9h); Vitality (VT, 4 items: #9a – #9i); Bodily Pain (BP, 2 items: #7 & #8) and General Health perceptions (GH, 5 items: #1, #11a – #11d) and an additional item on reported Health Transition to perceive changes in health status in the past 12 months (HT, 1 item: # 2).

In addition, Physical and Mental Component Summary (PCS and MCS) measures are calculated based on aggregate domain scores.

For each of the eight domains, scores will be coded, summed and transformed to generate a score from 0 (worst possible health state) to 100 (best possible health state). These domain scores as well as PCS and MCS will then be standardized to Z scores and then again will be transformed to create norm based scoring also known as t-scores. Details for generating these domain scores and aggregate component scores, handle missing data items and duplicated item entries are given in Appendix C.4.

The SF-36v2™ is completed by the subject on an electronic device at Visit 2 (Baseline), at Visits 6, 7, 10, 12 (Weeks 12, 24, 60, and 84), and 14 (W108) and at unscheduled visits due to relapses (R1, R2, etc.).

For analysis, visits are assigned based on time windowing, as detailed in Section 11.3. Baseline is defined as last available score up to or on randomization date. A variable with LOCF imputation by visit is derived for the domain and aggregated scores.

5.7 Pharmacoeconomic variables

5.7.1 Change from baseline by visit up to Week 108 in Work productivity and activity impairment (WPAI:MS) scores

The WPAI:MS V2.0 assesses work productivity and activity impairment. It has a recall period of 7 days. The following variables are considered:

- **Current paid employment status (Y/N)**

Four WPAI:MS outcome scores are considered that are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity:

- Percentage of work time missed due to MS (absenteeism),
- Percentage of impairment while working due to MS (presenteeism),
- Percentage of overall work impairment due to MS,
- Percentage of activity impairment due to MS.

The first three of these outcomes will be calculated only for those subjects who are working for pay, and the last outcome will be calculated for all subjects. See Appendix C.5 for details on the scoring.

WPAI is completed by the subject on an electronic device at Visit 2 (Baseline), and at Visits 6, 7, 10, and 12 (Weeks 12, 24, 60, 84, respectively), 14 (Week 108). For analysis, visits are assigned based on time windowing, as detailed in Section 11.3.

Absolute change from baseline is derived with baseline defined as last available score up to and including randomization date.

5.7.2 Health care resource utilization from baseline up to Week 108

Health care resource utilization data is taken from the following CRF pages: Hospitalization as recorded on the Adverse events, and 'Relapse summary' forms. Intensive care unit admissions for MS relapses are taken from 'Intensive Care Unit Admissions for MS Relapses' form. Emergency room visits for MS are taken from the 'Emergency medical facility visits for MS' form.

The following health care resource utilization variables will be derived:

- Number of admissions / visits from baseline to Week 108 (Frequency of health care resource utilizations)
- Length of stay (days) from baseline to Week 108 (Duration of health care resource utilizations)

The variables above are derived for the following categories

- Any hospitalization: Derived from the Adverse event, Relapse summary and 'Intensive Care Unit Admissions for MS Relapses eCRFs.
- Hospitalizations due to AEs: Derived from the AE page
- Hospitalizations due to relapses: Derived from the Relapse summary eCRF
- Intensive Care Unit (ICU) admissions for MS Relapses: Derived from the 'Intensive Care Unit Admissions for MS Relapses' form and includes all admissions (irrespectively if due to MS relapse or due to MS)
- Emergency room visits for MS relapses: Derived from the 'Intensive Care Unit Admissions for MS Relapses' form and includes visits due to MS relapse and due to MS. Note that emergency room visits are not considered hospitalizations.

The length of stay (in days) per admission is defined as: Length of hospitalization = End date of hospitalization − Start date of hospitalization + 1 (in days).

The length of stay (in days) from baseline to Week 108 per category taking into account multiple encounters is derived as the sum of length of stay across all admissions / visits only considering days from baseline (randomization date) to Week 108 (Day 756). In case of overlapping hospitalizations, each day is counted only once.

The number of admissions / visits from baseline to Week 108 is defined as a count variable per category. Two hospitalizations with identical admission date will be counted as one hospitalization.

5.8 Pharmacodynamic variables

Peripheral blood lymphocyte counts from central laboratory (absolute and change from baseline counts) are derived, as described in Section 5.5.12.

Lymphocyte counts are further categorized into the following categories:

- $< 0.2 \times 10^{9}/L$;
- $\geq 0.2 \times 10^{9}/L - < 0.5 \times 10^{9}/L$;
- $\geq 0.5 \times 10^{9}/L - < 0.8 \times 10^{9}/L$;
- $\geq 0.8 \times 10^{9}/L - < 1.0 \times 10^{9}/L$;
- $\geq 1.0 \times 10^{9}/L$.

Lymphocyte counts from central laboratory are flagged for analysis according to their sample dates as 'Baseline', 'Last on treatment', 'Day-15 follow-up', 'Day-30 follow-up', and 'Last follow-up assessment' following the rules described in Sections 5.5.1 and 11.4.

6 DEFINITION OF PROTOCOL DEVIATIONS

This section refers to all protocol deviations as recorded in the database following the specifications provided in the protocol deviation code list.

Protocol deviation 'Study drug taken from non-allocated kit' (PD_MM.058) will be classified at unblinding into the following two sub-protocol deviations: 'Study drug taken from non-allocated kit: non-assigned treatment received' (PD_MM.0581) or 'Study drug taken from non-allocated kit: assigned treatment received' at unblinding (PD_MM.0582).

Protocol deviations are categorized while being entered into the database into the following categories:

- **Important** protocol deviations (Yes/No)
- **Timing** (before Randomization / after Randomization / after end of treatment/ anytime during the study)

In addition, protocol deviations are categorized by high-level topic in line with the Protocol Deviation code list [see Appendix B].

7 ANALYSIS SETS

7.1 Definitions of analysis sets

7.1.1 Screened analysis set

The Screened Analysis Set (SCR) includes all subjects who were screened and received a subject number.

7.1.2 Full analysis set

The Full Analysis Set (FAS) includes all randomized subjects.

In order to preserve the randomization, subjects will be evaluated according to the treatment they have been randomized to.

The main analyses of the primary and all secondary endpoints will be based on the FAS. For the efficacy analyses on the FAS, unless otherwise stated, all available efficacy data up to the EOS will be included.

7.1.3 Per-protocol analysis set / Per-protocol analysis

The Per-Protocol analysis (PP analysis) will be based on data from all subjects in the FAS not affected by major protocol deviations. Major protocol deviations are defined as protocol deviations that impact the assessment of the primary/secondary endpoints.

The Per-protocol Set (PPS) comprises all subjects included in the FAS without any major protocol deviations occurring prior to or at randomization. Due to the nature of the primary endpoint which is assessed over the entire study period rather than at a fixed time point, for subjects developing major protocol deviations during the study, only data collected after a major protocol deviation will be excluded from the PP analysis.

Subjects will be excluded from the PPS if:

- **Subject is randomized and did not satisfy certain eligibility criteria:**
  - Subject presenting no diagnosis of MS as defined by the revised (2010) McDonald Diagnostic Criteria for MS or with no relapsing course from onset (no diagnosis of RRMS or SPMS with superimposed relapses) [PD_PP.013];
  - No active disease as required per study protocol inclusion criterion 5 [PD_PP.014];
  - Diagnosis of MS with progressive course from onset (i.e., diagnosis of primary progressive MS or progressive relapsing MS) [PD_PP.047];
  - Treatment within 90 days prior to randomization with medication listed in study protocol exclusion criterion 8 [PD_PP.024], medications: Plasmapheresis, cytapheresis, i.v. immunoglobulin, Treatment with an investigational drug (within 90 days or five half-lives of the drug, whichever is longer), except biological agents;
  - Treatment within 180 days prior to randomization with medication listed in study protocol exclusion criterion 9 [PD_PP.025], medications: Azathioprine, methotrexate, or cyclophosphamide, Natalizumab, Other systemic immunosuppressive treatment (e.g., cyclosporine, sirolimus, mycophenolic acid), Non-lymphocyte-depleting experimental biological agents (e.g., daclizumab);
  - Treatment within 24 months prior to randomization with lymphocyte-depleting biological agents (e.g. rituximab or ocrelizumab) or cladribine [PD_PP.026];
  - Treatment at any time prior to randomization with alemtuzumab, mitoxantrone, leflunomide, teriflunomide, fingolimod, ponesimod or other investigational S1P modulators, stem-cell transplantation [PD_PP.027];

- **Subject received no treatment or did not comply with treatment:**
  - No treatment taken: Any randomized subject who never received study drug [PD_MM.175] is excluded from the Per Protocol Set;
  - Compliance < 80 %, as per compliance based on days on study drug, derived from the eCRF, as described in Section 5.3.2.1, is excluded from the Per Protocol Set

Data will be excluded from the per protocol analysis for subjects after the occurrence of one of the following criteria:

- **Subject received wrong study treatment or incorrect dose:**
  - Wrong study treatment taken for > 14 days (> 2 consecutive kits as one kit contains 7 doses): Derived as protocol deviation 'Study drug taken from a non-allocated kit' [PD_MM.058] and only applies if at least 2 non-allocated kits containing wrong treatment (derived at unblinding: PD_MM.0581) have been taken, exclusion from first intake start date + 14 days onwards. *Note: Wrong study*

treatment taken for ≤ 14 days is not expected to affect efficacy considerably based on pharmacodynamic profile of the study treatment.

- Premature treatment discontinuation: Any randomized subject with premature treatment discontinuation, derived from the eCRF, as described in Section 5.1.5, exclusion from EOT + 7 days onwards.

- **Subject took an unauthorized concomitant medication:**
  - Treatment with systemic corticosteroids and ACTH during treatment, except for MS relapses and short-term treatments with low dose/inhaled corticosteroids for pulmonary conditions defined in study protocol [PD_MM.087], from date of protocol deviation onwards;
  - Any administration of disease modifying therapy for MS during the treatment period [PD_MM.088, PD_MM.093], exclusion from medication start date onwards;
  - During treatment period, any investigational procedure for MS [PD_MM.096], from date of procedure onwards;

- **Occurrence of major protocol deviation that could confound the interpretation of analysis conducted on the FAS;**
  - EDSS assessments not conducted for more than 6 months, exclusion from 'last EDSS date prior to 6-month gap + 6 months' onward; derived from eCRF: time between two consecutive EDSS assessments (or last EDSS assessment and EOS date) is > 6 months (182 days).
  - Subject's randomized treatment allocation was unblinded. Treatment code broken before unblinding of the study (for a reason not related to management of a clinical event [PD_MM.054], or any other unblinding documented in the IRT/ eCRF), exclusion from date of unblinding onwards.

7.1.4 Safety analysis set

The safety analysis set (SAF) includes all subjects who received at least one dose of study treatment. Irrespective of the randomized treatment group, subjects are summarized grouped:

- by actual treatment received, if the same kind of study treatment (teriflunomide or ponesimod) was taken throughout the study.
- within the treatment group they were exposed to for the majority of time on study treatment in case both study treatments (teriflunomide and ponesimod) were taken by a subject at some point during the study. If a subject received teriflunomide and ponesimod for exactly the same number of days, they are summarized within the treatment they are assigned to by randomization.

Some analyses are conducted based on the following safety set subsets:

- Subset of subjects with at least one re-initiation (based on eCRF Study Drug Administration - Re-Initiation)
- Subset excluding subjects from centers in Canada, the United States, and Mexico. Subset is used to address differences in assessment schedule of laboratory assessments in local protocol versions as compared to the global protocol version.
- Subset of subjects who received non-randomized study drug during the study (derived from protocol deviation PD_MM.058.1)
- Subset of subjects who resumed ponesimod maintenance treatment without up-titration after an interruption of >3 days [see Section 5.3.2.2]
- Subset of subjects who deviated from planned up-titration regimen [see Section 5.3.2.2]

7.1.5 Other analysis sets

The following other analysis sets are defined:

7.1.5.1 Non-conventional MRI sub-study full analysis set

Includes all subjects in the full analysis set who have consented to participate in the non-conventional MRI techniques sub-study ('Non-conventional MRI techniques sub-study' ticked on 'Informed Consent' eCRF).

7.1.5.2 Patient preference sub-study full analysis set

Includes all subjects in the full analysis set who have consented to participate in the Patient Preference sub-study ('Patient Preference Questionnaire' ticked on 'Informed Consent' eCRF).

7.1.5.3 DL_{CO} sub-study safety analysis set

Includes all subjects in the safety set who have consented to participate in the DL_{CO} sub-study ('DL_{CO} sub-study' ticked on 'Informed Consent' eCRF). Subjects are summarized in the same treatment group as in the safety analysis set.

7.1.5.4 Post-treatment observation period analysis set (PTOP Set)

Includes all subjects in the full analysis set who prematurely discontinued treatment (see definition in Section 5.1.5), with EOS date after EOT + 7 days.

7.1.5.5 Post-treatment safety analysis set

Includes all subjects in the safety analysis set who have an EOS date > EOT + 15 days. Subjects are summarized in the same treatment group as in the safety analysis set.

7.2 Usage of the analysis sets

Table 7 provides an overview of the analysis set usage for the main variables of the study.

Table 7 Overview of the different main analysis sets and their usage

Analyses/Data Displays	Screened analysis set	Full analysis set	Safety analysis set	Per protocol analysis set
Inclusion/exclusion criteria	✓			
Demographic characteristics		✓		✓
Baseline characteristics		✓		✓
Medical history		✓		
Previous and concomitant medications		✓		
Treatment exposure			✓	
Efficacy: Primary endpoint		✓		✓
Efficacy: Secondary endpoints		✓		✓
Efficacy: Exploratory endpoints		✓		
Safety endpoints			✓	
All subject listings		✓		

Main analysis of the primary efficacy endpoint is based on the Full analysis set, selected listings will be prepared on the Screened analysis set.

8 DEFINITION OF SUBGROUPS

The below described subgroups are considered in selected efficacy analyses. If applicable, the category in bold type is the reference category used in the statistical analysis of these subgroups.

- Prior use of MS disease-modifying treatment in the last two years prior to randomization (yes, **no**) as documented by the IRT system and used as randomization strata.
- Baseline Expanded Disability Status Scale score (**EDSS ≤ 3.5**, EDSS > 3.5) as documented by the IRT system and used as randomization strata.
- Prior use of MS disease-modifying treatment in the last two years prior to randomization (yes, **no**) as derived based on medications entered in the eCRF [see Section 5.2.2].
- Any prior treatment for MS (yes, **no**).
- Baseline Expanded Disability Status Scale score (**EDSS ≤ 3.5**, EDSS > 3.5) as documented in the eCRF.
- Geographical region of site (Europe Non-EU + Russia, European Union (EU) + UK, North America, Rest of World), as defined in Section 5.2.1;
- Sex (**male**, female);
- Age at screening (**< 40 years**, ≥ 40 years);
- MS subtype (**relapsing remitting**, secondary progressive with superimposed relapses)

- Number of documented MS relapses within the last 12 months prior to screening (**≤ 1**, ≥ 2), as defined in Section 5.2.2;
- Gd+ T1 lesions at baseline, on MRI scan as provided by the central reader (**absent**, present);
- Highly active disease (yes, **no**), as defined in Section 5.2.2;
- Protocol version at randomization:
  - **Randomized pre-PV4**: Subjects who have consented originally to protocol versions 1 to 3; these subjects can have data collected under protocol versions 1 to 3 only or under protocol versions 1 to 4.
  - Randomized post-PV4: Subjects who have consented to PV4 only; data of these subjects is only being collected under PV4.

  derived as described in Section 5.1.8.

In case of missing information, the corresponding subject is not considered for the respective subgroup analysis.

9 STATISTICAL ANALYSES

9.1 Overall testing strategy

9.1.1 Overall testing strategy

A multiple testing strategy is applied for the primary and secondary endpoints. The testing strategy starts with testing the primary endpoint at full alpha, followed hierarchically by a fallback type procedure for the secondary endpoints.

The multiple testing strategy starts with testing the primary endpoint (ARR using confirmed relapses), comparing the investigational treatment (ponesimod 20 mg) with the comparator (teriflunomide 14 mg) at full alpha.

If the primary endpoint null hypothesis is rejected, the alpha will be split evenly (1/3 of alpha) between the following three secondary endpoints using the fallback procedure, allowing for passing on alpha within the fallback procedure as per the following sequence:

- Change from baseline to Week 108 in fatigue-related symptoms as measured by the symptoms domain of the FSIQ–RMS [Fatigue]
- Cumulative number of unique active lesions from baseline to Week 108 [CUALs]
- Time to 12-week CDA from baseline to EOS [12-week CDA]

If a secondary endpoint listed as part of the fallback procedure above is successful, the preserved alpha in the sequence is passed along to the next secondary endpoint in the sequence and the summed alpha is used for testing that endpoint. Thus, as sequential tests are successful, the alpha accumulates for the endpoints later in the sequence; these endpoints are then potentially tested with a progressively larger alpha, but always with at least 1/3 of the alpha.

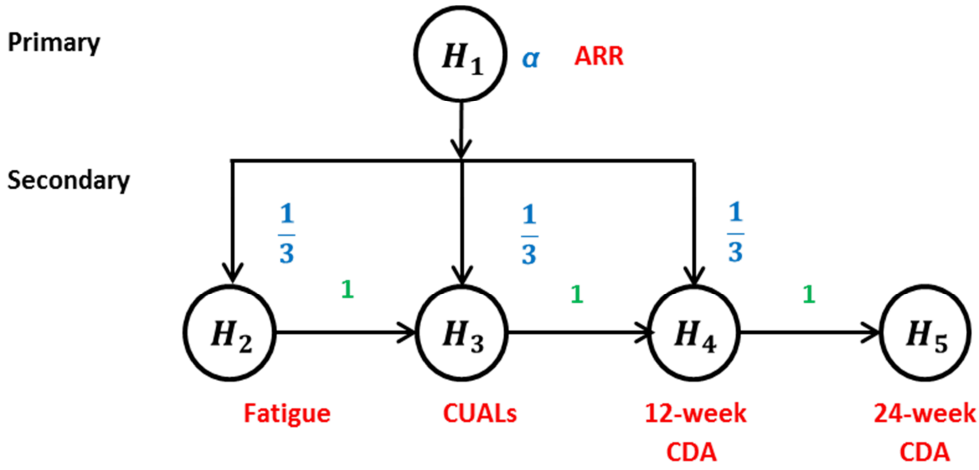
In a last step, the secondary endpoint

- Time to 24-week CDA from baseline to EOS [24-week CDA]

is tested in a subsequent stage following Time to 12-week CDA with the remaining alpha.

A graphical description of this testing strategy for the primary and secondary endpoints following the approach of Bretz et al. [Bretz 2009] is provided in Figure 2.

Figure 2 Overall testing strategy



ARR = Annualized relapse rate; CDA = Confirmed disability accumulation; CUAL = Cumulative number of combined unique active lesions; H = Hypothesis.

The primary null hypothesis will be tested at a two-sided 1% alpha level, and if rejected the study will be declared to show conclusive evidence of efficacy on the primary endpoint. If the primary null hypothesis can only be rejected at the two-sided 5% alpha level, the study will be declared positive.

The alpha level for the overall testing strategy across primary and secondary endpoints is set at 5%. Multiplicity adjustment as per the testing strategy outlined above will be performed. If the null hypothesis for the primary endpoint is rejected at two-sided alpha of 5% this leads to the example scenarios for alpha allocation for the secondary endpoints detailed as in Table 8.

Any further conducted tests will be considered exploratory.

Table 8 Example scenarios secondary endpoint alpha allocation

Scenario	Endpoint	Proportion of α	p-value for test	Statistically significant
1	Fatigue	1/3	< 1/3 α	Yes
	CUALs	2/3	< 2/3 α	Yes
	12-week CDA	1	< α	Yes
	24-week CDA	1	< α	Yes
2	Fatigue	1/3	> 1/3 α	No
	CUALs	1/3	< 1/3 α	Yes
	12-week CDA	2/3	< 2/3 α	Yes
	24-week CDA	2/3	< 2/3 α	Yes
3	Fatigue	1/3	< 1/3 α	Yes
	CUALs	2/3	> 2/3 α	No
	12-week CDA	1/3	< 1/3 α	Yes
	24-week CDA	1/3	< 1/3 α	Yes
4	Fatigue	1/3	< 1/3 α	Yes
	CUALs	2/3	< 2/3 α	Yes
	12-week CDA	1	< α	No
	24-week CDA	0		
5	Fatigue	1/3	> 1/3 α	No
	CUALs	1/3	> 1/3 α	No
	12-week CDA	1/3	< 1/3 α	Yes
	24-week CDA	1/3	< 1/3 α	Yes
6	Fatigue	1/3	>1/3 α	No
	CUALs	1/3	>1/3 α	No
	12-week CDA	1/3	>1/3 α	No
	24-week CDA	0		

Scenarios for secondary endpoint alpha allocation are conditional on rejecting the primary null hypothesis.
CDA = Confirmed disability accumulation; CUAL = Cumulative number of combined unique active lesions.

9.2 General rules for data presentations

This section describes the general rules applied for all data displays, unless otherwise specified in each corresponding section. Standard Guiding Rules and Principles and standard outputs are followed where applicable.

SAS version 9.4 is used for the preparation of all tables, listings and figures. Listings are sorted by treatment group, subject number (includes sorting by country and center) and timing (dates and/or visits as applicable). For analyses performed on multiple analysis sets, only one listing containing all analyzed subjects is presented. Names of outputs have a suffix that indicates the analysis set (e.g., _SAF for Safety set).

Data are listed and summarized using appropriate descriptive statistics:

- Number of non-missing observations, number of missing observations, mean, standard deviation, minimum, Q1, median, Q3, and maximum for continuous variables;
- Number of non-missing observations, and frequency with percentage per category (percentages based on the number of non-missing observations) for categorical endpoints;
- Some continuous variables obtained at scheduled visits are graphically presented utilizing box plots, on a time axis scaled by the target days of nominal visits. The box indicates the interquartile range, a horizontal line within the box indicates the median, and a diamond shape indicates the mean. Whiskers denote minimum and maximum values within the boundary of +/- 1.5 times the size of the interquartile range. Circles indicate outliers from +/- 1.5 times the size of the interquartile range.
- Some continuous variables are graphically presented based on mean change +/- SE on a time axis scaled to the target days of the corresponding visit. The standard error SE is derived as SD/$\sqrt{N}$.

Example table layouts:

```
ACT-128800/JNJ-67896153                Protocol: AC-058B301<TITLE>
Analysis set: <Set>
_____________________________________________________________________
            Ponesimod          Teriflunomide        Total
              20 mg                14 mg
             N = XX               N = XX            N = xx
_____________________________________________________________________

<XXX>       XX                 XX                   XX
_____________________________________________________________________
```

Treatment groups are displayed in the order (left to right) of Ponesimod 20 mg, Teriflunomide 14 mg, Total (if applicable). For figures, ponesimod is displayed in red with solid line style and teriflunomide is displayed in black with dashed line style.

9.3 Display of subject disposition, protocol deviations and analysis sets

9.3.1 Subject disposition

The number and percent of subjects in the FAS is summarized by country and site.

The number and percent of subjects screened and re-screened (incl. reasons for screening failure), subjects randomized, subjects completing study as per protocol, subjects treated, and subjects completed treatment per protocol, as well as subjects completing both treatment and study as per protocol are summarized for the SCR. For subjects that failed screening more than once, only the reason of the last failure is reported in the summary table.

All reasons for screening failures are included in the listing. The listing also includes the date of screening / re-screening, an indicator variable ('Yes', 'No') for failed screening attempts, the randomization date, and the number of days from successful screening to randomization.

The number and percent of subjects by reason for premature discontinuation from study is tabulated for the FAS. The number and percent of subjects by reason for premature discontinuation from study treatment is tabulated for the SAF.

Time in study and Time on treatment is presented with an inverse cumulative distribution function by treatment arm, descriptive statistics with cumulative time are displayed on the FAS.

Based on the FAS, the number and percent of subjects randomized under protocol versions 1–3 and 4 or higher are summarized by randomized treatment group. Based on the FAS, the number of subjects with data collected and the time on study (from randomization up to EOS) under protocol version 1–3 and under version 4 or higher are summarized by randomized treatment group by using descriptive statistics for continuous variables, overall and by time on study prior to protocol version 4 and time under study protocol version 4 or higher. Cumulative time on study (from randomization up to EOS) is presented overall, and cumulative time on study (from randomization up to EOS) is presented by time relative to version 4 along with percentage of overall time on study (from randomization up to EOS).

All subjects with unmet eligibility criteria are listed. For subjects re-screened unmet criteria from the initial screening visit are flagged. Subjects randomized with unmet criteria in their last screening attempt are flagged in the listing.

9.3.2 Protocol deviations

All protocol deviations as well as important protocol deviations are summarized by category (first timing, then high-level topic), per treatment group and overall on the FAS.

A listing of all protocol deviations (coded term, reported term) is provided on the SCR.

9.3.3 Analysis sets

The number and percentage of subjects in each analysis set and subjects included in the Per protocol analysis with data up to EOS are summarized in a table, per treatment group and overall.

Reasons for excluding subjects from the PPS, as well as reasons for excluding data from the Per protocol analysis are listed and summarized in a table. The time in study up to per protocol exclusion is summarized descriptively on the FAS.

9.4 Analyses of subject characteristics

9.4.1 Demographics

Demographic characteristics [defined in Section 5.2.1] are summarized using descriptive statistics for continuous and categorical data using the FAS, PPS and DL_{CO} safety set. Tables are presented by treatment group and overall. Demographic characteristics on the FAS will also be summarized by geographical region. All variables are also presented in a subject data listing based on the SCR.

9.4.2 Baseline disease characteristics

Baseline disease characteristics [defined in Section 5.2.2] are summarized using descriptive statistics for continuous and categorical data using the FAS and PPS. Tables are presented by treatment group and overall. Baseline disease characteristics on the FAS will also be summarized by geographical region. All variables are also presented in subject data listings based on the SCR.

9.4.3 Other baseline characteristics

Other baseline disease characteristics [defined in Section 5.2.3] are listed based on the SCR.

9.4.4 Medical history

Previous and concomitant medical histories [defined in Section 5.2.4] are summarized separately, by treatment group and overall, by tabulating the number and percentages of subjects who had/have at least one event and had/have each disease/diagnosis (by system organ class [SOC] and preferred term) using the FAS. SOCs are sorted by descending order of frequency. If the frequencies of SOCs are the same, alphabetical order is used. The same rule applies for preferred terms within SOC.

Complication or Symptoms Associated with MS within the last 24 months prior to the study [defined in Section 5.2.4] are summarized, by treatment group and overall, by tabulating the number and percentages of subjects who had/have at least one event and had/have each complication/symptom (by predefined term sorted by descending order of frequency) separately for previous conditions (not ongoing at screening) and conditions ongoing at screening using the FAS.

All variables including - as applicable - the treatment day relative to study treatment start date of start and end date of diagnosis/disease are presented in a subject data listing based on the SCR.

9.4.5 Previous and concomitant therapies

Number and percentages of subjects having taken at least one treatment are presented by at least one therapy, ATC class and individual preferred term within each ATC class (ATC class level 4). All summaries are tabulated by ATC class, and individual preferred terms within each ATC class, for the FAS, by treatment group and overall. ATC classes

are sorted by descending order of frequency. If the frequencies of ATC class are the same, alphabetical order is used. The same rule applies for preferred terms within ATC class. If the frequencies of ATC class are the same, alphabetical order is used. The same rule applies for preferred terms within ATC class.

The above described frequencies are summarized by treatment group and overall for the following therapies:

- Previous therapies (stopped prior to study treatment start)
- Study concomitant therapies (taken between study treatment start and EOS)
- Treatment concomitant therapies (taken between study treatment start and EOT)
- Therapies taken between study treatment start and EOT + 15 days

All therapies taken are reported in one subject listing with all information collected on the eCRF presented, using flags to mark previous, study concomitant, treatment concomitant as well as ongoing at study treatment start using the screened set.

9.4.5.1 Disease modifying therapies (DMTs) for MS

The following therapies are summarized in the same way as for the previous / concomitant therapies above by treatment group and overall on the FAS:

- Previous DMTs for MS
- Study concomitant DMTs for MS
- Treatment concomitant DMTs for MS
- DMTs for MS started after EOT + 7 days

All DMTs for MS taken are reported in one subject listing using flags to mark previous, study concomitant, treatment concomitant as well as ongoing at study treatment start using the FAS.

9.4.5.2 Concomitant steroid for treatment of relapses

Study concomitant steroids (taken between study treatment start and EOS) are summarized in the same way as concomitant therapies by treatment on the FAS.

Accumulated steroid dose for treatment of relapse (i) up to EOS, and (ii) up to EOT + 7 days is summarized using descriptive statistics by treatment group on the FAS. A listing of corticosteroids for treatment of relapses on the FAS is prepared.

9.4.5.3 Beta blocking agents

Treatment concomitant beta blocking agents (taken between study treatment start and EOT) are summarized in the same way as for treatment concomitant therapies on the FAS. A listing of beta blocking agents on the FAS is prepared.

9.4.5.4 Therapies used for accelerated elimination

Therapies used for accelerated elimination started between EOT and EOT + 15 days are summarized in the same way as for treatment concomitant therapies on the FAS. A listing of therapies used for accelerated elimination on the FAS is prepared.

9.4.6 Other subject characteristics

Other subject characteristics will be listed. This comprises the Chest X-Ray and Reproductive system findings.

9.5 Analysis of study treatment exposure and compliance

9.5.1 Exposure

Exposure to study drug [defined in Section 5.3.1] is summarized using descriptive statistics for continuous and categorical data using the SAF. In addition, this table displays the overall subject year exposure (sum over all subjects' exposure). Tables are presented by treatment group. All variables are also presented in subject data listings based on the SAF.

Subjects with at least one interruption, duration of interruptions as well as subjects with interruptions with reported reason will be summarized by treatment group on the SAF.

9.5.2 Compliance with study treatment

Compliance with study treatment variables [defined in Section 5.3.2] are summarized using descriptive statistics for continuous and categorical data using the SAF. Tables are presented by treatment group. All variables are also presented in subject data listings based on the FAS.

9.6 Analysis of the primary efficacy variable

9.6.1 Primary estimand

The primary estimand, the target for estimation in this study, is defined by the following components:

Population: Subjects with relapsing multiple sclerosis (RMS), as defined by the inclusion/exclusion criteria from the study.

Variable: ARR (number of confirmed relapses per subject-year) up to EOS, with EOS being at least 108 weeks after randomization [see Section 5.4.1].

Intercurrent events with corresponding strategies:

- **Treatment discontinuation:** Treatment Policy Strategy; including all confirmed relapses regardless of treatment discontinuation.

- **Start of alternative DMTs for MS:** Treatment Policy Strategy; including all confirmed relapses, regardless of start of alternative DMTs. Note: As per protocol alternative DMTs for MS are only allowed to be started after treatment discontinuation, therefore this intercurrent event is not expected to occur prior to or independent from treatment discontinuation.
- **Study discontinuation:** Hypothetical Strategy; effect if all subjects remained on study as planned per protocol. The assumptions for handling study discontinuation are detailed for the main analysis in Section 9.6.3 and for the sensitivity analyses on premature study withdrawal in Section 9.6.4.3.

Summary Measure: Rate ratio of ponesimod versus teriflunomide.

The primary analysis will be based on the Full Analysis set [see Section 7.1.2].

9.6.2 Hypothesis and statistical model

A generalized linear model with negative binomial distribution for the number of confirmed relapses will be assumed as described in Section 10.1.

Two-sided hypotheses are expressed in terms of the model parameters μ_{P20mg} and μ_{T14mg}. The primary null hypothesis is that the ARR (and consequently the model parameter μ) does not differ between ponesimod 20 mg and teriflunomide 14 mg. The alternative hypothesis is that the ARR differs between ponesimod 20 mg and teriflunomide 14 mg.

$$H_{0,\,ARR}: \mu_{P20mg} - \mu_{T14mg} = 0$$

versus

$$H_{1,\,ARR}: \mu_{P20mg} - \mu_{T14mg} \neq 0$$

The null hypothesis will be tested by a two-sided Wald test within the NB regression model, with a two-sided significance level of 0.01 for conclusive evidence and 0.05 for a positive study. Two-sided 99% and 95% Wald CIs and the associated p-value will be calculated for the relative reduction in mean ARR for ponesimod 20 mg compared to teriflunomide 14 mg.

9.6.3 Main analysis

The primary statistical analysis will be performed up to EOS on the FAS using an NB regression model for confirmed relapses, with treatment as a factor and the binary stratification variables (EDSS $\leq$ 3.5 versus EDSS > 3.5; DMT within last 2 years prior to randomization [Yes/No]) and the number of relapses in the year prior to study entry (categories $\leq$ 1 [or missing, in order to avoid excluding subjects from the analysis] and $\geq$ 2) included in the model. The model also includes an offset variable defined as the log of time on study (in years) from randomization up to EOS. The primary null hypothesis will be tested with a two-sided alpha level of 1% for conclusive evidence and 5% for a positive study.

Mean model-based estimates of the ARR (for confirmed relapses), by treatment arm, as well as 99% CIs and 95% CIs are presented. A rate ratio comparing ponesimod 20 mg with teriflunomide 14 mg will be derived from the model including 99% CIs, 95% CIs and the corresponding p-value. The dispersion parameter will also be displayed.

Descriptive summary statistics of the number of confirmed relapses as a continuous variable and the number of confirmed relapses in categories (0, 1, 2, 3, 4, ≥ 5), by treatment arm will be presented. The total cumulative number of confirmed relapses, as well as the total cumulative time on study (summed for all subjects) will also be displayed. From the total number of confirmed relapses and the total time on study the raw unadjusted ARR, by treatment arm, will be calculated and summarized.

The fit of the model will be assessed and other distributions such as the Poisson distribution will be considered in case of a lack of fit. The treatment-by-covariate interactions will also be assessed in separate models; if the interaction p-value is < 0.01 it will be investigated further to determine the nature of the interaction (quantitative or qualitative).

Handling of study discontinuation

For the main analysis it is assumed that follow-up time after study discontinuation is missing at random: a hypothetical scenario is considered assuming that subjects who discontinued the study would have similar efficacy as the subjects from the same treatment group who did not discontinue the study.

Sensitivity analyses on this assumption are described in Section 9.6.4.

9.6.4 Sensitivity analyses to main analysis

9.6.4.1 ARR based on confirmed relapses up to EOS on the FAS - unadjusted analysis

Same analysis as for the main analysis, but with treatment as the only factor in the model, i.e., not adjusting for covariates including stratification variables. This analysis is conducted to assess the effect of adjusting for covariates in the primary analysis.

9.6.4.2 ARR based on confirmed relapses up to EOS including derived stratification variable

Same analysis as for the main analysis, but instead of adjusting for the stratification variables from the IRT adjust for eCRF-derived stratification variables.

This analysis was requested by the FDA for the SPA review of study AC-058B302 and is performed to ensure that the results are consistent with the primary analysis, which are adjusted for the IRT stratification variables and – after eCRF data cleaning – might not match.

9.6.4.3 ARR imputation analysis for subjects withdrawing from the study prematurely (imputation from EOS to theoretical Week 108)

To assess the impact of subjects withdrawing from the study prior to Week 108, sensitivity analyses will be performed. These include the imputation of confirmed relapses using various assumptions for the period with missing data.

The proportion of missing data, i.e., missing observation time, will be calculated as

$$1 - \frac{\text{Cumulative time between randomization and EOS (or theoretical Week 108,if earlier)}}{\text{Cumulative time between randomization and theoretical Week 108}}$$

and will be summarized descriptively. Time up to EOS (i.e., time in study) will be presented based on an inverse cumulative distribution function [see Section 9.3.1].

Multiple imputation (MI) methods for count data are applied, following the approach from Keene 2014; see Section 10.1.1 for details.

Number of relapses between EOS and theoretical Week 108 will be imputed for subjects who have prematurely discontinued from the study.

Several approaches will be used for the imputation of events:

- **Missing at random (MAR) MI approach:** Under this approach, the rate after withdrawal for each treatment arm is assumed to be the rate in the respective arm before withdrawal.
- **Reference-based MI approaches:** The rate in the reference (teriflunomide 14 mg) arm is used for imputation in both the ponesimod 20 mg and teriflunomide 14 mg arms. While the post-withdrawal events in the teriflunomide 14 mg arm are imputed under the MAR assumption for that arm, two different approaches are applied for the imputation of the post-withdrawal events in the ponesimod arm:
    a. **Copy reference MI approach:** Impute based on the reference arm; For a subject in the ponesimod arm, if the event rate prior to withdrawal is ***better than would be expected on the reference arm***, under this approach, their imputed event rate after withdrawal will also be ***better than the expected event rate on the reference arm***. The copy reference imputation is expected to be less favorable for ponesimod 20 mg as compared to MAR approach as it is relative to the rate in the reference arm, but not as extreme as the jump to reference approach described below.
    b. **Jump to reference MI approach:** Impute based on the reference arm; For a subject in the ponesimod arm, if the event rate prior to withdrawal is ***worse than would be expected on the ponesimod arm***, under this approach, their imputed event rate after withdrawal will be ***worse than the expected event rate on the reference arm***. This approach assumes that withdrawn subjects on the ponesimod

20 mg treatment will quickly lose any positive effect of their treatment (relative to the reference arm) when it is stopped.

The resulting rate ratio and corresponding 99% and 95% CIs will be reported and presented in tables and graphically.

A delta-adjustment MI method [Keene 2014] will also be applied to further assess deviations from the MAR assumption:

- **Delta adjustment MI approach:** Starting with the MAR assumption for both treatment arms, a 2-way delta adjustment is applied.

  Ponesimod 20 mg arm: The rate after withdrawal is multiplied by δ_P, ranging from 1 (MAR, i.e. same rate as prior to withdrawal) to 3.1 (~Tripling of relapse rate) in steps of 0.3.

  Teriflunomide 14 mg arm: The rate after withdrawal is multiplied by δ_T, ranging from 0.1 (relapse rate reduced by 90%) over 1 (MAR, i.e. same rate as prior to withdrawal) to 3.1 (~tripling of relapse rate) in steps of 0.3.

For each imputed dataset the same negative binomial regression model as for the primary analysis will be fit and results will be combined using Rubin's rule [see Section 10.1.1].

The resulting rate ratios and corresponding 99% CIs and p-values will be reported and presented in tables and graphically.

9.6.5 Supplementary analyses

Supplementary analyses are conducted for other quantities of interest.

9.6.5.1 ARR up to EOT + 7 days for confirmed relapses

The primary efficacy model will be repeated only including confirmed relapses up to EOT + 7 days as a response variable, treatment as a factor, the stratification variables and the number of relapses in the year prior to study entry (≤ 1 [or missing] and ≥ 2) as covariates based on the FAS. The model is offset by the log of number of days from randomization up to EOT + 7 days.

The following sensitivity analyses related to confirmed relapses up to EOT + 7 days are conducted:

- *ARR based on confirmed relapses up to EOT + 7 days on the FAS – unadjusted analysis:*

Same analysis as for the main analysis, but with number of confirmed relapses up to EOT + 7 days as a response variable, and treatment as the only factor in the model, i.e., not adjusting for covariates, including stratification variables. The analysis is conducted to assess the effect of adjusting for covariates.

- *ARR based on confirmed relapses up to EOT + 7 days on the FAS – multiple imputation between EOT + 7 days and theoretical Week 108:*

Furthermore, sensitivity analyses based on multiple imputation are conducted based on different assumptions following treatment discontinuation to assess the hypothetical effect if subjects would have stayed on treatment up to Week 108. Multiple imputation methods for count data are applied as for the sensitivity analyses related to the primary analysis [Section 9.6.4.3 and Section 10.1.1]. For subjects who have prematurely discontinued treatment post-treatment events, i.e., the number of relapses between EOT + 7 and theoretical Week 108, will be imputed.

Imputation of events is conducted with the following MI approaches: Missing at Random (MAR), Copy reference, Jump to reference. The corresponding interpretations are detailed in Section 9.6.4.3, only that imputation is now conducted post EOT + 7 days instead of post-study withdrawal.

The resulting rate ratio and corresponding 99% and 95% CIs will be reported and presented in tables and graphically.

9.6.5.2 ARR based on confirmed relapses up to EOS - Per protocol analysis

Same analysis as for the main analysis, but with number of confirmed relapses up to EOS (per-protocol) as a response variable and the logarithm of time (in years) between randomization and EOS (or up to per protocol exclusion date) as an offset variable based on the PPS. Conducted to check the consistency of the treatment effect in the per protocol analysis with the effect in the main analysis. This analysis will evaluate whether a bias in evaluating the treatment effect has been introduced in the FAS due to major protocol deviations.

9.6.5.3 ARR based on all relapses up to EOS on the FAS

Same analysis as for the main analysis, but with number of all relapses up to EOS as a response variable. All other aspects of the model are as per the main analysis. This analysis is described to check the consistency of the treatment effect on all relapses and that no bias was introduced in the confirmation of relapses.

The following sensitivity analysis is conducted

- *ARR based on all relapses up to EOS including derived stratification variable:* Same analysis as for the main analysis, but instead of adjusting for the stratification variables from IRT adjust for eCRF-derived stratification variables.

9.6.5.4 ARR based on confirmed relapses up to start of new DMT for MS on the FAS

Same analysis as for the main analysis, but with number of confirmed relapses up to the minimum of EOS date, or the start of a new DMT for MS as a response variable, and the logarithm of time (in years) between randomization and the Min (EOS date, start of a new DMT for MS) as an offset variable. By checking the consistency of the treatment effect in this analysis with the effect observed in the main analysis, we evaluate whether a bias in evaluating the treatment effect has been introduced in the FAS analysis due to the decision to start a DMT.

9.6.5.5 ARR up to EOS for imputed confirmed relapses

Same analysis as for the main analysis, but with number of imputed confirmed relapses up to EOS as a response variable (relapses with missing confirmation status imputed as confirmed relapses). All other aspects of the model are as per the main analysis. This analysis is conducted to assess that no bias was introduced by missing confirmation status (e.g., due to EDSS not completed).

9.6.5.6 ARR post EOT + 7 days following premature treatment discontinuation

This analysis is similar to the analysis described for the primary efficacy variable, using a model with the number of confirmed relapses in the PTOP as a response variable, treatment, stratification variables and the number of relapses in the year prior to study (≤ 1 [or missing] and ≥ 2) entry as factors and the logarithm of time between EOT + 8 days and EOS (in years) as an offset variable. The analysis is based on the Post-treatment observation period analysis set (FAS subjects prematurely discontinuing treatment with EOS date after EOT + 7 days).

9.6.5.7 ARR by time period

Confirmed relapses are analyzed by time period (up to Week 60, post Week 60, and in 12 week periods). The analysis is conducted separately for each period using descriptive statistics and fitting a model as for the primary efficacy variable (adjusting for the same covariates), with logarithm of observation time in the respective period (in years) as an offset variable. The following are displayed for each period: ARR with 95% CI per treatment arm, Rate ratio with 95% CI together with number of subjects included in the analysis.

9.6.5.8 Forest plot of primary efficacy analysis and main sensitivity / supplementary analyses

Results of the analyses will be summarized in forest plots containing: ARR including 99% CIs, by treatment (as text); Number of subjects by treatment (as text); Rate ratio including 99% CIs (as graphic); Dotted reference line at rate ratio = 1; Solid reference line for the primary analysis rate ratio; Indicators to the right and to the left of the reference line favoring ponesimod 20 mg or teriflunomide 14 mg; Treatment effect p-value (as text). The plot will be repeated presenting 95% CIs.

The following sensitivity analyses for the primary estimand will be displayed in one forest plot:

- Primary efficacy analysis for all confirmed relapses up to EOS, adjusted for the stratification variables (IRT) and the number of relapses in the year prior to study entry based on the FAS (as a reference);
- Sensitivity analysis for all confirmed relapses up to EOS, unadjusted based on the FAS;
- Sensitivity analysis for all confirmed relapses up to EOS, adjusted for derived stratification variables and the number of relapses in the year prior to study entry on the FAS;

Sensitivity analyses on the FAS for the primary estimand on handling missing data [see Section 9.6.4.3] will be displayed in a forest plot as well. It will contain the primary efficacy analysis, the MAR, Copy reference, and Jump to reference as well as a subset of results from the delta adjustment method.

The main supplementary analyses to the primary estimand will be displayed in a second forest plot:

- Primary efficacy analysis for all confirmed relapses up to EOS, adjusted for the stratification variables (IRT) and the number of relapses in the year prior to study entry based on the FAS (as a reference);
- Supportive analysis of all confirmed relapses up to EOT + 7 days, adjusted for the stratification variables (IRT) and the number of relapses in the year prior to study entry based on the FAS;
- Supportive analysis for all confirmed relapses up to EOS collected per-protocol, adjusted for the stratification variables (IRT) and the number of relapses in the year prior to study entry based on the PPS;
- Supportive analysis for all relapses up to EOS, adjusted for the stratification variables (IRT) and the number of relapses in the year prior to study entry based on the FAS.
- Supportive analysis for all relapses up to start of new disease modifying therapy (DMT) for MS, adjusted for the stratification variables (IRT) and the number of relapses in the year prior to study entry based on the FAS.
- Supportive analysis for ARR up to EOS for imputed confirmed relapses.

9.6.6 Other analyses related to the primary endpoint

9.6.6.1 Relapse characteristics and relapse symptoms

Relapse characteristics will be summarized using descriptive statistics on the FAS. This includes the following:

Number (%) of subjects with at least one: relapse, relapse requiring hospitalization, relapse treated with corticosteroids; for all relapses up to EOS, confirmed relapses up to EOS, all relapses up to EOT + 7 days, confirmed relapses up to EOT + 7 days.

Number (%) of relapses (out of total number of relapses) requiring hospitalization, treated with corticosteroids and descriptive statistics for continuous data for relapse duration (days); for all relapses up to EOS as well as for confirmed relapses up to EOS, all relapses up to EOT + 7 days, confirmed relapses up to EOT+7 days. Number (%) of relapses (out of total number of relapses) per relapse symptom category, as well as number of functional systems affected per relapse (out of total number of relapses) is summarized descriptively per category by treatment arm. It is conducted for all relapses up to EOS as well as for confirmed relapses up to EOS, all relapses up to EOT + 7 days, confirmed relapses up to EOT + 7 days.

9.6.6.2 Number of new or worsening neurological symptoms up to EOS

New or worsening neurological symptoms up to EOS are summarized overall and by sub-categories (confirmed relapse, non-confirmed relapse, symptoms not considered a relapse) using descriptive statistics by treatment arm. Percentages for sub-categories are presented out of the total number of new or worsening neurological symptoms.

New or worsening neurological symptoms are reported via a telephone interview or at the site in a face-to-face interview and at both these times a decision must be taken by the treating neurologist to refer the subject to the efficacy assessor. Once referred to the efficacy assessor the symptom is considered a relapse and is assessed for confirmation. To ensure there is no bias in the selection of symptoms considered as relapses between treatments, an analysis will compare the proportion of these new or worsening neurological symptoms that are considered as relapses or as confirmed relapses. If there is an imbalance in the selection between treatments, this will be explored further.

9.6.6.3 Change in the collection of the primary endpoint introduced by protocol amendment 3

To investigate the impact of protocol amendment 3, the cumulative time in the study with data collected under protocol version 3 or earlier (prior to switching to PV4) as a proportion of the total cumulative time in the study will be calculated. Depending on the proportion of data collected under each protocol, further supportive analysis will be introduced to assess the impact of the change in the method of the primary endpoint collection. The analysis is detailed in Section 9.3.1.

9.6.7 Subgroup analyses

Subgroup analyses for confirmed relapses up to EOS will be conducted on the subgroup variables listed in Section 8 using the FAS and will be repeated for a Per Protocol analysis. The analysis is conducted to assess consistency of the overall treatment effect across subgroup variables.

A separate NB model will be fitted within each subgroup level for confirmed relapses up to EOS, with treatment as the only covariate and the log of the time (in years) from randomization up to EOS as an offset variable. Estimates of ARR by treatment arm, rate ratio with 99% CIs within each subgroup level will be presented, accompanied by the number of subjects per category. The analysis is repeated presenting 95% CIs within each subgroup level.

The treatment-by-subgroup interaction p-value is estimated for each subgroup variable from a separate model including treatment, the subgroup and the treatment-by-subgroup interaction term(s) using a Likelihood ratio test. Interactions with a p-value < 0.01 will be investigated further to determine the nature of the interaction (quantitative or qualitative) and the association with other subgroups.

Results are presented in a summary table and in a forest plot. The forest plot will include the 'overall' treatment effect, based on the primary efficacy analysis as a reference line. Point estimates are presented within the forest plot with its size proportional to the number of subjects in the subgroup.

9.7 Analysis of the secondary efficacy variables

9.7.1 Analysis of change from baseline in FSIQ-RMS weekly symptoms score (Fatigue) to Week 108

9.7.1.1 Hypothesis and statistical model

Change from baseline in FSIQ-RMS weekly score is analyzed using a repeated measurements ANOVA model to compare the change from baseline to Week 108 across both treatment groups. The underlying hypotheses are expressed as two means, assuming the endpoint follows a normal distribution:

$H_{0,\,FSIQ}$: $\mu_{P20mg,\,W108} - \mu_{T14mg,\,W108} = 0$ vs.

$H_{1,\,FSIQ}$: $\mu_{P20mg,\,W108} - \mu_{T14mg,\,W108} \neq 0$.

The hypotheses are assessed at a statistical significance level of 1/3 $\alpha = 0.05/3$ as per the overall testing strategy [see Section 9.1.1] based on a t-test for contrasts of mean estimates, once the null hypothesis for the primary endpoint has been rejected.

9.7.1.2 Statistical Analysis

The main analysis will be based on the FAS. A mixed effect model repeated measures (MMRM) which includes baseline (FSIQ-RMS $SWS_{c\ Base}$) and the two stratification factors, i.e., prior use of DMT (Yes/No), baseline EDSS category (EDSS $\leq$ 3.5, EDSS > 3.5) as covariates, treatment (fixed effects), visit, treatment by visit interaction, and baseline (FSIQ-RMS $SWS_{c\ Base}$) by visit interaction is applied. An unstructured co-variance structure shared across treatment groups will be used to model within patient errors. The Kenwood Roger approximation will be used to estimate denominator degree of freedom and adjusted standard errors. This model assumes that missing data follow the

missing at random assumption. Subjects with at least one available score at a post-baseline visit and at the baseline visit are included in the analysis.

Treatment effects per visit are derived based on (difference in) least-square means with corresponding 95% CIs and p-values derived using the t-distribution. They will be presented in tabular and graphical form.

The main comparison for this secondary endpoint is conducted on the difference in least-square means at Week 108. The corresponding p-value is compared to the local significance level of 1/3 $\alpha = 0.05/3$.

The SAS code for this model is as follows (considered "draft" until fully validated at analysis stage):

```
proc mixed data=;
    class subject treatment visit DMT EDSS;
    model  Chg = Base  DMT  EDSS  treatment visit treatment*visit Base*visit/ noint
ddfm=kr;
    repeated visit /  sub = subject type = un;
    lsmeans Treatment*Visit / pdiff CL OM;
run;
```

If the normality assumptions seems grossly violated the analysis will be conducted based on a rank ANCOVA of change from baseline to Week 108 using a stratified test [see Section 10.3.3].

9.7.1.3 Supportive / Sensitivity analyses

Further supportive and sensitivity analyses will be carried out as indicated below. Results in terms of treatment difference (95% CI) at Week 108 will be tabulated and forest plots will be provided.

9.7.1.3.1 Change from baseline in FSIQ-RMS weekly symptom score to Week 108 - delta worsening adjustment in ponesimod

To evaluate the robustness of the main analysis to deviations from the missing at random assumption a delta worsening multiple imputation analysis is conducted. Starting with a MAR multiple imputation approach the mean effect following drop-out (last score available) is assumed to worsen in the ponesimod 20 mg arm by delta, δ_P, using a marginal approach. δ_P is increased by small increments until the outcome of the analysis shifts from significance to non-significance. An increment size of 0.5 is applied (clinical meaningful threshold for change / 10). For the teriflunomide 14 mg arm a MAR multiple imputation is performed, i.e., without adding any worsening ($\delta_T = 0$ for all analyses). Analysis for the imputed datasets follows the main analysis and a pooled treatment difference estimate, 95% CI and p-value are obtained for each delta. Results are displayed graphically versus delta.

The analysis is only conducted if the main analysis is significant.

Non-monotone missing data (in-between missing data points) are imputed based on MCMC methods using Proc MI in SAS in a first step of the multiple imputation procedure. The covariates baseline FSIQ score, prior use of DMT, and baseline EDSS category are considered in the imputation model.

9.7.1.3.2 Change from baseline in FSIQ-RMS weekly symptom score to Week 108 (LOCF) - ANCOVA

An ANCOVA of change from baseline to Week 108 based on LOCF is conducted, adjusting for the same covariates as in the main analysis (baseline value and stratification variables). Mean, mean difference and 95% CIs are presented.

9.7.1.3.3 Change from baseline in FSIQ-RMS weekly symptom score to Week 108 - based on on-treatment data (up to EOT + 7 days)

The above MMRM for the main analysis will be repeated fitting the model only to data collected up to EOT + 7 days. A treatment comparison at Week 108 is conducted.

9.7.1.3.4 Change from baseline in FSIQ-RMS weekly symptom score to Week 108 - Per protocol analysis

A per-protocol analysis will be conducted as supportive. This analysis will be based on subjects in the PPS and will only include data in the analysis that were collected prior to a protocol deviation leading to exclusion from the per protocol analysis. A model as for the primary analysis is applied on the reduced data and a treatment comparison at Week 108 is conducted.

9.7.1.4 Descriptive analyses

Daily responses to FSIQ-RMS symptoms items 1, 2, 3, 4, 5, 6, and 7, FSIQ-RMS Symptoms Daily Scores will be listed, as well as the FSIQ-RMS weekly symptoms score. For each treatment group, by visit, Symptoms Weekly Scores based on observed data and LOCF, and change from baseline of these scores will be summarized using descriptive statistics. Change from baseline by visit is in addition presented graphically using mean (± SE). A separate by visit summary displaying only change from baseline up to EOT + 7 days will be provided.

Frequency and percentages of subjects with an available score will be summarized by visit.

9.7.1.5 Subgroup analyses

Subgroup analyses for change from baseline to Week 108 in FSIQ-RMS symptoms score will be conducted on the subgroup variables listed in Section 8, apart from recruitment period, using the FAS. The analysis is conducted to assess consistency of the overall treatment effect across subgroup variables.

The above main MMRM analysis will be repeated within subgroup levels using a model as for the main analysis without adjustments for stratification factors. Estimates of mean at Week 108 by treatment arm, difference in means with 95% CIs within each subgroup level will be presented, accompanied by the number of subjects per category.

The treatment-by-subgroup interaction p-value is estimated for each subgroup variable from a separate model including treatment, the subgroup and the treatment-by-subgroup interaction term(s). Interactions with a p-value < 0.01 will be investigated further to determine the nature of the interaction (quantitative or qualitative) and the association with other subgroups.

Results are presented in a summary table and in a forest plot. The forest plot will include the 'overall' treatment effect, based on the main efficacy analysis as a reference line. Point estimates are presented within the forest plot with its size proportional to the number of subjects in the subgroup.

9.7.2 Analysis of cumulative unique active lesions from baseline up to Week 108 (CUAL)

9.7.2.1 Hypothesis and statistical model

A generalized linear model with negative binomial distribution as described in Section 10.1 will be assumed for the number of cumulative unique active lesions from baseline to Week 108.

Two-sided hypotheses are expressed in terms of the model parameters μ_{P20mg} and μ_{T14mg} from the negative binomial regression model. The null hypothesis is that the CUAL (μ) does not differ between ponesimod 20 mg and teriflunomide 14 mg. The alternative hypothesis is that the CUAL differs between ponesimod 20 mg and teriflunomide 14 mg.

$$H_{0,\,CUAL}: \mu_{P20mg} - \mu_{T14mg} = 0$$

vs

$$H_{1,\,CUAL}: \mu_{P20mg} - \mu_{T14mg} \neq 0$$

The treatment effect will be evaluated based on the relative reduction in mean CUAL for ponesimod 20 mg compared to teriflunomide 14 mg, with corresponding 95% Wald CI. The null hypothesis will be tested by a two-sided Wald test, comparing the p-value to the local significance level as per the overall testing strategy [see Section 9.1].

9.7.2.2 Statistical analysis

The main analysis of the secondary endpoint CUAL will be performed on the FAS using a negative binomial regression model with treatment as a factor and the binary stratification variables (EDSS ≤ 3.5 versus EDSS > 3.5; DMT within last 2 years prior to randomization [Yes/No]) and Gd+ T1 lesions at baseline (absent or present) included in

the model. The model also includes an offset variable defined as the log of the time up to last MRI (in years) considered for analysis.

Mean model-based estimates of the number of CUALs per year will be provided by treatment group, as well as 95% Wald CIs. A rate ratio comparing ponesimod 20 mg with teriflunomide 14 mg will be derived from the model including 95% Wald CIs and the corresponding p-value. The dispersion parameter will also be displayed.

The fit of the model will be assessed and other distributions such as the Poisson distribution will be considered in case of a lack of fit.

Descriptive summary statistics of the number of CUALs up to Week 108 continuous and in categories (0, 1–5, 6–10, 11+) will be presented. The total number of CUALs as well as the total cumulative time up to the last MRI scan (summed for all subjects) will also be displayed. From the total number of CUALs and the total cumulative time up to the last MRI scan the raw number of CUALs per year will be calculated and summarized.

A listing of CUAL data will be presented.

9.7.2.2.1 Handling of missing data

CUAL is derived based on all available MRI scans up to the Week 108 visit (or EOS if missing) [see Section 5.4.2.2]. To account for missing MRI scans and consequently varying follow-up time, the log-transformed time up to the last post-baseline MRI is used as an offset in the model. Subjects with no post-baseline MRI result will be excluded from the analysis.

9.7.2.3 Supportive/sensitivity analyses

9.7.2.3.1 Analysis of CUALs from baseline to EOS - Missing data sensitivity: Imputation from last MRI scan to Week 108

To assess the impact of missing data on CUAL, i.e., last MRI scan conducted prior to Week 108, sensitivity analyses will be performed. These include the imputation of CUALs between last MRI scan and the theoretical Week 108 using different assumptions for the period with missing data on the FAS.

Time up to the last MRI scan is presented with a cumulative distribution function by treatment arm.

Sensitivity analyses for the main analysis will be conducted based on the MI method for negative binomial count data [see Section 10.1.1].

For subjects with missing Week 108 MRI scan the number of CUALs from last MRI scan up to the theoretical Week 108 will be imputed under different assumptions.

For all MI methods, subjects without a post-baseline MRI scan will be imputed based on baseline covariates; they will not be excluded from the analyses. No imputation is performed for subjects with missing baseline MRI; they are excluded from the analyses.

Imputation of events is conducted under the following assumptions: Missing at Random (MAR), Copy reference, Jump to reference. The corresponding interpretations are detailed in Section 9.6.4.3, only that imputation is now conducted post last MRI scan instead of post study-withdrawal.

For each imputed dataset the analysis is adjusted for the same covariates as for the main analysis, i.e., the binary stratification variables (EDSS ≤ 3.5 versus EDSS > 3.5; DMT within last 2 years prior to randomization [Yes/No]) and Gd+ T1 lesions at baseline (absent or present).

The rate ratio estimates with corresponding 95% CIs will be computed for each method.

9.7.2.3.2 Analysis of CUAL - per protocol analysis

The main analysis will be repeated using a per protocol analysis approach, i.e., on the PPS and considering MRI scans up to the first deviation leading to exclusion from the per protocol analysis. The number of CUALs (per protocol) is used as a response variable and the log of the time to the last MRI scan (years) considered in the per-protocol analysis as an offset variable [see Section 5.4.2.2]. Covariate adjustment as for the main analysis will be applied.

9.7.2.3.3 Analysis of CUALs from baseline to EOT + 7 days

The main analysis will be repeated using the number of CUALs up to EOT + 7 days as a response variable, and the log of the time to the last MRI scan (years) considered in this analysis as an offset variable [see Section 10.1], using the FAS. Covariate adjustment as for the main analysis will be applied.

9.7.2.4 Subgroup analyses

Subgroup analyses for CUAL from baseline to Week 108 will be conducted on the subgroup variables listed in Section 8, apart from recruitment period, using the FAS. The analysis is conducted to assess consistency of the overall treatment effect across subgroup variables.

From a methodological point of view subgroup analyses as for the primary endpoint will be applied, with results presented in a table and a forest plot.

9.7.3 Analysis of Time to 12-week CDA

9.7.3.1 Hypothesis and statistical model

Hypotheses are formulated in terms of survival functions S(t), i.e., the probability that time to 12-week CDA is ≥ t for a day t relative to the date of randomization. Two-sided hypotheses are expressed in terms of the survival functions $S_{P20mg}(t)$ and $S_{T14mg}(t)$.

$H_{0,CDA}$: $S_{P20mg}(t) = S_{T14mg}(t)$ for all $t \geq 0$

vs.

$H_{1,CDA}$: $S_{P20mg}(t) \neq S_{T14mg}(t)$ for all $t \geq 0$

The null hypothesis will be tested using a two-sided stratified log-rank test. The p-value will be compared to the local significance level as per the overall testing strategy [see Section 9.1].

9.7.3.2 Statistical analysis

The main analysis on 'Time to 12-Week CDA up to EOS' will be performed on the FAS by a two-sided stratified log-rank test with stratification factors (EDSS ≤ 3.5 versus EDSS > 3.5; DMT within last 2 years prior to randomization [Yes/No]) as stratification variables.

In addition, hazard ratio and 95% CI based on a stratified Cox regression will be provided. The model will include treatment as factor, and the same covariates as for the log-rank analysis as stratification variables.

Data are summarized in a table including, number of subjects with event / censored, Kaplan-Meier estimates (unstratified) and corresponding CI (anticipated time points: 12, 24, …, 108 weeks), and median (as well as 25th and 75th percentiles) of the survival function together with CIs. A graphical display presenting a Kaplan-Meier plot (unstratified) going upward is presented including CIs at specific time points as well as the log-rank p-value, hazard ratio and CI from the main analysis.

A listing of 12-week CDA data will be prepared.

9.7.3.2.1 Handling of missing data

See variable derivation in Section 5.4.2.3.1.

9.7.3.3 Supportive / sensitivity analyses

To explore the effect of the stratification variables included in the primary analysis log-rank test and Cox regression the main analysis will be repeated by adding each stratification variable separately at a time, as well as performing equivalent un-stratified analyses (only treatment as factor).

9.7.3.3.1 Time to 12-Week CDA up to EOS - Per Protocol analysis.

The same analyses as for the main analysis will be repeated based on the Time to 12-Week CDA up to EOS or per protocol exclusion using the PPS, see variable from Section 5.4.2.3.2.1.

9.7.3.3.2 Time to 12-Week CDA up to EOT + 7 days

Repeat the main analysis of the secondary endpoint based on 12-Week CDA up to EOT + 7 days, as defined in Section 5.4.2.3.2.2, using the FAS.

9.7.3.3.3 Time to potential 12-Week CDA up to EOS

Repeat of the main analysis of this secondary endpoint based on time to potential 12-Week CDA, as defined in Section 5.4.2.3.2.3, using the FAS.

9.7.3.3.4 Time to potential 12-Week CDA up to EOS – Per protocol analysis

Repeat of the main analysis of this secondary endpoint based on time to potential 12-Week CDA up to EOS or per protocol exclusion using the PPS.

9.7.3.3.5 Time to 12-Week CDA up to EOS – alternative censoring

The same analyses as for the main analysis will be repeated based on time to 12-Week CDA up to EOS with alternative censoring algorithm [see Section 5.4.2.3.2.5].

9.7.3.4 Subgroup analyses

Subgroup analyses for time to 12-week CDA up to EOS on the FAS will be conducted on the subgroup variables listed in Section 8 (apart from 'recruitment period') using the FAS. The analysis is conducted to assess consistency of the overall treatment effect across subgroup variables.

Separate Cox regression models will be fitted within each subgroup level using only treatment as a factor and no other covariates. The hazard ratio and 95% CI of the treatment effect within each subgroup level will be displayed, accompanied by number of subjects with events / censored.

The treatment-by-subgroup interaction p-value is estimated for each subgroup variable from a separate model including treatment, the subgroup and the treatment-by-subgroup interaction term(s).

Results are presented in a summary table and in a forest plot. The forest plot will include the 'overall' treatment effect, based on the primary efficacy analysis as a reference line. Point estimates are presented within the forest plot with its size proportional to the number of events in the subgroup.

9.7.1 Analysis of Time to 24-week CDA

9.7.1.1 Hypothesis and statistical model

Hypotheses are formulated in terms of survival functions S(t), i.e., the probability that time to 24-week CDA up to EOS is ≥ t for a day t relative to the date of randomization. Two-sided hypotheses are expressed in terms of the survival functions $S_{P20mg}(t)$ and $S_{T14mg}(t)$.

$$H_{0,CDA}: S_{P20mg}(t) = S_{T14mg}(t) \text{ for all } t \geq 0$$

vs.

$H_{1,CDA}$: $S_{P20mg}(t) \neq S_{T14mg}(t)$ for all $t \geq 0$

The null hypothesis will be tested using a two-sided stratified log-rank test. The p-value will be compared to the local significance level as per the overall testing strategy [see Section 9.1].

9.7.1.2 Handling of missing data

See variable derivation in Section 5.4.2.4.1.

9.7.1.3 Statistical analysis

The main analysis on 'Time to 24-Week CDA up to EOS' will be performed on the FAS by a two-sided stratified log-rank test with stratification factors (EDSS ≤ 3.5 versus EDSS > 3.5; DMT within last 2 years prior to randomization [Yes/No]) as stratification variables, similar to the analysis for 'Time to 12-Week CDA up to EOS'.

Hazard ratio and 95% CI based on a stratified cox regression will be provided. The model will include treatment as factor, and the same covariates as for the log-rank analysis as stratification variables.

Data are summarized in a table including, number of subjects with event / censored, Kaplan-Meier estimates (unstratified) and corresponding CI (anticipated time points: 12, 24, …, 108 weeks), and median (as well as 25th and 75th percentiles) of the survival function together with CIs. A graphical display presenting a Kaplan-Meier plot (unstratified) going upward is presented including CIs at specific time points as well as the log-rank p-value, hazard ratio and CI from the main analysis.

A listing of 24-week CDA data will be prepared.

9.7.1.4 Supportive/sensitivity analyses

To explore the effect of the stratification variables included in the primary analysis log-rank test and cox regression the main analysis will be repeated by adding each stratification variable separately at a time, as well as performing equivalent un-stratified analyses (only treatment as factor).

9.7.1.4.1 Time to 24-Week CDA - Per Protocol analysis.

The same analyses as for the main analysis will be repeated based on the Time to 24-Week CDA up to EOS or per protocol exclusion using the PPS.

9.7.1.4.2 Time to 24-Week on treatment CDA up to EOT + 7

Repeat of the main analysis of this secondary endpoint based on 24-Week CDA up to EOT + 7 days using the FAS.

9.7.1.4.3 Time to potential 24-Week CDA up to EOS

Repeat the main analysis of this secondary endpoint based on time to potential 24-Week CDA using the FAS.

9.7.1.4.4 Time to potential 24-Week CDA -Per protocol analysis

Repeat of the main analysis of this secondary endpoint based on time to potential 24-Week CDA up to EOS or per protocol exclusion using the PPS.

9.7.1.4.5 Time to 24-Week CDA up to EOS – alternative censoring

The same analyses as for the main analysis will be repeated based on time to 24-Week CDA up to EOS with alternative censoring algorithm [see Section 5.4.2.4.2.5].

9.7.1.5 Subgroup analyses

Subgroup analyses for time to 24-week CDA up to EOS on the FAS will be conducted on the subgroup variables listed in Section 8 (apart from 'recruitment period') using the FAS using the same methods as for time to 12-week CDA up to EOS. The analysis is conducted to assess consistency of the overall treatment effect across subgroup variables.

9.8 Analysis of exploratory MRI efficacy variables

9.8.1 Number of [redacted] at [redacted] and [redacted]

The number of Gd+ T1 lesions at [redacted] and at [redacted] will be analyzed by visit on the FAS using a negative binomial regression model: Response variable is the [redacted] of [redacted] at [redacted] (observed, and separately for LOCF), or at [redacted] (observed, and separately for LOCF) respectively, treatment as a factor and [redacted], and [redacted] at baseline (absent or present) as covariates, and no offset. Mean model-based estimates, rate ratio, and corresponding 95% CIs and p-value will be presented

The number of [redacted] per visit will be summarized descriptively continuous and in categories (anticipated categories: 0, 1, 2, 3, 4+) based on observed data, as well as based on LOCF imputation, using the FAS. [redacted] data will be listed on the FAS.

The cumulative number of [redacted] up to [redacted] per scan will be analyzed on the FAS using a negative binomial regression model with treatment as a factor and [redacted] as covariates. The model also includes an offset variable defined as the log of the number of [redacted] considered for analysis. Mean model-based estimates, rate ratio, and corresponding 95% CIs and p-value will be presented. Descriptive statistics of cumulative number of [redacted] up to [redacted] and up to [redacted] will be displayed, separately based on MRI scans up to EOS and up to EOT + 7 days.

9.8.2 Cumulative number of [redacted]

The cumulative number of [redacted] will be analyzed on the FAS using a negative binomial regression model with treatment as a factor and [redacted] and [redacted] at baseline (absent or present) as covariates. The model also includes an offset variable defined as the log of the time up to last MRI (in years) considered for analysis. Mean model-based estimates by treatment arm, and rate ratio with corresponding 95% CIs and p-value will be presented. The analysis is repeated for cumulative number of [redacted]

Descriptive summary statistics for the number of [redacted] and to [redacted] will be presented for continuous data and frequency per categories (anticipated: 0, 1–5, 6–10, 11+), by treatment group, total counts and total cumulative observation time (time up to the last MRI scan in years) summed for all subjects, raw yearly rate (total count / total cumulative observation time).

[redacted] will be listed.

9.8.3 Percent change of [redacted] to [redacted]

[redacted] to [redacted] and to [redacted] will be summarized using descriptive statistics by visit based on observed data, as well as based on [redacted] and listed using the FAS.

[redacted] to [redacted] will be analyzed based on a linear mixed model using the FAS. The model is fitted on the response variable [redacted] from baseline using all available post-baseline MRI scans, and is assumed to follow a normal distribution. Time (days from randomization to each MRI scan) is considered as linear continuous covariate (a linear time effect is assumed, as supported by results seen for fingolimod [Radue 2012]). Fixed effects: Treatment is considered as a factor, and a treatment by time interaction is included, the analysis is adjusted for the following covariates: [redacted], [redacted] and normalized [redacted]. Subject is considered as random effect (intercept) and the within subject covariance matrix is assumed to follow a spatial power covariance structure of time (considering time as continuous covariate). Mean change at [redacted] ([redacted] and at [redacted] ([redacted] using least square means and difference between treatment arms based on contrasts (estimates, 95% CIs, corresponding p-value for treatment difference) will be provided.

[redacted] at [redacted] is analyzed in addition using a repeated measurement ANOVA model (MMRM) with fixed effects: Treatment, Visit ([redacted] as categorical variable and treatment by visit interaction. Covariates are considered as above. An unstructured covariance structure within subjects is considered. Mean change

by visit (and absolute differences in mean change between treatment arms are reported based on LS means (estimates, 95% CI, corresponding p-value for treatment difference).

9.8.4 Change from baseline in the and to

The change from baseline in at and will be analyzed on the FAS using a repeated measurements ANOVA model (MMRM), with treatment as factor, and time (visit) as categorical variable, a treatment by time interaction, and unstructured covariance matrix, assuming residuals follow a normal distribution. The analysis is adjusted for the following covariates:

Mean change by visit (using least square means) and difference between treatment arms by visit using a contrast (estimates, 95% CIs, corresponding p-value for treatment difference) will be provided by visit.

Change from baseline in per visit (at and is summarized descriptively based on observed data, as well as based on LOCF imputation, using the FAS and data will be listed.

9.8.5 Change from baseline in the to and to

Analyze in the same way as for change from baseline in to and to

9.8.6 at and

(at and separately at will be analyzed on the FAS using a logistic regression model [see Section 10.1.1] with treatment as factor, adjusting for the following covariates: Odds ratio for the treatment effect with 95% Wald CIs and corresponding p-value will be presented. Subjects with missing information are excluded from the analysis. A similar analysis is conducted for from to and

Descriptive statistics on s (will be presented by visit, and cumulative for up to each visit. A listing will be prepared using the FAS.

9.8.7 [redacted] at [redacted] and [redacted]

[redacted] up to each visit on the FAS is analyzed in the same way as for [redacted] [see Section 9.8.6], using logistic regression and descriptive statistics.

9.8.8 Proportion of [redacted]

[redacted] will be presented by visit (nominal and LOCF) in the subset of subjects with [redacted]. A listing will be prepared using the FAS.

9.8.9 Cumulative number of [redacted] to [redacted]

Analyzed in the same way as the main analysis of the secondary endpoint number of [redacted] up to [redacted] using a negative binomial regression model on the FAS [see Section 9.7.2.2].

Descriptive summary statistics of the number of [redacted] up to [redacted] will be presented in the same way as for the number of [redacted] up to [redacted] but using categories (anticipated: 0, 1–2, 3–5, 6+).

9.9 Analysis of other clinical efficacy variables

9.9.1 [redacted]

A two-sided stratified log-rank test with stratification variables (EDSS ≤ 3.5 versus EDSS > 3.5; DMT within last 2 years prior to randomization [Yes/No]) and number of [redacted] in the year prior to study entry, will be performed on the FAS. Hazard ratio and 95% CI based on a stratified cox regression will be provided. The model will include treatment as factor, and the same covariates as for the stratified log-rank test as stratification variables. In addition, an unstratified analysis is conducted using a log-rank test and Cox regression.

[redacted] will be summarized on the FAS in a table including, number of subjects with event / censored, Kaplan-Meier estimates (unstratified) and corresponding CI (anticipated time points: [redacted]), and median (as well as 25th and 75th percentiles) of the survival function together with CIs. A graphical display presenting a Kaplan-Meier plot (unstratified) going upward is presented including CIs at specific time points as well as the log-rank p-value, hazard ratio and CI from the main analysis.

[redacted] data will be listed using the FAS.

9.9.2 Absence of [redacted] to [redacted] and [redacted] ([redacted])

Absence of [redacted] to [redacted] and to [redacted] ([redacted]) are analyzed on the FAS as for absence of [redacted] by visit [see Section 9.8.6] using descriptive statistics and a logistic regression model adjusting for the following covariates: [redacted]
Subjects with missing outcome are included in the analysis and considered not meeting [redacted]

9.9.3 Change in [redacted]

The [redacted] score and change from baseline in [redacted] score by visit and treatment group will be summarized using descriptive statistics on the FAS. A separate by visit summary displaying only change from baseline up to EOT + 7 days will be provided. A listing of [redacted] / [redacted] scores will be produced on the FAS.

Change from baseline in [redacted] score will be analyzed on the FAS using an repeated measurements ANOVA model (MMRM) [see Section 10.3.2], with treatment, visit, and visit by treatment interaction as fixed effects and adjusting for the following covariates: [redacted] Least square means and 95% confidence intervals by visit will be presented using a graphical approach.

9.9.4 [redacted]

[redacted] at [redacted] as defined in Section 5.4.4.4, will be analyzed on the FAS using a logistic regression model [see Section 10.1.1] with treatment as factor, adjusting for the following covariates: [redacted].
Odds ratio for the treatment effect with 95% Wald CIs and corresponding p-value will be presented. Subjects with missing information are excluded from the analysis.

Descriptive statistics of [redacted] status [redacted] including its components [redacted] ill be presented at [redacted] and [redacted] using the FAS. A listing will be prepared.

9.9.5 [redacted]

[redacted] at [redacted] as defined in Section 5.4.4.5, will be analyzed in the same way as [redacted] status on the FAS [see Section 9.9.4].

9.9.6 Change in [redacted] to [redacted]

[redacted] as well as its subcomponents scores and changes from baseline will be descriptively summarized for each treatment group by visit on the FAS (observed and LOCF). Mean change (± SE) over time will be graphically presented by treatment group. A listing will be prepared.

A mixed effect model repeated measures (MMRM) analysis on change from baseline in [redacted] will be conducted following the same approach as for the FSIQ-RMS weekly symptoms score (main analysis) [see Section 9.7.1.2]. Results will be presented in a table and graphically by visit.

9.9.7 Change in the [redacted] to [redacted]

[redacted] scores and change from baseline will be descriptively summarized for each treatment group by visit on the FAS (observed and LOCF). Mean change (± SE) over time will be graphically presented by treatment group. A listing will be prepared.

A mixed effect model repeated measures (MMRM) analysis on change from baseline will be conducted following the same approach as for the FSIQ-RMS weekly symptoms score (main analysis) [see Section 9.7.1.2]. Results will be presented in a table and graphically by visit.

9.9.8 [redacted]

[redacted] will be summarized using frequencies and percentages by visit on the FAS.

[redacted] at [redacted] will be analyzed using a logistic regression model adjusting for the two stratification factors. i.e., [redacted] and the baseline [redacted] symptom score as covariates. Subjects with missing [redacted] at [redacted] will be considered to have not responded. Subjects with missing baseline [redacted] are not included in the analysis.

Cumulative distribution function (CDF) plots will be provided for change from baseline to [redacted] in [redacted] (observed and LOCF) to visualize the entire distribution of responses, i.e., the proportion of patients at each point along the scale score continuum who experience change at that level or lower, by treatment arm. This will allow for simultaneously evaluating [redacted]

9.9.9 Change from baseline by visit up to [redacted] in [redacted]

[redacted] nd change from baseline will be descriptively summarized for

each treatment group by visit on the FAS. Mean change (± SE) over time will be graphically presented by treatment group. A listing will be prepared.

A mixed effect model repeated measures (MMRM) analysis on change from baseline will be conducted for each of the [redacted] separately following the same approach as for the FSIQ-RMS weekly symptoms score (main analysis) [see Section 9.7.1.2]. Results will be presented in a table and graphically by visit.

9.9.10 Change from baseline in [redacted] to [redacted]

To evaluate the [redacted], summary statistics of the [redacted] (n, % per [redacted] as well as mean, SD, median and range for absolute values and changes from baseline) per visit by treatment group will be tabulated and change from baseline as well as % in each of the [redacted] will be graphically presented.

9.9.11 Change from baseline in [redacted]

To evaluate the [redacted] overtime, summary statistics of the [redacted] (n, % per [redacted] as well as mean, SD, median and range) per visit by treatment group will be tabulated and the change from baseline as well as % in each of the [redacted] will be graphically presented.

9.10 Analysis of safety variables

9.10.1 Adverse events

All AEs captured (i.e., from signature of informed consent up to EOS) are reported in the subject listings based on the SCR, treatment-emergent events are flagged.

Unless otherwise specified, the SAF is used for all analyses of AEs.

An overview of treatment-emergent AEs is presented, per treatment group, as the number and percentage of subjects having any AE, any severe AE, any drug-related AE, any AE leading to study drug discontinuation, any serious AE, any drug-related serious AE, or any fatal serious AE.

Also, AEs are summarized by system organ class and preferred term: Presenting, per treatment group, the number and percentage of subjects having any AE, having an AE in each primary SOC, and having each individual AE (preferred term). SOCs are sorted by descending order of frequency in the ponesimod arm. If the frequencies of SOCs are the same, alphabetical order is used. The same rule applies for preferred terms within SOC. The following summaries are presented:

- Treatment emergent AEs
- Treatment emergent AEs on Day 1
- Treatment emergent AEs on Day 1 of re-initiation, for subjects with at least one re-initiation

- Treatment emergent AEs leading to temporary interruption of study treatment
- Treatment emergent AEs considered to be related to study drug
- Treatment emergent AEs with fatal outcome
- AEs with onset prior to first study drug intake (based on the screened analysis set)
- Post-treatment AEs with onset after EOT + 15 days, for subjects in the post-treatment safety analysis set
- Treatment emergent AEs with onset prior to or at EOT + 1 day

Also, treatment-emergent AEs, treatment-emergent AE related to study treatment, and treatment-emergent AE leading to premature discontinuation of study treatment are summarized by preferred term: presenting, per treatment group, the number and percentage of subjects having any AE, and having each individual AE (preferred term). Preferred terms are sorted by descending order of frequency in the ponesimod arm. If the frequencies are the same, alphabetical order is used.

In addition, one summary table of treatment-emergent AEs by preferred term sorted by absolute percent difference (descending) between ponesimod and teriflunomide is presented.

For all treatment-emergent AEs, the number of occasions (i.e. recurrent events) and corresponding cumulative subject-years of observation, and – based on Poisson regression – the average annualized event rates are tabulated by preferred term sorted by descending rate ratio [see Section 10.1].

Treatment-emergent AEs by maximum intensity are summarized by presenting, per treatment group, the number and percentage of subjects having any AE, and having each individual AE (preferred term).

For the disclosure of the results to EudraCT and ClinicalTrials.gov (and inclusion in the clinical study report appendix), frequent treatment emergent non-serious AEs, i.e., preferred terms reported in ≥ 5% of subjects in at least one treatment group, are summarized as follows:

1) The number and percentage of subjects having any frequent non-serious AE, and having each individual non-serious AE (preferred term).
2) The overall number of frequent non-serious AEs (i.e., reported episodes), and the number and percentage of frequent AEs (preferred term). Here the denominator is the overall number of frequent non-serious AEs.

9.10.2 Deaths, other serious adverse events

9.10.2.1 Death

The number and percentage of subjects who died are summarized per treatment group, including the reported primary cause of death. A summary table is provided for all deaths

overall and split by time of death (treatment-emergent and more than 15 days after end of study treatment).

A separate listing including all deaths is provided based on the SCR; treatment-emergent deaths are flagged as such in that listing.

9.10.2.2 Serious adverse events (SAEs)

Unless otherwise specified, the SAF is used for all analyses of SAEs.

SAEs are summarized by SOC and preferred term. The following summaries are presented:

- Treatment emergent SAEs
- Treatment emergent SAEs on Day 1
- Treatment emergent SAEs on Day 1 of re-initiation, for subjects with at least one re-initiation
- Treatment emergent SAEs with fatal outcome
- Treatment emergent SAEs with fatal outcome considered to be related to study drug
- Treatment emergent SAEs considered to be related to study drug
- Treatment emergent SAEs leading to study drug interruption
- SAEs with onset prior to first study drug intake (based on the screened analysis set)
- Post-treatment SAEs with onset after EOT + 15 days, for subjects in the post-treatment safety analysis set.

In addition, treatment-emergent SAEs, treatment-emergent SAE related to study treatment, and treatment-emergent SAE leading to premature discontinuation of study treatment are summarized by preferred term.

For the disclosure of the results to EudraCT and ClinicalTrials.gov (and for inclusion in the clinical study report appendix), the overall number of treatment emergent SAEs (i.e., reported episodes), and the number and percentage SAEs (preferred term) are summarized. Percentages of subjects are based on the number of subjects in the safety analysis set and the percentages of events are based on the total number of events.

9.10.2.3 Treatment-emergent adverse events leading to study treatment discontinuation

The SAF is used for all analyses of (S)AEs leading to discontinuation. AEs and SAEs leading to premature discontinuation of study drug are summarized by SOC and preferred term. The following summaries are presented:

- Treatment-emergent AEs leading to premature discontinuation of study drug
- Treatment-emergent AEs leading to premature discontinuation of study drug on treatment day 1
- Treatment-emergent AEs leading to premature discontinuation of study drug on Day 1 of re-initiation, for subjects with at least one re-initiation

- Treatment-emergent SAEs leading to premature discontinuation of study drug
- Treatment-emergent SAEs leading to premature discontinuation of study drug on treatment day 1
- Treatment-emergent SAEs leading to premature discontinuation of study drug on Day 1 of re-initiation, for subjects with at least one re-initiation

In addition, treatment-emergent AEs leading to premature discontinuation of study drug are summarized by preferred term.

9.10.2.4 Other significant adverse events

9.10.2.4.1 Adverse events of special interest (AESIs)

For each of the AESI categories described in Section 5.5.9.1, the following information is presented: The number and percent of subjects with AESI is summarized by preferred term per treatment arm. In addition, the number and percentage of subjects having any event of that category, having any serious event, any fatal event or any event leading to premature discontinuation of study drug, the relative risk relative to teriflunomide (incl. 95% CI). The table will also include key estimates from a time to first event analysis (Kaplan-Meier estimates and hazard ratio along with 95% CI), the number of occasions (i.e., recurrent events) and corresponding cumulative subject-years of observation, and based on Poisson regression, the average annualized event rate and rate ratio with corresponding rate ratio incl. 95% CI. Corresponding time to first event (Kaplan-Meier) curves (going upwards) are graphically displayed as described in Section 10.4. Separate listings are prepared on the FAS.

These tables and graph are presented for the following subcategories:

- Treatment-emergent AESIs, based on the SAF for each AESI category separately;
- Hepatobiliary disorders / Liver enzyme abnormality AESI that occurred after start of study drug and up to and including EOT + 1 day

In addition, treatment-emergent AESI Effect on heart rate and rhythm related AEs are summarized by time of onset by presenting, per treatment group, the number and percentage of subjects having any AE, and having each individual AE (preferred term). Preferred terms are sorted by descending order of frequency in the ponesimod arm. This analysis is done

- By day of onset after study start, by groups: Day 1, Day 2, Day 3, Day 4, Day 5, Day 6, Day 7, Day 8, Day 9, Day 10, Day 11, Day 12, Day 13, Day 14, Day 15, Day 16, > Day 16

- By day of onset after any re-initiation of up-titration, for the following (daily) by groups: Day 1 of any re-initiation, Day 2 of any re-initiation… , Day 16 of any re-initiation

9.10.2.4.2 Major adverse cardiovascular events (MACE)

Based on the SAF, treatment emergent MACEs are summarized by presenting, per treatment group, the number and percentage of subjects having any MACE, and having an event of the MACE subcategories (cardiovascular death, non-fatal myocardial infarction, non-fatal stroke), and the relative risk relative to teriflunomide (including 95% CI). The table will also include key estimates from a time to first event analysis (Kaplan-Meier estimates and hazard ratio along with 95% CI), the number of occasions (i.e., recurrent events) and corresponding cumulative subject-years of observation, the average annualized event rate and corresponding Poisson rate ratio including 95% CI. Corresponding time to first event (Kaplan-Meier) curves are graphically displayed as described in Section 10.4. A listing is prepared.

9.10.3 Electrocardiography (ECG)

The following summary displays are provided by treatment group, all based on the SAF:

- Quantitative ECG results and absolute change from baseline, as described in Section 5.5.11.1, are summarized, by parameter and visit. Except for data summarized under 'Baseline', 'Day-15 follow-up' and 'Day-30 follow-up', only treatment-emergent results from nominal scheduled visits are included, premature EOT visit is mapped to a scheduled visit [see Section 11.3]. These are summarized descriptively and mean changes (with standard error) from baseline in heart rate are also presented graphically over time (plots starting with zero change at baseline);

- Quantitative ECG results and absolute change from pre-dose on Day 1, by time point (pre-dose, 1–12 hours post dose), are summarized descriptively and mean changes (with standard error) from pre-dose in heart rate are also presented graphically over time (plots starting with zero change at pre-dose), plots are only to include Hours 1–4;

- Descriptive summary statistics of quantitative ECG result and absolute change from pre-dose on Day 1 of first Re-initiation, by time point (pre-dose, 1– 12 hours post dose); Note: Data from re-initiations beyond the first are listed only;

- Number and percentage of subjects with any treatment-emergent PR, QTcF prolongations and HR outliers together with 95% CIs for proportions (Clopper-Pearson exact method) are presented overall and by visit. Percentages are based on the number of subjects with at least one treatment emergent result available for the corresponding parameter. Abnormalities are presented cumulatively, i.e., a subject with a treatment emergent QTcF result > 500 ms is also summarized in the category QTcF > 450 ms;

- Number and percentage of subjects with any post-dose PR, QTcF prolongations and HR outliers on Day 1 or Day 1 of Re-initiation together with 95% CIs for proportions (Clopper-Pearson exact method) are presented overall post-dose hours and by individual time point. Percentages are based on the number of subjects with at least one post-dose result available for the corresponding parameter and time point. Abnormalities are presented cumulatively;

- Number and percentage of subjects with any 3-hr post-dose PR, QTcF prolongations and HR outliers on Day 1, Week 12 Visit, and Day 1 of any re-initiation together with 95% CIs for proportions (Clopper-Pearson exact method). For this analysis the ECG results reported as 3-hours post-dose are considered, irrespective of actual time the ECG was performed. Percentages are based on the number of subjects with a result available for the corresponding parameter at 3-hours post-dose. Abnormalities are presented cumulatively;

- Number and percentage of subjects with prior morphological ECG findings, and new/pre-existing treatment-emergent morphological ECG findings (at any time during the respective periods, i.e., prior or treatment-emergent), by abnormality and group term following CDISC terminology are presented. New (as compared to any assessment prior to study drug start) findings are also presented by visit;

- Number and percentage of subjects with prior and new/pre-existing morphological ECG findings on Day 1 or on Day 1 of any re-initiation, by abnormality and group term following CDISC terminology are presented. New (as compared to any assessment prior to study drug start) post-dose hourly ECG findings are also presented and by individual time point;

- Number and percentage of subjects with prior and new/pre-existing treatment-emergent morphological ECG findings at Visit 4 (Week 2), by abnormality and group term following CDISC terminology.

All ECG data including QTc parameters are listed by treatment group and subject based on the SCR. All marked abnormalities are flagged and morphological findings are listed.

9.10.4 Vital signs and body weight

The following summary displays are provided by treatment group, all based on the SAF:

- Blood pressure and body weight including absolute change from baseline, as described in Section 5.5.10, are summarized, by parameter and visit. Except for ‘Baseline’, ‘Last on treatment’, ‘Day-15 follow-up’ and ‘Day-30 follow-up’, only treatment-emergent results from nominal scheduled visits are included; premature EOT visit is mapped to a scheduled visit [see Section 11.3]. These are summarized descriptively and mean changes (with standard error) from baseline in blood pressure are also presented graphically over time (plots starting with zero change at baseline);

- Blood pressure including absolute change from baseline on Day 1, by time point (pre-dose, 1–12 hours post dose), are summarized descriptively and mean changes (with standard error) from pre-dose are also presented graphically over time (plots starting with zero change at pre-dose);

- Descriptive summary statistics of blood pressure including absolute change from pre-dose on Day 1 of first Re-initiation, by time point (pre-dose, 1–12 hours post dose); Note: Data from re-initiations beyond the first are listed only;

- Number and percentage of subjects with any treatment-emergent blood pressure result meeting the criteria defined in Section 5.5.10.1. Percentages are based on the number of subjects with at least one treatment-emergent result available for the corresponding parameter;

- Number and percentage of subjects with post-dose blood pressure result on Day 1 and Day 1 of any re-initiation meeting the criteria defined in Section 5.5.10.1 are presented overall post-dose hours and by individual time point. Percentages are based on the number of subjects with blood pressure results available for the corresponding parameter and timepoint;

All blood pressure and body weight results are listed by treatment group and visit based on the SCR. Qualitative results ('normal' or 'abnormal') of body temperature and pulse rate are listed as collected in the eCRF, by treatment group and visit based on the SCR.

9.10.5 Laboratory tests

Descriptive summary statistics by visit (including protocol-mandated between visit assessments) and treatment group are provided for laboratory test results reported by the central laboratory and corresponding absolute changes from baseline, based on the SAF. Except for 'Baseline', 'Last on treatment', 'Day-15 follow-up' and 'Day-30 follow-up', only treatment-emergent results from nominal scheduled visits are included; premature EOT visit is mapped to a scheduled visit [see Section 11.3]. Data are displayed in SI units as provided by the central laboratory. In the same manner, mean absolute changes from baseline with their standard error are presented graphically (plots starting with zero change at baseline). Values from unscheduled visits are not included in by visit summaries but are included in the listings and in the summaries for marked abnormalities and liver function test abnormalities.

For lymphocyte counts, this table and plot are also provided on percent changes from baseline. The number of subjects with at least one treatment-emergent marked laboratory abnormality is summarized by treatment group, based on the SAF. The worst case in each direction is considered and summarized in all applicable categories, i.e., a subject counted in LLL is also counted in LL for a given parameter; also, a subject maybe counted in both the high and the low category for a given parameter. In addition, the number of subjects with treatment-emergent liver function test abnormalities is summarized by treatment

group. This tabulation is also repeated for abnormalities that occurred after start of study drug and up to and including EOT + 1 day. The denominator for percentages is the number of subjects with at least one post-baseline assessment.

A plot to evaluate Drug-Induced Serious Hepatotoxicity (eDISH plot) is provided based on the highest observed treatment-emergent ALT and total bilirubin values, considering both, central and local laboratory data. Individual subject's values expressed as × ULN are plotted on a log-log scatter plot (ALT on the horizontal axis, total bilirubin on the vertical axis); reference lines are drawn at Hy's law thresholds, i.e., at 3 × ULN for ALT and at 2 × ULN for total bilirubin. A footnote is presenting, for each treatment group, the number and percentage of subjects falling in each of the 4 quadrants defined by these reference lines. This eDISH plot is repeated for abnormalities that occurred after start of study drug and up to and including EOT + 1 day.

Safety laboratory hematology, clinical chemistry, urinalysis parameters planned as per protocol are also presented in subject data listings based on the SCR. All abnormalities and marked abnormalities including those at baseline are flagged in the listings.

Other laboratory parameters are presented in a separate listing on the SCR.

Further analyses related to lymphocytes are described in Section 9.13.

9.10.6 Pulmonary function tests - Spirometry

The following summary displays are provided by treatment group, all based on the SAF:

- Treatment-emergent FEV1, FVC, FEV1/FVC, %FEV1, and %FVC including absolute and percent change from baseline, as described in Section 5.5.13.1, are summarized, by parameter and visit. Except for 'Baseline', 'Last on treatment', 'Day-15 follow-up' and 'Day-30 follow-up', only treatment-emergent results from nominal scheduled visits are included; premature EOT visit is mapped to a scheduled visit [see Section 11.3]. These are summarized descriptively and mean changes from baseline are also presented graphically with their standard error over time (plots starting with zero change at baseline).

- A scatter plot of all measured treatment-emergent %FEV1 values is provided. Treatment groups are color-coded and linear regression lines (with time as explanatory variable) are drawn per treatment group. These regression lines are to visualize the course over time after an initial change in pulmonary function and for that reason, exclude measurements taken prior to study day 26. The regression equation is given in a footnote.

- Number and percentage of subjects with any occasion of treatment-emergent low spirometry values as defined in Section 5.5.13.1 are presented. Percentages are based on the number of subjects with at least one treatment-emergent result available for the corresponding parameter;

- In the subset of subjects experiencing a decrease of ≥ 200 mL or ≥ 12% in FEV1 (or FVC) from baseline to the last assessment on treatment, the number and percentage of subjects with the condition reversed, not reversed, or missing data on reversibility at the last follow-up result is provided;

- FEV1, FVC, FEV1/FVC, %FEV1, %FVC at the following three time points (for definition, see Section 5.5.1) are displayed graphically by box-and-whisker plots visualizing mean, median, interquartile ranges, SD, and outliers over time: Baseline, last assessment on treatment, and Day-15 follow-up assessment. All subjects with data available at all three time points are included. A corresponding table is included as well.

- Individual absolute change from baseline to last assessment on treatment versus absolute change from baseline to last follow-up assessment is graphically displayed for FEV1, FVC, %FEV1, and %FVC by means of scatterplots with treatment groups color-coded.

All spirometry results (best effort results) are listed by treatment group and visit based on the SCR.

9.10.7 Pulmonary function tests - DL_{CO}

The following summary displays are provided by treatment group, all based on the DL_{CO} sub-study safety analysis set:

- Treatment-emergent Hemoglobin corrected DL_{CO}, Percent predicted DL_{CO}, and DL_{CO} including absolute and percent change from baseline as described in Section 5.5.13.2, are presented, by parameter and visit. Except for 'Baseline', 'Last on treatment', 'Day-15 follow-up' and 'Day-30 follow-up', only treatment-emergent results from nominal scheduled visits are included; premature EOT visit is mapped to a scheduled visit [see Section 11.3]. These are summarized descriptively and mean changes from baseline are also presented graphically with their standard error over time (plots starting with zero change at baseline). The analysis is repeated based DL_{CO} measurements graded acceptable.

- A scatter plot of all measured treatment-emergent Percent predicted DL_{CO} values is provided. Treatment groups are color-coded, and linear regression lines are drawn per treatment group. This regression lines are to visualize the course over time after an initial change in pulmonary function and for that reason, excludes any measurements taken prior to study day 26. The regression equation is given in a footnote.

- Number and percentage of subjects with any occasion of treatment-emergent low Hemoglobin corrected DL_{CO}, Percent predicted DL_{CO}, and DL_{CO} values as defined in Section 5.5.13.2 are presented. Percentages are based on the number of subjects with

at least one treatment-emergent result available for the corresponding parameter. The analysis is repeated based DL_{CO} measurements graded acceptable.

- Hemoglobin corrected DL_{CO}, Percent predicted DL_{CO}, and DL_{CO} at the following three time points [for definition, see Section 5.5.13.2] are displayed graphically by box-and-whisker plots visualizing mean, median, interquartile ranges, SD, and outliers over time: Baseline, last assessment on treatment, and Day-15 follow-up assessment. All subjects with data available at all three time points are included. A corresponding table is included as well.

- Individual absolute change from baseline to last assessment on treatment versus absolute change from baseline to last follow-up assessment is graphically displayed for Hemoglobin corrected DL_{CO}, Percent predicted DL_{CO}, and DL_{CO} by means of scatterplots with treatment groups color-coded.

All DL_{CO} results (derived mean values per assessment) are listed by treatment group and visit based on the SCR.

9.10.8 Analysis of other safety variables

Other safety assessment results will be listed. This comprises the following assessments: Dermatological examination, OCT, and Ophthalmological Examination and Columbia Suicide Severity assessment.

9.10.8.1 Columbia Suicide ideation or behavior

Suicidal Ideation, Suicidal Behavior or Self-Injurious Behavior without Suicidal Intent will be listed for each patient, at each assessment visit on the SAF. Events occurring during the period up to EOT + 15 days (treatment-emergent period) will be flagged.

Number (%) of subjects with suicidal ideation (overall and by category), suicidal behavior (overall and by category), suicidal ideation score ≥ 4 or suicidal behavior and/or self-injurious behavior without suicidal intent will be tabulated by treatment group and visit, as well as up to EOT + 15 days (worst outcome). Analyses will be based on subjects in the SAF who have at least one eC-SSRS measurement during the respective visit or period available. The analysis will be repeated in the subgroup of patients with no baseline result available.

For subjects who have in addition a baseline eC-SSRS assessment available shifts from baseline (maximum score from pre-treatment recent history) to the worst reported outcome (the maximum score) up to EOT + 15 days will be tabulated by treatment group to demonstrate any changes in Suicidal Ideation and Suicidal Behavior score categories.

9.11 Analysis of quality of life variables

9.11.1 Change from baseline by visit up to Week 108 in SF-36v2™ Health Survey domain and component scores

Quality of life analysis based on the SF-36v2™ will be performed on the FAS.

The eight normative (t-scores) domain scores, PCS and MCS will be summarized descriptively over time (Baseline, Weeks 12, 24, 60, 84, 108) by treatment groups (observed and LOCF). Changes overtime will be graphically presented for each of the domains, PCS and MCS by treatment group (mean and ±SE). The health transition item (#2) will be summarized separately (in categories) by visit and a shift table from baseline to Week 108 (observed and LOCF) will be prepared.

9.12 Analysis of pharmacoeconomic variables

9.12.1 Change from baseline by visit up to Week 108 in WPAI:MS scores

Analyses will be summarized and listed on the FAS. WPAI scores and number of subjects currently employed by visit during the main study are summarized per treatment group. WPAI-MS outcome measures (absenteeism, presenteeism, overall work impairment and activity impairment) will be graphically presented over time using mean ± SE.

9.12.2 Health care resource utilization

Health care resource utilization will be summarized and listed on the FAS.

Descriptive summary statistics will be displayed for subjects with at least one and frequency of health care resource utilizations as well as length of stay (days) from baseline up to Week 108 for the following categories: all hospitalizations, hospitalizations due to AE, hospitalizations due to relapse, ICU admissions due to MS relapses, emergency room visits due to MS.

9.13 Analysis of pharmacodynamic variables

Descriptive analyses of lymphocyte data are described in Section 9.10.5. Only lymphocyte counts from central laboratory, irrespective of whether obtained at a scheduled or unscheduled visit, are included in any of the below describe analyses.

For each subject in the SAF, the

- Nadir (i.e., lowest) treatment-emergent lymphocyte value (up to EOT + 15 days)

is identified and summarized by means of frequency counts and percentages based on the categories specified in Section 5.8.

Lymphocyte counts and percent change from baseline in lymphocyte counts at the following time points [for definition, see Section 5.8] are summarized using descriptive

statistics and displayed graphically by box-and-whisker plots visualizing mean, median, interquartile ranges, SD, and outliers over time: Baseline, Last on-treatment, Day-15 follow-up (plus one display adding Day-30 follow-up in addition). Only subjects with data available at all three (four) time points are included. Corresponding descriptive data is tabulated.

Individual absolute and percent change from baseline in lymphocyte counts to last assessment on treatment versus absolute and percent change from baseline to last follow-up is graphically presented by means of scatterplots with treatment groups color-coded.

9.14 Analysis of pharmacokinetic variables

Will be detailed in a separate analysis plan.

10 GENERAL STATISTICAL METHODOLOGY

This section describes in general terms the statistical models and methods applied.

10.1 Statistical methodology for count data

Count data will be analyzed assuming data is negative binomially (NB) distributed.

A generalized linear model with NB distribution will be assumed.

t_j denotes the length of observation for subject j.
Y_j denotes the counts of interest for subject j during t_j.
μ_j denotes the mean of the NB distribution of Y_j.

The mean for the distribution of the ARR for subject j, denoted by μ_j / t_j, will be modeled by the following equation:

$$\log\left(\mu_j / t_j\right) = \mathbf{x}'_j\boldsymbol{\theta}, \textit{ i.e.,} \ \log\left(\mu_j\right) = \mathbf{x}'_j\boldsymbol{\theta} + \log\left(t_j\right), \text{ where}$$

$\mathbf{x}_j$ is the vector denoting study treatments and covariates for subject j
$\boldsymbol{\theta}$ is the vector of unknown fixed-model parameters.

The SAS code for the NB model with 99% CI is as follows (considered "draft" until fully validated at analysis stage):

```
proc genmod data=ADREL;
     class Treat Strat Covar;
     model COUNTS = Treat Strat Covar / dist=negbin link=log
     offset = offset;
      lsmeans Treat / cl exp alpha = 0.01 OM;
     estimate 'Ponesimod - Teriflunomide' ‖ Treat 1 -1 / exp alpha = 0.01;
run;
```

The offset variable used will be specified per analysis (e.g. for ARR it is the log-transformed observation time, for total T1 lesions it is the log-transformed number of available MRI scans, …).

The 'LSMEANS' statement will output the mean estimates for each of the treatment arms with 99% Wald CIs; the option OM ensures that the mean is derived for categorical covariates with weights as per observed marginal proportions.

The 'ESTIMATE' statement will output the rate ratio of the treatment effect with 99% Wald CIs. The direction of the estimate statement (governed by '1 – 1' or '−1 1') is chosen such that the rate ratio relative to teriflunomide is provided (with rate ratio < 1 indicating ponesimod is better).

To obtain 95% CIs the option alpha = 0.05 is used in the LSMEANS and ESTIMATE statements.

The number of covariates / stratification variables included in the model is different from analysis to analysis, and these are described in the respective analysis sections [see Section 9]. Interaction terms might be included in the model, depending on the analysis.

If the NB distribution is not considered to be appropriate (e.g. due to non-convergence), other distributions such as the Poisson and zero-inflated Poisson will be explored. For the Poisson distribution, a Poisson regression is conducted with model equation identical to the one for the negative binomial regression. The SAS code is as follows (considered "draft" until fully validated at analysis stage):

```
proc genmod data=ADREL;
     class Treat Strat Covar;
     model COUNTS = Treat Strat Covar / dist=poisson link=log
     offset = offset;
      lsmeans Treat / cl exp alpha = 0.01;
     estimate 'Ponesimod - Teriflunomide' ‖ Treat 1 -1 / exp alpha = 0.01;
run;
```

10.1.1 Imputation methods for count data

Multiple imputation for count data will be applied for sensitivity analyses of the primary and secondary count data.

A multiple imputation approach following Keene 2014 is followed, count data being assumed to follow a negative binomial distribution. Implementation will be based on a SAS Macro publicly available [see Roger 2017].

The general multiple imputation is a 3-step approach:

(i) Create n_{impute} imputed datasets. Details are specified below in the imputation algorithm.

(ii) Analyze each imputed dataset using a negative binomial model as for the corresponding main analysis. (Details such as response, covariates, and offset are specified per analysis).

(iii) Derive multiple imputation inferential statistics by combining step (ii) results using Rubin's rules.

In step (iii) the rate ratio with corresponding CI (99% and/or 95%) will be derived.

For creation of the imputed datasets in step (i) the following imputation algorithm will be followed:

Imputation algorithm for negative binomial count data

1. Fit a bayesian log-linear negative binomial regression model on non-missing data,
   - For imputation between EOS and Week 108 fit to count data up to EOS; for imputation between EOT+7 and Week 108 fit to count data up to EOT + 7;
   - Use logarithm of time (in years) as an offset variable, treatment as a factor and adjust for covariates as for the main analysis;
   - Use non-informative priors for the model parameters (Default option for proc genmod with bayes statement);
     – Generate samples for the model parameters (linear regression parameters β and dispersion parameter k) from the posterior distribution using MCMC (seed = 24092010).

       The following initial MCMC specifications will be used: 1000 burn-in iterations (nbi = 1000), 20000 iterations after burn-in (nmc = 20000), with a thinning of 5 (thin = 5). This yields a total of 4000 posterior draws (nmc ÷ thin), which corresponds to the number of multiple imputations (nimpute = 4000). The MCMC specifications, may be adjusted based on diagnostic checks for the Bayesian model.
2. Impute the count data for the missing follow-up time conditional on the follow-up information available based on the parameter estimates samples from step 1, under the considered assumption (MAR, copy reference, jump to reference, or delta adjustment). The conditional distribution follows also a negative binomial distribution; see Keene 2014 for details (seed = 30112013).
3. Derive the overall imputed count variable by adding up the observed count data (e.g., confirmed relapses up to EOS for imputation between EOS and Week 108) and the simulated count data for the missing follow-up time (from step 2).
4. Repeat this n_{impute} times to obtain n_{impute} complete datasets.

Imputation based on baseline covariates only

Imputation based on baseline covariates only (required for imputation of CUALs from baseline to Week 108 for subjects without any post-baseline MRI scan) can be implemented with the above algorithm as follows: Subjects with missing information are included in the input dataset with count variable of zero and time variable of zero. Interpretation is that no events (e.g., CUALs) have been observed since baseline for zero follow-up time. The imputation algorithm steps 2–3 will impute the counts for the entire time period from randomization to Week 108 which is the missing follow-up time in this case. The conditional distribution simplifies in this case to the marginal distribution and thus imputation will be based on baseline covariates only [see Keene 2014]. Note that in step 1 of the imputation algorithm, these zero follow-up time records will be ignored when fitting the Bayesian negative binomial regression model as the offset (log of the follow-up time) does not exist.

10.2 Statistical methods for binary data

10.2.1 Logistic regression

A logistic regression model is considered for binary response variables separately by visit. The model is applied, e.g., to absence of Gd+ T1 lesions by visit. The model includes treatment as factor and covariates as specified in the respective analysis section.

Specifically, the model with the structure below is used:

$$logit(p) \sim \beta_0 + \beta_1 \text{TREAT} + \beta_2 \text{COVAR}$$

Odds ratio estimates of the treatment effect (ponesimod vs teriflunomide) and 95% Wald CIs as well as the corresponding p-value are derived. The SAS code for this model is as follows (considered "draft" until fully validated at analysis stage):

```
proc logistic data=;
     class Treat Strat Covar;
     model BINARY_RESPONSE = Treat Strat Covar;
     oddsratio Treat /cl=wald;
     ods output OddsRatiosWald = oddsr;
     lsmeans trt / ilink cl oddsratio OM;
run;
```

10.3 Statistical methods for continuous data

10.3.1 ANOVA / ANCOVA model

An analysis of variance (ANOVA) model is considered for continuous responses that are expected to be approximately normally distributed. Treatment is included as a factor. Adjustment for additional covariates, as well as interaction terms might be included, depending on the analysis, in which case it is usually referred to as analysis of covariance (ANCOVA) model.

The SAS code for this model is as follows (considered "draft" until fully validated at analysis stage):

```
proc glm data= ;
     class Treat Strat Covar;
     model Response = Treat Strat Covar ;
     lsmeans Treat / cl alpha = 0.05 OM;
run;
```

10.3.2 Mixed effects repeated measurements model

A general mixed model for repeated measurements (linear mixed effects model) is considered for continuous responses that are expected to be approximately normally distributed and measured at scheduled visits over the course of the study.

Mixed effect repeated measurement model (MMRM) by visit data

The model includes fixed effects for treatment, visit as categorical covariate, the stratification variables as well as other categorical covariates and the interaction between treatment and visit. Additional covariates may be added. An unstructured covariance matrix for within-subject errors is applied. Treatment effect at a specific time point inferential statistics are derived based on contrasts and t-distribution, overall fixed effect treatment effect is assessed based on Type III effects.

The SAS code for this model is as follows (considered "draft" until fully validated at analysis stage):

```
Proc mixed data=;
   class subject Treat Visit Strat Covar;
   model Response = Treat Visit Treat*Visit Strat Covar / noint DDFM = KR solution;
   repeated Visit /subject=subject type=un;
   lsmeans Treat*Visit / pdiff CL OM;
run;
```

Mixed effect model with time as continuous variable

A mixed model for repeated measurements similar to the model above but considering time as a continuous variable and assuming a linear time effect is considered. Fixed effects are treatment, time (continuous) and treatment by time interaction. Further covariates can be added to the model. Subject is considered as random effect with spatial power correlation function (generalization of the autoregressive order one structure for unequally spaced data) of time (continuous) for the within-subject errors. This model can be applied using the following SAS code:

```
Proc mixed data=;
   class subject Treat Strat;
   model Response = Treat Time Treat*Time Strat Covar / noint DDFM = KR solution;
   random Int / subject = subject type = un;
   repeated / subject=subject type= SP(POW)(Time);
   lsmeans Treat / cl diff at time = 108 OM; *Time presented in Weeks;
run;
```

10.3.3 Rank ANCOVA

A rank analysis of covariance will be considered for continuous responses that are not approximately normally distributed. The non-parametric ANCOVA procedure incorporates three steps:

1. Transformation of baseline and post-baseline change values for all patients (regardless of treatment groups) to standardized ranks (i.e., ranks divided by the number of patients ranked plus 1, mean ranks in case of ties). For a stratified approach standardized ranks are obtained in each stratum separately.

2. Fit linear regression of post-baseline standardized ranks (by stratum for stratified analysis) on baseline standardized ranks and determine residuals.

3. Application of the Wilcoxon-Mann-Whitney test to these residuals (or van Elteren test for a stratified analysis) to test null hypothesis of no difference in location between treatment arms against the alternative of a difference in location.

SAS code for the analysis:

```
PROC RANK DATA=dataset TIES=MEAN NPLUS1 OUT=stdrank;
 BY strata; *optional for stratified analysis;
 VAR base outcome;
RUN;

PROC GLM DATA=stdrank;
 BY strata; *optional for stratified analysis;
 MODEL outcome = base / PREDICTED;
 ODS OUTPUT PredictedValues = pred;
 ID treat;
RUN;

* Step 3 for unstratified analysis;
PROC NPAR1WAY DATA=pred WILCOXON;
 CLASS treat;
 VAR residual;
RUN;

* Step 3 for stratified analysis;
PROC FREQ DATA=pred SCORES=MODRIDIT;
 TABLES strata*treat*residual / CMH2;
RUN;

* base represents the Baseline value of the endpoint;
* outcome represents the post-Baseline endpoint;
* strata represents the stratification variable(s);
* treat represents the treatment group;
```

10.4 Statistical methods for time-to-event data

The analysis of time-to-event data are conducted using Kaplan-Meier estimates of events over time (including graphical representation), stratified log-rank test and Cox proportional hazard model.

10.4.1 Time to Event and log-rank test

Estimates of the event rate are obtained for each treatment group using the Kaplan-Meier method as implemented in SAS Proc Lifetest. The graphical representation follows the recommendations from Pocock [Pocock 2002]. Two-sided CIs at specific time points are constructed, with confidence limits calculated using Greenwood's formula for the estimate of the standard error. Median time to event (as well as 25^{th} and 75^{th} percentiles) for each group are provided with the corresponding two-sided CIs calculated using the method of Brookmeyer [Brookmeyer 1982].

The stratified log-rank test is conducted with SAS Proc Lifetest where the STRATA statement includes the specified stratification variables and the GROUP option includes the treatment variable (treat). The TIME statement includes a variable with times to event

(time) and an indicator variable for right censoring (censor) with 1 representing censoring.

```
Proc lifetest data= method=KM;
   time survtime*censor(1);
   strata STRAT COVAR / group=treat diff=control('Teriflunomide');
run;
```

10.4.2 Cox proportional hazard model

SAS Proc Phreg is used to estimate the hazard ratio, the p-value associated with the treatment effect hazard ratio and the CI of the hazard ratio (using Wald-based methods). The Cox regression model can be implemented using the following code:

```
Proc phreg data =;
  class treat (ref='Teriflunomide')
        Strat;
  model TimeTo*censor(1)= treat /risklimits ties=exact;
  strata Strat;
run;
```

An investigation into the assumption of proportional hazards for treatment is performed informally using a plot of the complementary log-log of the survival against the log of time (for each group). If the hazards are proportional, the lines should be approximately parallel. A plot of the Schoenfeld residuals against time and log(time) is also constructed, including a Loess curve [Collett 1994]. Departure from a horizontal line indicates violations of the proportional hazard assumption. A test of a non-zero slope is also derived based on traditional linear regression model. Since certain types of non-proportionality are not detected by the test, both the test and the plot are used to assess the non-proportionality assumption.

11 GENERAL DEFINITIONS AND DERIVATIONS

11.1 Dates, times and days

Study treatment start date see definition in Section 5.3.1.

End-of-Treatment (EOT) / Study treatment end date, defined as the latest study treatment end date as recorded on the Study Drug Log (SDL) eCRFs [see Section 5.3.1].

End-of-Study (EOS) date defined as the 'Date of End of Study' collected on the end of study eCRF, unless a subject is lost to follow-up or dies, in which case 'Date of Last Contact' or 'Date of Death' from Death eCRF are used [see Section 5.1.6].

Screening date defined as the visit date of the last available screening visit, i.e., date of Visit 1A for re-screened subjects, date of Visit 1 for all others.

Study Day refers to the number of days elapsed since randomization date plus 1 (e.g., Study Day 1 is the day of randomization). For dates prior to randomization, study day is

the negative number of days elapsed between the date under consideration and the day of randomization. Therefore, the study day is always different from 0.

For efficacy analyses also referred to as 'Day'.

Treatment Day refers to the number of days elapsed since study treatment start date plus 1 (e.g., Treatment Day 1 is the day of study treatment start). For dates prior to study treatment start, treatment day is the negative number of days elapsed between the date under consideration and the day of study treatment start. Therefore, it is always different from 0.

For safety analyses also referred to as 'Day'.

11.2 Analysis period definitions

11.2.1 Efficacy study periods

The 'Efficacy Study Period' of primary interest in this study is the period from randomization up to the EOS date, defined as: [Randomization; EOS].

For sensitivity efficacy analyses an 'Efficacy Treatment Period' is defined as: [Randomization; EOT + 7 days].

11.3 Re-assignment of premature EOT visits / Visit windowing

For all parameters unless stated below: For subjects who discontinued study drug prematurely with available premature EOT visit, the visit may be re-assigned to a scheduled visit as follows:

- If a result is missing for a scheduled visit, it can be replaced with a result from a premature EOT visit available in the respective visit window [see Table 9].

Table 9 Premature end of treatment remapping

Visit window	Nominal value Day	Lower limit for Day	Upper limit for Day
Week 2	15	2	22
Week 4	29	23	70
Week 12	85	71	126
Week 24	169	127	210
Week 36	253	211	294
Week 48	337	295	378
Week 60	421	379	462
Week 72	505	463	546
Week 84	589	547	630
Week 96	673	631	714
Week 108	757	715	Open end

Day refers to treatment day, i.e., days from study treatment start.

MRI: For MRI measurements visit windowing of premature EOT visits is conducted using ± 3 months around the target day (Week 60: Day 330 to Day 512; Week 108: ≥ Day 665) to account for the less frequent assessment schedule.

Patient reported outcomes collected on electronic device: For patient reported outcome data collected via an electronic device visit windowing is applied as a general process. Rationale is that incorrect visit labels are not corrected via a query process whereas the questionnaire date is tracked automatically by the device therefore, the latter is considered to provide most reliable information. Questionnaires based on dates will be assigned to the visits based on Table 10. If multiple questionnaires are available, the one closest to target is selected.

Physician reported questionnaire and other composite assessment data [redacted] Nominal visits are considered. If missing at the nominal visit, data from a premature end of treatment visit (if not available or appropriate then from an unscheduled visit) is assigned to the visits based on Table 10. If multiple unscheduled assessments qualify the one closest to target is selected.

Table 10 Visit windowing for questionnaire and psychometric assessment data

Visit window	Nominal value **Day**	Lower limit for **Day** (Week)	Upper limit for **Day** (Week)
Baseline	-1	Open end	1
Week 12	85	2 (W0+)	126 (W18)
Week 24	169	127 (W18^{+})	294 (W42)
Week 60	421	295 (W42^{+})	504 (W72)
Week 84	589	505 (W72^{+})	672 (W96)
Week 108	757	673 (W96^{+})	Open end

Day refers to study day, i.e., days from randomization date. Wx+ = after week x.

11.4 Summaries by visit

Visit based safety and PD assessments are generally summarized according to the nominal visit. Except for Visit 14 (EOT/Week 108) in subjects prematurely discontinuing study treatment, Visit 15 (FU1, Day-15 follow-up assessment), and Visit 16 (FU2, Day-30 follow-up assessment). Those are derived following the rules below:

- Premature EOT visits are re-assigned following the visit windowing described in Section 11.3
- Day-15 follow-up assessment: Any assessment within EOT date + 8 days and EOT date + 22 days. If multiple assessments fall into that period, the one corresponding to the regular FU1 visit (Visit 15) is selected, if none of the multiple assessments corresponds to that visit, the one closest to EOT date +15 days is selected, if there are two closest assessments, the later one is selected.
- Day-30 follow-up assessment: Any assessment within EOT date + 23 days and EOT date + 37 days. If multiple assessments fall into that period, the one corresponding to the regular FU2 visit (Visit 16) is selected, if none of the multiple assessments corresponds to that visit, the one closest to EOT date + 30 days is selected, if there are two closest assessments, the later one is selected.

In addition, the 'Last on-treatment' assessment is flagged. This is defined to be the latest assessment prior to or on EOT date +1 day.

Summaries by visit are presented by Baseline (derived), and further scheduled nominal visits up to Week 108 (irrespectively if conducted during PTOP or not), Last on-treatment, and derived Follow-up visits (FU1, FU2). Generally, by visit tables are tabulating the baseline and nominal and derived visits as follows:

- Baseline (Note: This is not a nominal visit but the derived baseline; see Section 5.5.1)
- Visit 4 - Week 2

- Visit 5 - Week 4
- Visit 6 - Week 12
- Visit 7 - Week 24
- Visit 8 - Week 36
- Visit 9 - Week 48
- Visit 10 - Week 60
- Visit 11 - Week 72
- Visit 12 - Week 84
- Visit 13 - Week 96
- Visit 14 - Week 108
- Last on-treatment (derived; see Section 5.5.1)
- Day-15 follow-up (derived)
- Day-30 follow-up (derived)

Premature EOT visits are mapped to and summarized within a scheduled visit following the window approach described in Section 11.3; they will not be summarized as "Visit 14 - Week 108" unless they are mapped to this visit as per approach described in Section 11.3.

As per protocol, WBC and total lymphocyte counts will be assessed every 4 weeks up to Week 24 (i.e., there are 3 protocol mandated between-visit assessments for those tests at Week 8, Week 16, and Week 20 when no regular visit is scheduled). Likewise, liver tests (ALT, AST, INR, alkaline phosphatase and total bilirubin) will be assessed in-between visits at Weeks 6, 8, 10, 14, 16, 18, 20, and 22 in all countries except Canada, the United States, and Mexico. In Canada, the United States, and Mexico there are only 3 protocol mandated between-visit assessments for the liver function tests at Week 8, Week 16, and Week 20.

These assessments are considered scheduled and are included in corresponding by-visit tabulations labeled as follows:

- Additional Visit - Week 6
- Additional Visit - Week 8
- Additional Visit - Week 10
- Additional Visit - Week 14
- Additional Visit - Week 16
- Additional Visit - Week 18
- Additional Visit - Week 20
- Additional Visit - Week 22

Visit-based efficacy, quality of life or pharmaeconomic assessments are summarized according to the nominal visit up to EOS, irrespective if conducted during PTOP or not:

- Baseline (Derived)
- Visit 6 - Week 12
- Visit 7 - Week 24
- Visit 8 - Week 36
- Visit 9 - Week 48
- Visit 10 - Week 60
- Visit 11 - Week 72
- Visit 12 - Week 84
- Visit 13 - Week 96
- Visit 14 - Week 108

Premature EOT visits are mapped to and summarized within a scheduled visit following the window approach described in Section 11.3; they will not be summarized as "Visit 14 - Week 108" unless they are mapped to this visit as per approach described in Section 11.3.

11.4.1 LOCF imputation

LOCF imputation by visit is applied for several variables and is derived as follows: If a scheduled visit is missing or the result at a scheduled visit is missing, the outcome is imputed with the last available post-baseline result up to the upper limit of the visit window [see Section 11.3]. Baseline observations are not carried forward.

11.5 Listings of safety data for subjects with deviations from study treatment schedule

For safety data analysis based on the safety analysis set, subjects who - in error - were treated with both study treatments (teriflunomide and ponesimod) are summarized under the treatment group the subjects were exposed to for the longest time [see Section 7.1.4]. In order to individually review and – as needed – discuss their safety results, safety data listings are repeated in this subset of subjects.

Likewise, listings of adverse events, blood pressure and ECG results are provided for subjects who resumed ponesimod maintenance treatment without up-titration after an interruption of > 3 days and for subjects who deviated from planned up-titration regimen.

12 HANDLING OF MISSING/INCOMPLETE DATE AND TIME FIELDS

In the following, 'lower limit' and 'upper limit' refer to the minimum or maximum, respectively, of a possible date. For example, if the day is missing, the lowest limit is the first day of the given month and the upper limit is the last day of the given month. If the day and month are missing, the lower limit refers to the first day of the given year and the upper limit to the last day of the given year. *The earliest and the latest dates refer to the first or last date, respectively, when ordered in sequence.*

12.1 Previous / Concomitant therapies

For previous and concomitant therapies with missing or partial start and end dates the following rules for assignation to previous therapies and study concomitant therapies are applied [see Table 11].

Table 11 Handling of missing or partial start and end dates for previous/concomitant therapies

eCRF page	Start date	End date	Previous / Concomitant
Prev*	Unless start date clearly after/on STS (Start date either: prior to STS date, partial with either lower or upper limit prior to STS date, or missing)	Unless end date after/on STS (Either prior STS, partial with either lower or upper limit prior STS, or missing)	Previous
Con**	Unless start date clearly after/on STS (as above)	Prior to STS date; or Partial with upper limit prior to STS date;	Previous
Con**	Start date confirmed prior STS (Start date either: prior to STS date, partial with upper limit prior to STS date)	End date on STS date and 'Ongoing at start of Treatment' ticked 'No'; or Partial with lower limit prior to STS date and upper limit after/on STS date and 'Ongoing at start of Treatment' ticked 'No';	Previous
Any	All other cases not listed above	All other cases not listed above	Study concomitant

STS = Study treatment start; For subjects randomized but not treated the randomization date is used instead of the STS.
* Includes 'Previous Medications' or 'MS Specific Treatment History Log' eCRF: On these forms as per CRF completion guidelines only medications that stopped prior to signature of informed consent are to be recorded.
** Includes all forms collecting therapies apart from 'Previous Medications' or 'MS Specific Treatment History Log eCRF.

To derive treatment concomitant therapies in case of missing or partial start dates for a study concomitant therapy the following rules apply:

- Start date is available and after EOT, or start date is partial with lower limit after EOT: not considered treatment concomitant irrespective of end date
- Start date is available and prior or on EOT, start date is partial with lower limit prior or on EOT, or start date is missing: considered treatment concomitant irrespective of end date
- Subjects randomized but not treated are not considered to have any treatment concomitant medications

To further flag study concomitant therapies as 'ongoing at study treatment start' or 'started during study drug administration' medications the following rules apply:

- Study concomitant therapies with missing start date or partial start dates potentially falling into the treatment period (i.e., lower limit prior or on EOT) are considered to

have been started during study drug administration unless 'Ongoing at start of Treatment' ticked 'Yes' then it is considered ongoing at study treatment start.

12.1.1 Previous / Concomitant therapies date imputation

The following imputation of partial dates is made:

Previous therapies: Impute partial start date with lower limit, and partial end date with minimum of upper limit and study treatment start date − 1. Missing dates are not imputed.

Study concomitant therapies ongoing at study treatment start: Impute partial start date with lower limit and partial end date with upper limit. Missing dates are not imputed.

Study concomitant therapies not ongoing at study treatment start: Impute partial start date with maximum of lower limit and study treatment start date, impute partial end date with upper limit. Missing dates are not imputed.

12.2 Relapse

Type of date/time	Date/time is incomplete	Date/time is missing
Relapse start date (collected on 'Relapse Summary' eCRF)	Maximum of lower limit and randomization date Unless upper limit is prior to randomization date, then upper limit	Randomization date

eCRF = electronic case report form.

12.3 Adverse event onset and death dates

The following imputation rules are applied for (partially missing) AE onset dates and (partially missing) dates of death:

- Onset day missing: If month and year is clearly on or after study treatment start (short: STS, the date of first study drug intake) month and year, and clearly before or on the month and year of last intake date + 15 days, consider the event as treatment-emergent. With regard to the date imputation, the following rules are applied:
  - If the record's origin is the First-Dose eCRF and month and year correspond to month and year of STS, impute onset date and time as the date and time of the first study drug intake.
  - If the record's origin is the First-Dose eCRF and month and year are clearly after STS month and year, but a date of re-initiation with corresponding month and year is documented, impute onset date and time as the date and time of this re-initiation study drug intake (in case more than one study drug re-initiations are

documented in this month and year, impute with the date and time of the earliest of those).
- If the record's origin is the First-Dose eCRF and month and year are clearly after STS month and year, but no date of re-initiation with corresponding month and year is documented, the onset date is imputed to the 1st day of the month and year given and time to 00:00.
- If the record's origin is the main CRF and month and year is clearly on or after STS month and year, the onset date is imputed as the maximum of (date of Treatment day 2, 1st day of the month and year given) and time is imputed to 00:00.
- If event onset month and year is clearly prior to the STS month and year, the onset date is imputed to the last day of the given month (i.e., 28th, 29th, 30th, or 31st depending on month) and year and time is imputed to 00:00. This imputation is done irrespective of eCRF origin (First-Dose or Main).

- Onset day and month missing: If the year is the same year as the year of STS or later, and if the year is prior to or in the same year as the last intake date + 15 days, consider AE as treatment-emergent. With regard to the date imputation, the following rules are applied:
  - If the record's origin is the First-Dose eCRF and the year corresponds to the year of STS, impute onset date and time as the date and time of the STS.
  - If the record's origin is the First-Dose eCRF and the year is clearly after the year of STS, but a date of re-initiation with corresponding year is documented, impute onset date and time as the date and time of this re-initiation study drug intake (in case more than one study drug re-initiations are documented in this year, impute with the date and time of the earliest of those).
  - If the record's origin is the First-Dose eCRF and the year is clearly after the year of STS, but no date of re-initiation with corresponding year is documented, the onset date is imputed to January 1st of the given year and time to 00:00.
  - If the record's origin is the main CRF and year is clearly on or after the year of STS, the onset date is imputed as date of maximum of (Treatment day 2, January 1st of the given year) and time is imputed to 00:00.
  - If the event onset year is clearly prior to the year of STS, the onset date is imputed to 31-December of the given year and time is imputed to 00:00. This imputation is done irrespective of eCRF origin (First-Dose or Main).

- Missing onset time is imputed to
  - the time of the STS, if the onset date equals the STS.
  - the time of the study drug intake at re-initiation, if the onset date equals the date of a documented study drug re-initiation.
  - 00:00, for any other onset date.

- Onset date is completely missing: Consider as treatment-emergent.
  The onset date and time is imputed as date and time of the STS (if death date or if AE onset date with record's origin is the First-Dose eCRF), if AE onset date with record origin in the main CRF, the onset date is imputed as Study Day 2.

12.4 Study treatment start and EOT

The following imputations of study treatment start and EOT are considered for assigning safety events and assessments to the treatment-emergent period and used for deriving efficacy variables with definitions requiring EOT date. It is not considered for derivation of exposure variables. Details for assigning safety events and assessments to the treatment-emergence period, as described in Section 5.5.1, apply.

Type of date/time	Date/time is incomplete	Date/time is missing
Study treatment start date	Maximum of lower limit and randomization date.	Randomization date
Study treatment start time	See missing.	0:00 or randomization time if study treatment start date (after imputation) is equal to randomization date.
EOT	Earliest between treatment start date + 756 days (108 weeks), the upper limit, EOS date, and Death date.	Earliest between treatment start date + 756 days (108 weeks), EOS date and Death date.

EOS = End of Study; EOT = End of Trial.

13 LIST OF SUMMARY TABLES, LISTINGS AND FIGURES

The list of summary tables, listings, and figures is stored in a separate document in Excel format.

14 REFERENCES

[Bretz 2009] Bretz F, Maurer W, Brannath W, et al. A graphical approach to sequentially rejective multiple test procedures. Statist Med. 2009;28(4):586-604.

[Brookmeyer 1982] Brookmeyer R, Crowley JA. CI for the median survival time. Biometrics. 1982;38:29-41.

[Collett 1994] Collett D. Modelling survival data in medical research. London: Chapman & Hall; 1994.

[Cotes 1993] Cotes JE, Chinn DJ, Quanjer PH, Roca J, Yernault JC. Standardization of the measurement of transfer factor (diffusing capacity). Report Working Party Standardization of Lung FunctionTests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. Eur Respir J. 1993;16(Suppl.):41-52.

[Cuzick 2005] Cuzick J. Forest plots and the interpretation of subgroups. The Lancet. 2005;365(9467):1308.

[Dmitrienko 2010] Dmitrienko A, Tamhane AC, Bretz F. Multiple testing problems in pharmaceutical statistics. CRC Press; 2010.

[Keene 2014] Keene, O. N., Roger, J. H., Hartley, B. F., & Kenward, M. G. Missing data sensitivity analysis for recurrent event data using controlled imputation. Pharm Stat. 2014;13(4):258-264.

[Macintyre 2005] Macintyre N, Crapo RO, Viegi G, et al. ATS/ERS Task Force. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. Eur Respir J. 2005;26(4):720-35.

[Pocock 2002] Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. The Lancet. 2002;359:1686-9.

[Radue 2012] Radue EW, et al. "Impact of fingolimod therapy on magnetic resonance imaging outcomes in patients with multiple sclerosis." Arch Neurol. 2012 Oct;69(10):1259-1269.

[Reilly Associates 2002] WPAI coding; Coding Rules for Self-Administration. Margaret Reilly Associates Inc; 2002. Downloaded from http://www.reillyassociates.net/ WPAI_Coding.html on 21 Nov 2016.

[Roger 2017] Roger, J. Reference-based MI for Negative Binomial discrete data – SAS macros. Downloaded from https://missingdata.lshtm.ac.uk/2017/04/07/reference-based-mi-for-negative-binomial-discrete-data/ on 15 Nov 2018.

[Ware 2000] Ware JE, Kosinsky M, Dewey JE. How to score version 2 of the SF-36 Health Survey. Lincoln RI: QualiMetric Incorporated, 2000.

15 ATTACHMENTS

Attachment 1 Protocol Deviation Code List

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Document Revision History:

Version	Date	Reason
1	23Mar15	New
2	25Jul16	Updated to: - reflect last study protocol version - include new study protocol deviations covering findings reported by CRAs as well as issues detected during data review
3	28Oct16	- New PD codes were added to include monitor-reported deviations - Assessment-related PDs were split into specific important / not important PDs - PDs were reworded and/or modified to improve clarity and applicability - Yes/No (important PDs) was reconsidered for specific PDs - PDs were added or reworded to consider the post-treatment observation period (PTOP)
4	08Jun18	- New PD codes were added to include monitor-reported deviations - Assessment-related PDs were split into specific important / not important PDs - PDs were reworded and/or modified to improve clarity and applicability - Yes/No (important PDs) was reconsidered for specific PDs
5	22Feb19	- New PD codes were added to further characterize deviations of blinding procedures - Assessment-related PDs were split into specific important / not important PDs - PDs were reworded and/or modified to improve clarity and applicability - Yes/No (important PDs) was reconsidered for specific PDs
6	06May19	- Yes/No (important PDs) was reconsidered for specific PDs - PD_MM.043 was split into 3 codes specific for FSIQ-RMS, MRI and EDSS
7	03June19	- PDs instructions were reworded and/or modified to improve clarity and applicability - Yes/No (important PDs) was reconsidered for specific PDs

ACTELION	**TPL-100062** **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

8	17June19	- New PD code addition: PD_MM_212

Authoring and Review:

Name	Role /Function	Date	Signature
[redacted]	CLINICAL PROJECT PHYSICIAN **Clinical Science**	17 Jun. 2019	[redacted]
	CLINICAL TRIAL PHYSICIAN **Clinical Science**	17 June 2019	
	CLINICAL PROJECT SCIENTIST **Clinical Science**	17 June 2019	
	CLINICAL TRIAL SCIENTIST **Clinical Science**	17 June 2019	
	TRIAL DATA SCIENTIST **Clinical Development Data Management**	17-jun-2019	
	TRIAL STATISTICIAN **Biostatistics**	17 Jun 2019	

1: Author for 'Condition', 'Additional information expected?', 'Important PD'; Reviewer of the document.
2: Author for ' Type', 'Identifier, 'Source/instructions'; Reviewer of the document.
3. Reviewer of the document.

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

1. Protocol Deviations Before Randomization

1.1. Informed Consent and Patient Rights

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
No informed consent form (ICF) signed	M	*PD_MM.001*	Monitor PD / *To be reported even if subject is not randomized*		Yes
Informed consent form signed after first study procedure	M	*PD_MM.106*	Monitor PD / *Subject number assignation excluded - not a procedure*		Yes
Informed consent procedures conducted by a non-qualified person	M	*PD_MM.107*	Monitor PD		Yes
During pre-randomization period, new ICF version not signed at first applicable protocol scheduled visit	M	*PD_MM.143*	Monitor PD		Yes
DLco addendum to ICF not signed or signed after initiation of DLco sub-study mandated procedure	M	*PD_MM.003*	Monitor PD		Yes
MRI addendum to ICF not signed or signed after initiation of supplementary MRI sub-study mandated procedure	M	*PD_MM.004*	Monitor PD		Yes
PPQ addendum not signed or signed after initiation of PPQ sub-study mandated procedure	M	*PD_MM.108*	Monitor PD		Yes
Other violations of Informed consent procedures for any sub-study	M	*PD_MM.109*	Monitor PD		No

1.2. General Eligibility

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD

ACTELION

TPL-100062

Protocol Deviation Code List

Protocol Number: AC-058B301

Acronym: OPTIMUM

Randomization performed without taking into account centrally provided quality review (i.e. MRI, spirometry)	M	*PD_MM.113*	Monitor PD		No
Randomization performed based on local laboratory results	M	*PD_MM.114*	Monitor PD		No
Any pre-randomization safety assessment required for eligibility not performed prior to randomization	M	*PD_MM.112*	eCRF / Monitor PD / *i.e. Blood pressure, Ophthalmological exam, OCT, Lab samples, Pregnancy tests, ECG, Spirometry*		Yes
Any pre-randomization assessment not required for eligibility not performed prior to randomization	M	*PD_MM.044*	eCRF / Monitor PD / *e.g., Weight, Physical exam, Body temperature, Urinalysis, eC-SSRS, Smoking status, Concomitant medications (ACL)*		No
Pre-randomization assessments not performed using centrally provided devices	M	*PD_MM.115*	Monitor PD / *e.g., ECG, spirometry*		No
Subject re-screened without sponsor approval	M	*PD_MM.186*	Monitor PD / *GCS&E input*		No
Any pre-randomization safety assessment required for eligibility performed, but with at least one repeated assessment or test missing	M	*PD_MM.042*	eCRF / Monitor PD / *e.g., V1 and V3 ECG confirming eligibility but V2 ECG not done; INR V1 available and V2 result unavailable*		No
Any pre-randomization safety assessment required for eligibility performed, but results not available at randomization (subject retrospectively eligible)	M	*PD_MM.188*	eCRF / Monitor PD / *e.g., Lab results available after randomization and confirming eligibility*		No
Any pre-randomization assessment required for eligibility not performed as per study protocol and not covered by another PD code	M	*PD_MM.187*	eCRF / Monitor PD		No
Pre-randomization EDSS not performed according to study protocol or performed after randomization	M	*PD_MM.209*	eCRF / Monitor PD / *Includes assessments not done;*		Yes
Pre-randomization MRI not performed according to study protocol or performed after randomization	M	*PD_MM.210*	eCRF / Monitor PD / *Includes assessments not done;*		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

			MRI not repeated in case of unacceptable quality as per central overread, missing sequences		
Pre-randomization FSIQ-RMS not performed according to study protocol or performed after randomization	M	*PD_MM.211*	eCRF / Monitor PD / *Includes assessments not done; e.g., symptoms domain not completed or completed less than 4 days, impact domain not completed*		No

1.3. Inclusion Criteria

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Age <18 or >55 at Visit 1 (Screening)	P	*PD_PP.005*	eCRF		Yes
Woman of childbearing potential (WOCBP) without negative serum pregnancy test at Visit 1 (Screening) or negative urine pregnancy test at Visit 2 (Baseline)	P	*PD_PP.009*	eCRF		Yes
WOCBP not agreeing to undertake 4-weekly urine pregnancy tests during the study and up to 6 weeks after the first of two tests showing teriflunomide plasma level < 0.02 mg/L	P	*PD_PP.010*	eCRF		Yes
WOCBP not agreeing to use reliable methods of contraception from Visit 1 (Screening) until 6 weeks after the first of two tests showing teriflunomide level < 0.02 mg/L	P	*PD_PP.011*	eCRF		Yes
Fertile male subjects not agreeing to use a condom during the treatment period and for an additional 6 weeks after the first of two tests showing teriflunomide level < 0.02 mg/L	P	*PD_PP.012*	eCRF		Yes
Subject presenting no diagnosis of MS as defined by revised (2010) McDonald Diagnostic Criteria for MS or with no relapsing course from onset (no diagnosis of RRMS or SPMS with superimposed relapses)	P	*PD_PP.013*	eCRF		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

No active disease as required per study protocol inclusion criterion 5	P	*PD_PP.014*	eCRF		Yes
Subject not assessed as ambulatory on EDSS or with EDSS > 5.5 at Visit 1 (Screening) or Visit 2 (Baseline)	P	*PD_PP.016*	eCRF		Yes
Not agreeing to use an accelerated elimination procedure for teriflunomide after last dose of study drug	P	*PD_PP.017*	eCRF		Yes

1.4. Exclusion criteria

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Lactating or pregnant woman	P	*PD_PP.018*	eCRF		Yes
Subject wishing to parent a child during the study	P	*PD_PP.019*	eCRF		Yes
Evidence of a relapse of MS with onset date < 30 days prior to baseline EDSS or between baseline EDSS and randomization	P	*PD_PP.020*	eCRF		Yes
Diagnosis of MS with progressive course from onset (i.e. diagnosis of primary progressive MS or progressive relapsing MS)	P	*PD_PP.047*	eCRF		Yes
Treatment within 7 days prior to randomization with IFN b-1a, IFN b-1b or glatiramer acetate	P	*PD_PP.021*	eCRF		Yes
Treatment within 15 days prior to randomization with b-blockers, diltiazem, verapamil, digoxin or any other anti-arrhythmic or HR lowering systemic therapy, cholestyramine or activated charcoal	P	*PD_PP.022*	eCRF		Yes
Treatment within 30 days prior to randomization with adrenocorticotropic hormone (ACTH), systemic corticosteroids, dimethyl fumarate or vaccination with live vaccines	P	*PD_PP.023*	eCRF		Yes
Treatment within 90 days prior to randomization with medication listed in study protocol exclusion criterion 8	P	*PD_PP.024*	eCRF		Yes
Treatment within 180 days prior to randomization with medication listed in study protocol exclusion criterion 9	P	*PD_PP.025*	eCRF		Yes

ACTELION

TPL-100062

Protocol Deviation Code List

Protocol Number: AC-058B301

Acronym: OPTIMUM

Treatment within 24 months prior to randomization with lymphocyte-depleting biological agents (e.g. rituximab or ocrelizumab) or cladribine	P	*PD_PP.026*	eCRF		Yes
Treatment at any time prior to randomization with alemtuzumab, mitoxantrone, leflunomide, teriflunomide, fingolimod, ponesimod or other investigational S1P modulators, stem-cell transplantation	P	*PD_PP.027*	eCRF		Yes
Any infection or infection risk as per exclusion criteria 12, 13 and 14	P	*PD_PP.028*	eCRF		Yes
Known Progressive Multifocal Leukoencephalopathy (PML) infection / evidence of new neurological symptoms / MRI signs within 6 months prior to randomization compatible with a PML infection diagnosis	P	*PD_PP.110*	eCRF		Yes
Any malignancy as per exclusion criteria 15 and 16	P	*PD_PP.029*	eCRF		Yes
Presence of macular edema	P	*PD_PP.030*	eCRF		Yes
Any cardiovascular condition as per exclusion criterion 18	P	*PD_PP.031*	eCRF		Yes
Type 1 or 2 diabetes poorly controlled according to investigator's judgment, or diabetes complicated with organ involvement such as nephropathy or retinopathy	P	*PD_PP.032*	eCRF		Yes
Any clinically significant pulmonary condition as per exclusion criterion 20	P	*PD_PP.033*	eCRF		Yes
Active or latent TB, except if there is documentation of a previous adequate treatment for latent TB infection or TB disease	P	*PD_PP.034*	eCRF		Yes
Any hematology abnormal laboratory value at Visit 1 (Screening) or Visit 2 (Baseline) as per exclusion criterion 22	P	*PD_PP.035*	eCRF		Yes
Any hepatic condition as per exclusion criteria 23, 24 and 25	P	*PD_PP.036*	eCRF		Yes
Any abnormal liver laboratory value as per exclusion criterion 26	P	*PD_PP.037*	eCRF		Yes
Any renal condition as per exclusion criteria 27 and 28	P	*PD_PP.038*	eCRF		Yes
Any other category as per exclusion criteria 29, 30, 31,32 and 33	P	*PD_PP.039*	eCRF		Yes
Any contraindication for MRI as per exclusion criterion 34	P	*PD_PP.040*	eCRF		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Unlikely to comply with study protocol as per exclusion criterion 35	P	*PD_PP.041*	eCRF		Yes

1.5. Pregnancy and Contraception

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Woman of childbearing potential with a negative serum pregnancy test at Visit 1 (Screening) and a negative urine pregnancy test at Visit 2 (Baseline) performed less than 3 weeks apart	M	*PD_MM.145*	Vendor data / Monitor PD		No
Contraceptive methods requirements for WOCBP not followed as per study protocol by subject	M	*PD_MM.111*	eCRF / Monitor PD / *Refer to Section 4.5.1 of the study protocol*		No
Female partner of male participant not agreeing to use reliable methods of contraception as described in study protocol	M	*PD_MM.144*	Monitor PD / *Refer to section 4.5.2 of the study protocol*		Yes

1.6. Other

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Pre-randomization period > 45 days	P	*PD_PP.116*	eCRF		No
Incorrect stratum factor "EDSS <= or > 3.5" communicated to IVRS at randomization	P	*PD_PP.147*	eCRF		No
Incorrect stratum factor "MS treatment within two years prior to randomization (yes, no)" communicated to IVRS at randomization	M	*PD_MM.148*	eCRF / Monitor PD		No
Additional serum sample for viral serology not taken prior to randomization	M	*PD_MM.189*	eCRF / Monitor PD		No

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Missing information in source documents related to pre-randomization period	M	*PD_MM.118*	Monitor PD / *e.g. assessment time*		No
Any repeated assessment during pre-randomization period visits performed without respecting the minimum interval between assessments as defined per study protocol (except pregnancy tests)	M	*PD_MM.117*	eCRF / Monitor PD / *For CDDM only: to be coded once*		No

2. After Randomization

Note: Applies any time after randomization, unless specified otherwise, (e.g., during PTOP)

2.1. Informed consent and subject rights

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
New ICF version not signed at first applicable protocol scheduled visit or signed after procedure related to new protocol version was performed	M	*PD_MM.150*	Monitor PD		No

2.2. Safety

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
On Day 1 or on day 1 of re-initiation, any post-dose assessment not performed within 24 hours after pre-dose assessments	M	*PD_MM.155*	Monitor PD / *To be reported in FDM database*		Yes
On Day 1 or first day of re-initiation of study drug when post-dose monitoring is required, subject discharged from hospital before 4 hours post-dose or before criteria for discharge were met	M	*PD_MM.062*	eCRF / Monitor PD / *To be reported in FDM database*		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Post-dose monitoring not performed according to study protocol on Day 1 or first day of re-initiation of study drug when post-dose monitoring is required	M	*PD_MM.063*	Monitor PD / *To be reported in FDM database*		No
During maintenance period, any interruption > 3 days where study drug was re-initiated without up-titration and post-dose cardiac monitoring	M	*PD_MM.056*	eCRF / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon subject experiencing an event of suspected opportunistic infection	M	*PD_MM.067*	eCRF / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any cardiovascular discontinuation criterion	M	*PD_MM.064*	Vendor data / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any hematological abnormalities discontinuation criterion	M	*PD_MM.065*	Vendor data / Monitor PD / *GCS&E input*		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any respiratory system (PFT decrease and persistent respiratory AEs) monitoring and discontinuation criterion	M	*PD_MM.068*	Vendor data / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon pregnancy determination	M	*PD_MM.069*	eCRF / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any liver abnormalities monitoring and discontinuation criterion	M	*PD_MM.071*	Vendor data / Monitor PD / *For CDDM only: to be coded once*		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any renal function discontinuation criterion	M	*PD_MM.072*	Vendor data / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any ocular abnormalities discontinuation criterion	M	*PD_MM.073*	eCRF / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any skin reaction discontinuation criterion	M	*PD_MM.076*	eCRF / Monitor PD		Yes
Study drug not interrupted or discontinued (if applicable) upon subject reaching any peripheral neuropathy discontinuation criterion	M	*PD_MM.077*	eCRF / Monitor PD		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

During treatment period, QT-prolonging drug with known risk of Torsades de pointes started or increased in dose without adhering to recommendation defined in study protocol	M	*PD_MM.086*	eCRF / Monitor PD / *Refer to Appendix 3 of the study protocol*		Yes
During treatment period, treatment with systemic steroids or ACTH, except for MS relapses and short-term treatments with low dose/inhaled steroids for pulmonary conditions	M	*PD_MM.087*	eCRF / Monitor PD / *Refer to section 5.2.5 of the study protocol; not valid for low dose for non-pulmonary condition*	Medication start date (CM\CMSPID)	Yes
During treatment period, treatment with any disease modifying drug for MS other than prescribed as per study protocol	M	*PD_MM.088*	eCRF / Monitor PD	Medication start date (CM\CMSPID)	Yes
Active uveitis not monitored and managed as per Guidance described in study protocol	M	*PD_MM.075*	eCRF / Monitor PD / *Refer to Section 5.1.13.7.1 of the study protocol*		Yes
No follow-up monitoring provided until resolution of any event of clinical concern, or until stabilization of the condition, or until change not regarded as any longer clinically relevant	M	*PD_MM.078*	Monitor PD		Yes
No unscheduled OCT examination performed upon experiencing suspected clinically significant findings indicative of macular edema	M	*PD_MM.184*	eCRF / Monitor PD / *e.g., blurred vision*		Yes
During treatment period, treatment with Immunosuppressive	M	*PD_MM.089*	eCRF / Monitor PD / *Refer to section 5.2.5 of the protocol*		Yes
During treatment period, treatment with intravenous immunoglobulins	M	*PD_MM.090*	eCRF / Monitor PD		Yes
During treatment period, treatment with plasma exchange or total lymphoid irradiation	M	*PD_MM.091*	eCRF / Monitor PD		Yes
During treatment period, vaccination with live vaccines	M	*PD_MM.092*	eCRF / Monitor PD		Yes
During treatment period, treatment with investigational drug	M	*PD_MM.093*	eCRF / Monitor PD	Medication start date (CM\CMSPID)	Yes
During treatment period, treatment with cholestyramine or activated charcoal unless needed for accelerated elimination procedure	M	*PD_MM.094*	eCRF / Monitor PD		Yes
During treatment period, treatment with b-blockers, diltiazem, verapamil, digoxin, or any other anti-arrhythmic or HR-lowering systemic therapy as listed in study protocol	M	*PD_MM.095*	eCRF / Monitor PD / *Refer to Appendix 4 of the protocol*		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

During treatment period, any investigational procedure for MS	M	*PD_MM.096*	eCRF / Monitor PD	Date of procedure (CM\CMSPID OR AE\AEREFID)	Yes
Two consecutive safety assessments not performed or performed but results not available and no re-test done	M	*PD_MM.195*	eCRF / Vendor data / Monitor PD / *Valid for following assessments not performed as a whole at 2 or more consecutive visits: Spirometry, ECG, Physical exam, Body weight, Body temperature, Blood pressure, both Ophthalmology and OCT, Chemistry, Urinalysis, Hematology, eC-SSRS*		Yes
Single scheduled non-blood safety assessment missing	M	*PD_MM.125*	eCRF / Monitor PD / *Valid for following assessments not performed as a whole: Spirometry, Physical exam, Urinalysis, Body temperature, Body weight, eC-SSRS, ECG, both Ophthalmological exam and OCT; Not valid if PD_MM.195 is applicable; For CDDM only: to be coded once*		No
Any missing single blood safety assessment not performed or performed but with unavailable results	M	*PD_MM.197*	eCRF / Vendor data / Monitor PD / *Valid for whole Hematology or Chemistry; Not valid if PD_MM.195 is applicable; For CDDM only: to be coded once*		No
Any spirometry assessment with "unacceptable" grading as per central overread and not repeated	M	*PD_MM.172*	Vendor data / Monitor PD / Valid only if repeated test not done; *For CDDM only: to be coded once per occurrence*		No
Lymphocytes or neutrophils repeat testing not performed	M	*PD_MM.066*	Vendor data / Monitor PD / *i.e. in case of abnormal results*		Yes
Spirometry repeat testing not performed	M	*PD_MM.166*	Vendor data / Monitor PD / *i.e. in case of abnormal results in previous scheduled assessment*		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Liver function repeat testing not performed	M	*PD_MM.167*	Vendor data / Monitor PD / *i.e. in case of abnormal results*		Yes
Dalfampridine not taken as instructed in the protocol	M	*PD_MM.122*	eCRF / Monitor PD		Yes
Treatment of relapse with ACTH or other corticosteroids, dose or route of administration than recommended per study protocol	M	*PD_MM.083*	eCRF / Monitor PD / *Includes Tapering, Treatment > 5 days, Oral or IM administration*		No
Any repeated safety assessment not performed following the interval of time defined in the study protocol	M	*PD_MM.135*	eCRF / Vendor data / Monitor PD / *e.g., Lab tests, Spirometry;* *For CDDM only: to be coded once*		No

2.3. Contraception and Pregnancies

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Contraception requirements for WOCBP not followed by subject	M	*PD_MM.081*	eCRF / Monitor PD		Yes
Contraception requirements for fertile male not followed by subject	M	*PD_MM.082*	Monitor PD		Yes
Any urine pregnancy test to be done at home or at site not performed or any test result not shared with study personnel	M	*PD_MM.102*	eCRF / Monitor PD / *Includes serum pregnancy test at FU 2 visit*		No
No reminder or follow-up from study personnel to use the methods of contraception defined in the study protocol	M	*PD_MM.104*	eCRF / Monitor PD		No
Absence of follow-up on pregnancies	M	*PD_MM.070*	Monitor PD		Yes

2.4. Efficacy / Endpoint

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Any scheduled EDSS assessment not done	M	*PD_MM.185*	eCRF / Monitor PD		No
Any EDSS assessment not performed as per study protocol (impact on validity of EDSS score)	M	*PD_MM.190*	eCRF / Monitor PD / *Any deviation not matching to codes: PD_PP.097, PD_MM.134, PD_MM.185 or PD_MM.191;* *To be reviewed by GCS&E for coding to either PD_MM.190 or PD_MM.161*		Yes
Any EDSS assessment not performed as per study protocol (no impact on EDSS score validity)	M	*PD_MM.161*	eCRF / Monitor PD / *Any deviation not matching to codes PD_MM.097, PD_MM.134, PD_MM.185 or PD_MM.191;* *To be reviewed by GCS&E for coding to either PD_MM.190 or PD_MM.161*		No
No EDSS assessment performed to confirm relapse	M	*PD_MM.134*	eCRF / Monitor PD		Yes
EDSS assessment confirming relapse performed > 7 days after onset of symptoms	P	*PD_PP.097*	eCRF		No
EDSS for unconfirmed relapse performed after start of treatment with steroids or > 7 days after onset of symptoms	M	*PD_MM.191*	eCRF / Monitor PD		Yes
FSIQ-RMS symptoms or impact domain not assessable at any given visit or FSIQ-RMS impact domain not assessable at Week 108 (Visit 14 or 14A)	M	*PD_MM.163*	Monitor PD / *e.g., symptoms domain not completed or completed less than 4 days, impact domain not completed;* *For CDDM only: to be coded once*		No
During treatment period or PTOP, any MRI missing or not performed as per study protocol (except Week 108)	M	*PD_MM.162*	eCRF / Monitor PD / *e.g., not repeated in case of MRI unacceptable as per central overread, missing sequences*		No

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

2.5. Blinding

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
At least one Actelion study team member (except site monitor, CCMM, FDM, first dose data scientist) had access to data potentially revealing treatment assignment as defined in study protocol	M	*PD_MM.052*	CTT PD / *Except alerts as defined per study protocol*		Yes
PI/study nurse or any other personnel involved in clinical care and management of subject deliberately tried to find out treatment assignment	M	*PD_MM.053*	Monitor PD / *e.g., local lab results of lymphocytes*		Yes
Any MRI report containing MS-related information without justification documented on the report	M	*PD_MM.154*	Monitor PD / *Except in case of pre-defined threshold documented at site*		No
Treatment code broken before unblinding of the study for a reason not related to management of a clinical event	M	*PD_MM.054*	eCRF / Monitor PD / *Date of unblinding as per IxRS + GDS*	Date of unblinding	Yes
Efficacy assessor involved in clinical care and management or made aware of any data potentially revealing treatment assignment	M	*PD_MM.050*	Monitor PD / *e.g., AEs, ECGs, Lab results;* *Monitor to provide date of potential unblinding*	Date of potential unblinding	Yes
Any site personnel involved in clinical care and management of subject who have had a potential access to day 1 or re-initiation data	M	*PD_MM.051*	Monitor PD / *Monitor to provide date of potential unblinding*	Date of potential unblinding	No
Any site personnel made aware of any data with unblinding potential assessed as high or moderate not related to management of a clinical event (except day 1 or day of re-initiation of study drug data)	M	*PD_MM.152*	Monitor PD / *Data reviewed by at least one site member (except alerts as defined per protocol) or broken capsules in more than one patient at one site. Monitor to provide date. Unblinding type to be provided to CDDM based on "Potential Unblinding Event Documentation Form"); Type of unblinding to be reported for BST only for lymphocytes and/or WBC*	Date of potential unblinding\ Type: Lymphocytes\ WBC	Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Any site personnel made aware of data with unblinding potential assessed as low (except day 1 or day of re-initiation of study drug data)	M	*PD_MM.202*	Monitor PD / *Data received but site claiming not reviewing the results (except alerts as defined per protocol) or single broken capsule; Monitor to provide date. Unblinding type to be provided to CDDM based on "Potential Unblinding Event Documentation Form"); Type of unblinding to be reported for BST only for lymphocytes and/or WBC*	Date of potential unblinding \ Type: Lymphocytes \ WBC	No
Any site personnel involved in clinical care and management of subject who has reviewed Day 1 or re-initiation data	M	*PD_MM.203*	Monitor PD / *Monitor to provide date of unblinding*	Date of potential unblinding	Yes

2.6. Study drug

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Incorrect request of kit type via IVRS (i.e. up-titration kit instead of maintenance kit or vice-versa)	M	*PD_MM.140*	Vendor data / Monitor PD		Yes
Study drug taken from non-allocated kit	M	*PD_MM.058*	Vendor data / Monitor PD / *e.g., incorrect kit dispensed, accidental exchange/ NB: PD to be classified after unblinding into the following two protocol sub-deviations: 'Study drug taken from non-allocated kit: non-assigned treatment received' (PD_MM.058.1) or 'Study drug taken from non-allocated kit: assigned treatment received' at unblinding (PD_MM.058.2)*	Date of intake from incorrect kit (start and end date), kit number: (start date\end date\kit number)	Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

During up-titration period, lack of compliance with study drug	P	*PD_PP.055*	eCRF / *e.g., drug interrupted >1 day without re-up-titration, drug not taken in the correct sequence;* *For CDDM only: to be coded once per up-titration; to be raised manually in case of temporary interruption and no re-initiation*		Yes
During up-titration period, at least one study drug dose not taken in the morning	M	*PD_MM.120*	Monitor PD / *For CDDM only: to be coded once*		No
IMP temperature excursion of any dispensed kit with GQM (GMP-DP) approval for use	M	*PD_MM.156*	Monitor PD / *For CDDM only: to be coded once*		No
IMP temperature excursion of any dispensed IMP kit without GQM (GMP-DP) approval for use	M	*PD_MM.157*	Monitor PD / *To be reported for each affected subject*		No
Misuse and abuse of study treatment	M	*PD_MM.080*	eCRF / Monitor PD / *>3 doses taken on the same day*		Yes
Study drug accountability and compliance check (if e-diary data available) was not performed at any applicable visit	M	*PD_MM.059*	Monitor PD		No
Subject did not return all used, partially used and unused study treatment blister packs at any given visit	M	*PD_MM.060*	Monitor PD		No
Any missing recording of study drug intake in the e-diary	M	*PD_MM.057*	Vendor data / Monitor PD / *e.g., no record of more than 3 days at the time of study drug intake* *For CDDM only: to be coded once*		No
Study drug not available at site in due time	M	*PD_MM.192*	Monitor PD / *Valid for late IMP delivery causing any interruption*		No
Overdose of study treatment	M	*PD_MM.204*	eCRF / Monitor PD / *≤3 additional doses taken*		No

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

2.7. Other

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Any protocol scheduled visit not done (except follow-up and EOT visits)	M	*PD_MM.178*	eCRF / Monitor PD		No
Study drug administered but randomization not done per IRT system	M	*PD_MM.119*	eCRF / Monitor PD		Yes
Study drug administered but date/time of first dose < date/time of randomization	P	*PD_PP.048*	eCRF		Yes
Any pre-dose assessment performed post-dose (except pre-dose PK sampling)	M	*PD_MM.129*	eCRF / Monitor PD/ *For CDDM only: to be coded once per visit if all assessments done post-dose*		No
Any randomized subject who never received study drug	M	*PD_MM.175*	eCRF / Monitor PD		No
Any pre-dose PK sampling performed post-dose	M	*PD_MM.158*	eCRF / Monitor PD		No
Any PK sample not taken	M	*PD_MM.193*	eCRF / Monitor PD		No
Missing information in source documents related to treatment period	M	*PD_MM.137*	Monitor PD / *e.g., assessment time*		No
Vaccine-specific antibody titers assessment before and/or after non-live vaccination not performed	M	*PD_MM.128*	eCRF / Monitor PD		No
Any scheduled visit performed out of time-window (except EOT and follow-up visits)	P	*PD_PP.182*	eCRF/ *For CDDM only: to be added manually in case of Day 15 or first post re-initiation visit out of time-window; to be inactivated manually when present on Visit 14*		No

3. After End of Treatment

ACTELION	**TPL-100062** **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Note: Applies any time after end of treatment unless specified otherwise, (e.g., during PTOP)

3.1. Safety

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Procedure for accelerated elimination of teriflunomide not followed	M	*PD_MM.079*	eCRF / Monitor PD / *e.g.: not done, duration < 7 days, doses < 21, compliance not assessed or assessed as not sufficient for subjects participating to the extension study*		No
Any safety follow-up assessment not done or done after first study drug intake in OPTIMUM-LT	M	*PD_MM.205*	eCRF / Monitor PD / *Valid for Visit 14-16 assessments including: Physical examination, Body weight, both Ophthalmological examination and OCT, Body temperature, SBP/DBP, Spirometry, Hematology, Chemistry, Urinalysis, Pregnancy test, ec-SSRS*		No

3.2. Efficacy / Endpoints

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
FSIQ-RMS symptoms domain not assessable at Week 108 (Visit 14 or 14A)	M	*PD_MM.176*	Monitor PD / *e.g., symptoms domain not completed or completed less than 4 days*		No
Week 108 (Visit 14 or 14A) MRI not performed as per study protocol	M	*PD_MM.177*	eCRF /		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

			e.g., not done or not repeated in case of unacceptable quality as per central overread, missing sequences, done < 14 days after treatment with corticosteroids		
Visit 14 (EOT) was not performed within 7 days after study drug discontinuation for a subject who completed 108 week-treatment as per study protocol.	P	*PD_PP.179*	eCRF		No
Visit 14 (EOT) was not performed within 7 days after study drug discontinuation for a subject who prematurely discontinued study treatment	P	*PD_PP.180*	eCRF		No
Any applicable follow-up visit not performed as per study protocol	M	*PD_MM.181*	eCRF / Monitor PD / *e.g. done out of time-window, full instead of abbreviated Follow-up 2*		No
Visit 14 (EOT) was not performed for a subject who prematurely discontinued study treatment	M	*PD_MM.206*	eCRF		Yes
Visit 14 (EOT) was not performed for a subject who completed 108 week-treatment as per study protocol	M	*PD_MM.207*	eCRF		Yes
Any applicable follow-up visit not performed	M	*PD_MM.208*	eCRF		Yes

4. Anytime during the study

4.1. Safety and Pregnancy

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
SAE not reported as per protocol	M	*PD_MM.105*	Vendor data / Monitor PD / *e.g., delayed*		Yes
Any MRI not reviewed by local neuroradiologist	M	*PD_MM.099*	Monitor PD		Yes

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Pregnancy / pregnancy of partner not reported as per study protocol	M	*PD_MM.194*	Monitor PD		Yes
Any assessment or procedure performed by a non-qualified or non-trained person (except EDSS and MRI)	M	*PD_MM.159*	Monitor PD		No
PFT done without at least 5 minutes of rest prior to testing or done without refraining from taking LABA/SABA for 6/24 hours prior to testing	M	*PD_MM.169*	Monitor PD / *For CDDM only: to be coded once*		No
Blood Pressure or ECG done without at least 5 minutes of rest prior to testing	M	*PD_MM.170*	Monitor PD / *For CDDM only: to be coded once*		No
Any blood pressure or ECG not performed in supine position at any assessment	M	*PD_MM.130*	eCRF / Monitor PD / *Visit 3 assessments: to be reported in FDM database;* *For CDDM only: to be coded once*		No
Any Dermatology or Chest X-ray assessment not performed, or Chest X-ray: performed but results not available	M	*PD_MM.196*	eCRF / Monitor PD		No
Any sub-study assessment not performed, not performed as per study protocol or performed but with at least one test missing	M	*PD_MM.127*	eCRF / Vendor data / Monitor PD / *MRI, DLco, PPQ; DLco: e.g., unacceptable or incomplete QC grade as per central overread;* *For CDDM only: to be coded once per sub-study*		No
No assessment conducted for sub-study with signed ICF	M	*PD_MM.201*	eCRF / Vendor data / Monitor PD		No
Blood pressure not measured on the same arm at every assessment	P	*PD_PP.131*	eCRF / *Visit 3 assessments: to be reported in FDM database;* *For CDDM only: to be coded once*		No
Contraceptive methods not documented in the source notes	M	*PD_MM.198*	Monitor PD / *Valid for female subjects and female partners of male subjects; For CDDM only: to be coded once*		No
Any safety or efficacy assessment or any questionnaire performed out of time-window	M	*PD_MM.136*	eCRF / Vendor data / Monitor PD / *For CDDM only: to be coded once*		No

ACTELION	**TPL-100062** **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

4.2. Efficacy / Endpoints

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Any EDSS assessment performed by personnel not qualified or not trained and certified or re-certified	M	*PD_MM.049*	eCRF / Monitor PD / *"Neurostatus e-Test" web-based interactive test not passed or not re-certified as per protocol*		Yes
Any ophthalmology related EDSS assessment not performed by the efficacy assessor	M	*PD_MM.160*	Monitor PD / *For CDDM only: to be coded once*		No
Any relapse assessment not performed as per study protocol anytime during the study	M	*PD_MM.168*	eCRF / Vendor data / Monitor PD / *Includes:* *- scheduled calls for relapse detection not done* *- any of the following assessments for relapse detection not done: Relapse Assessment Questionnaire, Symptoms form, Body temperature, Physical exam*		No
Any QoL questionnaire not performed or performed but with one or more question missing	M	*PD_MM.126*	Vendor data / Monitor PD / *For CDDM only: to be coded once*		No
Any MSFC/SDMT tests missing or not done as per study protocol	M	*PD_MM.171*	eCRF / Monitor PD / *Includes: two PASAT alternate forms (i.e. A and B) not administered in a counterbalanced way across visits (< 40% of one type), incorrect sequence of tests, tests not timed;* *For CDDM only: to be coded once*		No

ACTELION	TPL-100062 **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

4.3. Other

Condition	Type	Identifier	Source/Instructions	Additional information expected?	Important PD
Any performed assessment not required by study protocol	M	*PD_MM.141*	Monitor PD / *i.e., extra labs, PFTs or MRIs;* *For CDDM only: to be coded once*		No
Other violations of Informed consent procedures	M	*PD_MM.002*	Monitor PD / *e.g., wrong or draft version, wrong language, missing information (signature time...)*		No
E-diary not used at any visit	M	*PD_MM.199*	Monitor PD / *For CDDM only: to be coded once*		No
Data provided by site to external vendors containing potential patient identifiers	M	*PD_MM.153*	Monitor PD / Vendor data / *e.g., real date of birth, subject's initials*		No
Any applicable assessment or report not signed (except central laboratory) and/or not dated	M	*PD_MM.174*	Monitor PD / *e.g., ECG, PFT or MRI reports;* *Includes reports not signed and/or not dated;* *For CDDM only: to be coded once*		No
Any central laboratory report signed and dated more than 5 calendar days after receipt	M	*PD_MM.138*	Monitor PD / *Includes reports not signed and/or not dated;* *For CDDM only: to be coded once*		No
Any performed assessment not following the study protocol requirements	M	*PD_MM.142*	Monitor PD / *Valid only if not covered by another PD code* *For CDDM only: to be coded once per deviation type*		No
Any other study protocol deviation	M	*PD_MM.183*	eCRF / Monitor PD / *For CDDM only: to be coded once per deviation type*		No
Plasma concentration of teriflunomide measured earlier than 20 weeks after last study intake	M	*PD_MM.200*	Monitor PD		Yes

ACTELION	**TPL-100062** **Protocol Deviation Code List**
Protocol Number: AC-058B301 Acronym: OPTIMUM	

Any site personnel made aware of Central laboratory report containing flags (e.g. L or H flags) for blinded lymphocytes or white blood cell counts (except Visit 1 and Visit 2).	M	*PD_MM.212*	Vendor data / Monitor PD / *For CDDM only: to be coded once per occurrence.* *To be reported in FDM database* ONLY based on ACM listings as the information is potentially unblinding.	\Original report date\\\	No

16 APPENDICES

A. Adverse Events of special interest

Adverse events of special interest (AESIs) include the anticipated risks of treatment with ponesimod and events that may be related to MS comorbitities.

The definitions for AESIs are based on the systematic approach using Standardized MedDRA Queries (SMQ). The additional relevant terms can be added to the search or deleted appropriately providing the rationale for the change. The proposal is based on MedDRA version 21.0. The following safety areas are addressed by the pre-defined AESIs:

- **Effect on heart rate and rhythm AESI (including hypotension)**

Effect on heart rate and rhythm AESI are identified by the preferred terms (PT) in the following SMQ: Bradyarrhythmias (including conduction defects and disorders of sinus node function) (SMQ) [20000053]. In addition, the following PT will be added to the search for AEs addressing effects on heart rate and rhythm: 'Bradycardia', 'Electrocardiogram RR interval prolonged', 'Heart rate decrease', 'Presyncope', 'Syncope', 'Loss of consciousness', 'Chronotropic incompetence', and 'Central bradycardia'.

Hypotension will be identified searching the following PT: 'Blood pressure decreased', 'Blood pressure diastolic decreased', 'Blood pressure orthostatic decreased', 'Blood pressure systolic decreased', 'Diastolic hypotension', 'Hypotension', 'Mean arterial pressure decreased', 'Orthostatic hypotension', 'Procedural hypotension', 'Circulatory collapse', 'Blood pressure fluctuation', 'Labile blood pressure' and 'Blood pressure ambulatory decreased'.

- **Hypertension AESI**

Hypertension AESI are identified by the PT in the following SMQ: Hypertension SMQ (narrow scope) [20000147].

- **Hepatobiliary disorders / Liver enzyme abnormality AESI**

Hepatobiliary disorders/ Liver enzyme abnormality AESI are identified by the PT in the following SMQ: Drug related hepatic disorders – comprehensive (SMQ) (broad scope) [20000006]. This SMQ is included in the SMQ Hepatic disorder (SMQ) but only the PT included in Drug related hepatic disorders - comprehensive search (SMQ) are included to identify hepatobiliary disorders / Liver enzyme abnormality AESI.

- **Pulmonary AESI**

These AEs are identified by the PT the in the following SMQs: Asthma/bronchospasm (SMQ) (broad scope) [20000025] or Interstitial lung disease (SMQ) (broad scope)

[20000042]. The PT 'Dyspnoea at rest', 'Dyspnoea', 'Dyspnoea exertional', 'Carbon monoxide diffusing capacity decreased', 'Pulmonary function test abnormal', 'Pulmonary function test decreased', 'Vital capacity abnormal', and 'Vital capacity decreased' are added to the search pre-defined by SMQs Asthma/bronchospasm or Interstitial lung disease.

- **Macular edema AESI**

Macular edema AESI are identified by the following PT: 'Macular oedema', 'Macular hole', 'Macular pseudohole', 'Macular rupture', 'Macular cyst', 'Retinal oedema', 'Diabetic retinal oedema', 'Cystoid macular oedema', 'Papilloedema', and 'Pseudopapilloedema'.

- **Infection AESI**

Infection AESI are identified by the adverse events belonging to the SOC (System Organ Class) Infections and Infestations (SOC), only if reported as serious or severe.

- **Herpetic infection AESI**

Herpetic infection AESI are identified by the PT in the following high level terms: Herpes viral infections (HLT) and the following preferred terms will be added to the search for AEs addressing varicella zoster infection: 'Encephalitis post varicella', 'Herpes gestationis', 'Herpes simplex test positive', 'Human herpes virus 6 serology positive', 'Human herpes virus 8 test positive', and 'Herpes virus test abnormal'.

- **Skin malignancy AESI**

Skin malignancy AESI are identified by the PT in the following SMQs: Skin neoplasms malignant and unspecified (SMQ) (broad scope) [20000173].

- **Non-skin malignancy AESI**

Non-skin malignancy AESI are identified by the PT in the following SMQ: Malignant or unspecified tumours (SMQ) (broad scope) [20000091] excluding the PT which included in the following SMQs: Skin neoplasms, malignant and unspecified (SMQ) (broad scope) [20000173].

- **Seizure AESI**

Seizure AESI are identified by any preferred term in the following SMQ: Convulsions (narrow scope) (SMQ) [20000079].

B. Protocol Deviation code list

Protocol deviations are defined as per the Protocol Deviation code list, version 4, 8 June 2018. Protocol deviations are categorized according to the categories in the Protocol

Deviation code list (level 2 heading, e.g. 'Informed Consent and Patient Rights', 'General Eligibility', …) [Attachment 1].

C. PRO and other composite assessment scoring and handling of missing data

[redacted]

the MSFC Z-score

[redacted]

[redacted]

C.2. FSIQ-RMS Symptoms score

Daily Symptoms Scores:

FSIQ-RMS Symptoms Daily scores (**FSIQ-RMS SDS_{c_Dj}**) will be calculated for each day D_j of assessment as 0 to 100 score using following equation:

$$\textbf{FSIQ-RMS } SDS_{c\_Dj} = \frac{\textbf{(Sum of non-missing items 1, 2, 3, 4, 5, 6, 7) at } D_j}{\textbf{Number of non-missing items} * \textbf{Highest rating (10)}} * 100,$$

Daily scores will be derived at each day (where day refers to study day from randomization) with assessment. If for a given day, more than 3 out of the 7 items are missing in the Diary, then the Symptoms daily score **FSIQ-RMS $SDSc_{\_Dj}$** will be set to missing.

Weekly Symptoms Scores:

FSIQ-RMS Symptoms Weekly Scores (**FSIQ-RMS SWS_{c_Wi}**) within a 7 consecutive day period (Week i) are calculated from daily scores **FSIQ-RMS $SDSc_{\_Dj}$** as follows:

$$\textbf{FSIQ-RMS } SWS_{c\_Wi} = \frac{\textbf{Sum of the non-missing FSIQ-RMS } SDSc_{\_Dj}}{\textbf{Number of non-missing diary days at Week}_i},$$

if the number of missing daily scores is < 4 (i.e. at least 4 daily scores are available).

If more than 3 daily scores are missing (i.e., less than 4 daily scores are available), the weekly score **FSIQ-RMS SWS_{c_Wi}** cannot be calculated and is considered missing.

Weekly scores are derived for the following time points W_i: Baseline (Pre-randomization period), Visits 6 (Week 12), 7 (Week 24), 10 (Week 60), 12 (Week 84) and 14 (Week 108).

Visits will be assigned based on time windowing [see Section 11.3]. The weekly score for a visit is derived based on the 7-day time period with non-missing weekly score. Preference is given to 7-day time periods with (i) minimum number of missing daily scores and (ii) with the seventh day being closest to target. If two periods are equally close to target the worst score is considered.

Baseline Symptoms Weekly Score will be derived during the closest 7-day time period prior to and including randomization date in which a non-missing weekly score can be derived.

Change from baseline **FSIQ-RMS SWS_c** to each assessment visit (**CHN FSIQ-RMS $SWS_{c\ Base\_Wi}$**) will be derived as: Weekly score (**FSIQ-RMS $SWS_{c\ Wi}$**) − Baseline score (**FSIQ-RMS $SWS_{c\ Base}$**).

C.3. FSIQ-RMS impacts domain

FSIQ-RMS impact sub-domain scores are derived as a weekly 0 to 100 score. They are derived based on the second part of the questionnaire based on questions referring to the last 7 days. Weekly scores are therefore derived based on questions answered during a single day.

Each sub-domain score is defined based on 5 items. If at least 3 of these 5 items are available, the scores can be derived. Otherwise it is considered missing. Formulae for derivation are provided below.

FSIQ-RMS Physical Impacts Weekly score:

$$\text{FSIQ-RMS-PIWS}_{c\ Wi} = \frac{(\text{Sum of non missing items } 1, 5, 6, 8, 13 \text{)at Week}_i}{\text{Number non missing items} * \text{Highest rating (4)}} * 100$$

FSIQ-RMS$_c$ Cognitive/Emotional Impacts Weekly score:

$$\text{FSIQ-RMS-CEIWS}_{c\ Wi} = \frac{(\text{Sum of non missing items } 2, 3, 7, 9, 10) \text{ at Week}_i}{\text{Number of non missing items} * \text{Highest rating (4)}} * 100$$

FSIQ-RMS Coping Impacts Weekly score:

$$\text{FSIQ-RMS-CIWS}_{c\ Wi} = \frac{(\text{Sum of non missing items } 4, 8, 11, 12, 13) \text{ at Week}_i}{\text{Number of non missing items} * \text{Highest rating (4)}} * 100$$

Scores are derived for the following time points W_i: Baseline (Pre-randomization period), Visits 6 (Week 12), 7 (Week 24), 10 (Week 60), 12 (Week 84) and 14 (Week 108). Visits will be assigned based on time windowing [see Section 11.3].

Baseline is defined as last available score prior to and including randomization date. Change from baseline is defined as absolute change from baseline.

C.4. SF-36 scoring

Recoding items

Scores for each question response are indicated on the questionnaire. Firstly, seven of the questions are reverse scored, to ensure that a higher item value indicates better health (Items 11b & 11d, 9a & 9e, 6, 9d & 9h). Three of the questions also need to be recalibrated. These are question 1 and questions 7 and 8. Recalibration and reverse score items are as follows:

Precoded and final values for Item 7

Response choices	Precoded Value	Final Value
None	1	6.0
Very mild	2	5.4
Mild	3	4.2
Moderate	4	3.1
Severe	5	2.2
Very Severe	6	1.0

Precoded and final values for item 8 where both items 7 and 8 are answered

Response choices	Item 8 Precoded	Item 7 Precoded Value	Item 8 Final value
Not at all	1	1	6
Not at all	1	2 through 6	5
A little bit	2	1 through 6	4
Moderately	3	1 through 6	3
Quite a bit	4	1 through 6	2
Extremely	5	1 through 6	1

Precoded and final values for item 8 where item 7 is not answered

Response choices	Item 8 Precoded	Item 8 Final Value
Not at all	1	6.0
A little bit	2	4.75
Moderately	3	3.5
Quite a bit	4	2.25
Extremely	5	1.0

Precoded and Final Values for Item 1

Response choices	Precoded Value	Final Value
Excellent	1	5.0
Very Good	2	4.4
Good	3	3.4
Fair	4	2.0
Poor	5	1.0

Precoded and Final Values for Items 11a & 11c, and 11b & 11d

Response choices	Precoded Item Value	Final Value Items 11a & 11c	Final Value Items 11b & 11d
Definitely true	1	1	5
Mostly true	2	2	4
Don't know	3	3	3
Mostly false	4	4	2
Definitely false	5	5	1

Precoded and Final Values for Items 9a - 9i

Response choices	Precoded Item Value	Final Value Items 9a, 9d, 9e, 9h	Final Value Items 9b, 9c, 9f, 9g, 9i
All of the time	1	5	1
Most of the time	2	4	2
Some of the time	3	3	3
A little of the time	4	2	4
None of the time	5	1	5

Precoded and Final Values for Item 6

Response choices	Precoded Value	Final Value
Not at all	1	5
Slightly	2	4
Moderately	3	3
Quite a bit	4	2
Extremely	5	1

Computing domain scores

The raw score for each domain is then calculated as the sum of the scores for the items in that domain, as shown in Table 12. For example the raw score for role physical is the sum of the scores for items 4a, 4b, 4c and 4d. The next step is to transform each raw domain score to a domain scale 0–100 score. This is calculated as follows:

$$\text{(Actual raw score} - \text{lowest possible raw score)} / \text{Possible raw score range} \times 100$$

where the possible raw score range is the highest possible raw score for that domain minus the lowest possible raw score for that domain. This transformation converts the lowest and highest possible scores to 0 and 100, respectively. Scores between these values represent the percentage of the total possible score achieved. All except one of the 36 items (Item 2 self-reported health transition) are used to score the eight SF-36 domains. Each item is used in scoring only one domain. The table below provides the information necessary to apply the above formula to each domain.

Table 12 Scoring and transforming scales

Domain	- Raw Domain scores (individual final item value range)	Lowest possible raw score	Highest possible raw score	Possible raw score range
Physical Functioning	3.a to j (1-3)	10	30	20
Role-Physical	4.a to d (1-5)	4	20	16
Bodily Pain	7# +8# (1-6)	2	12	10
General Health	1# + 11.a + b* + c + d* (1-5)	5	25	20
Vitality	*9.a** + *e** + *g* + *i* (1-5)	4	20	16
Social Functioning	6* + 10 (1-5)	2	10	8
Role-Emotional	5.a + b + c (1-5)	3	15	12
Mental Health	9.b + c + d* + f + h*(1-5)	5	25	20

[!] Sum of the final item values after recalibrating and recoding,
* indicates items with reversed scores
# indicates items with recalibrated scores

Computing normalized domain scores

For each of the individual domain the scale 0-100 score is standardized using a z-score transformation. The z score is computed by subtracting the mean 0-100 general US population score (1998) from the computed scale 0-100 score and dividing the difference by the corresponding domain standard deviation from the 1998 general U.S. population as detailed in SF-36 Physical and Mental Health Summary Scales: How to score Version 2 of the SF-36 Health Survey [Ware 2000].

Calculations of the Z scores using 1998 General US population Means and SD are as follows:

$PF_Z = (PF - 83.29094) \div 23.75883$

$RP\ Z = (RP - 82.50964) \div 25.52028$

$BP\ Z = (BP - 71.32527) \div 23.66224$

$GH\ Z = (GH - 70.84570) \div 20.97821$

$VT\ Z = (VT - 58.31411) \div 20.01923$

$SF\ Z = (SF - 84.30250) \div 22.91921$

$RE\ Z = (RE - 87.39733) \div 21.43778$

$MH\ Z = (MH - 74.98685) \div 17.75604$

Each of the individual domain z-score is then transformed to norm based (mean 50, standard deviation 10) scoring (also known as t scores) by multiplying each domain z-score by 10 and adding the resulting product to 50.

Normative domain scores:

PFt = (PF Z × 10) + 50

RPt = (RP Z × 10) + 50

BPt = (BP Z × 10) + 50

GHt = (GH Z × 10) + 50

VTt = (VT Z × 10) + 50

SFt = (SF Z × 10) + 50

REt = (RE Z × 10) + 50

MHt = (MH Z × 10) + 50

Computing PCS and MCS measures

Physical and Mental Component Summary measures (PCS and MCS respectively) are also calculated. These are aggregate scores for the physical and mental components which are computed by multiplying each of the individual domain z-scores (calculated as described above) by their respective physical/mental factor score coefficients and summing the eight products.

Aggregate component summaries are derived as follows:

$PCS_Z = (PF_Z \times 0.42402) + (RP_Z \times 0.35119) + (BP_Z \times 0.31754) + (GH_Z \times 0.24954) + (VT_Z \times 0.02877) + (SF_Z \times -0.00753) + (RE_Z \times -0.19206) + (MH_Z \times -0.22069)$

$MCS_Z = (PF_Z \times -0.22999) + (RP_Z \times -0.12329) + (BP_Z \times -0.09731) + (GH_Z \times -0.01571) + (VT_Z \times 0.23534) + (SF_Z \times 0.26876) + (RE_Z \times 0.43407) + (MH_Z \times 0.48581)$

Each summary measure is then transformed to the norm-based (mean 50, standard deviation 10) scoring (t scores). This is accomplished by multiplying each aggregated score by 10 and adding the resulting product to 50.

Normative Physical Component Summary score: $PCS_t = (PCS_Z \times 10) + 50$

Normative Mental Component Summary score: $MCS_t = (MCS_Z \times 10) + 50$

Handling of Missing data

Imputation for missing items in a domain

For an individual domain where 50% or more of the item questions are non-missing, the average of the non-missing items for that patient in that domain is calculated. This

average is then be used to impute the remaining missing item questions of that domain. Note: imputation should be based on the final domain score after recalibration / reverse scores of pre coded scores.

However, if more than 50% of item questions are missing the domain is recorded as missing.

Aggregated scores are computed after the imputation rule has been applied to the component domains, but are set to missing if any domain is missing.

Handling of duplicate entries on Questionnaire form

Duplicate entries within the same day for a question item are handled as follows:

- If a respondent marks two responses which are adjacent to each other, then the worst case (lowest value) is selected. Note this is selected based on the item score after recalibration / reverse scores of pre coded scores.
- If a respondent marks two responses which are not adjacent to each other, then the score is missing for that item
- If a respondent marks three or more responses then the score is missing for that item

C.5. WPAI:MS scoring

The four WPAI:MS outcome scores are defined based on the outcome Q1–Q6 of Questions 1–6 (ref: http://www.reillyassociates.net/WPAI_Scoring.html):

Q1 = currently employed (Y/N)

Q2 = hours missed due to MS

Q3 = hours missed other reasons

Q4 = hours actually worked

Q5 = degree MS affected productivity while working

Q6 = degree MS affected regular activities

WPAI:MS outcome scores are defined as detailed below. At each assessment visit for those who are currently employed the following 3 scores are defined:

- **Absenteeism**: Percent work time missed due to MS = Q2 / (Q2 + Q4) × 100
- **Presenteeism**: Percent impairment while working due to MS = Q5 / 10 ×100
- **Work productivity loss**: Percent overall work impairment due to MMS = {Q2 / (Q2 + Q4) + [(1–Q2 / (Q2 + Q4)) × (Q5 / 10)]} × 100

For all subjects irrespective of their current employment status

- **Activity impairment**: Percent activity impairment due to MS = Q6 / 10 × 100

Scoring: Handling of missing or inconsistent data

The following rules for handling of missing or inconsistent data within the same day are implemented as described in the section "Coding Rules for Self-Administration" by Reilly Associates [Reilly Associates 2002].

- *Employment status (Q1):*

  If Q1 = 'YES' or Q1 = 'NO' or 'missing' and hours missed (Q2) or hours worked (Q4) > 0, then subjects is considered employed.

  If Q1 = 'missing' and hours missed (Q2) and hours worked (Q4) = 0, then considered not employed (i.e., Q1 considered No).

- *Hours missed (Q2) and hours worked (Q4)*

  If hours worked (Q4) = 0, then productivity while at work (Q5) is not applicable and considered as missing.

  If subject enters a range of hours for (Q2 or Q4), use the midpoint.

  If a subject is considered employed but hours missed (Q2) is missing, then impute as zero.

- *Productivity Questions (Q5 and Q6)*

  If for one question, two responses are provided, enter the midpoint and round to nearest integer.

Missing response after applying the above scoring

WPAI scores cannot be calculated if there is a missing response to one or more corresponding individual questions.

D. Laboratory marked abnormalities

Table 13 Thresholds for marked laboratory abnormalities

Parameter (SI unit)	LL	LLL	HH	HHH
Hemoglobin (g/L)	< 100	< 80	> 20 g/L above ULN (for pre-treatment assessments, and post-treatment assessments when baseline ≤ ULN) or increase from baseline > 20 g/L (for post-treatment assessments when baseline is > ULN)	> 40 g/L above ULN (for pre-treatment assessments, and post-treatment assessments when baseline ≤ ULN) or increase from baseline > 40 g/L (for post-treatment assessments when baseline is > ULN)
Hematocrit (L/L)	< 0.28 (female) < 0.32 (male)	< 0.20	> 0.55 (female) > 0.60 (male)	> 0.65
Platelet count (10^9 /L)	< 75	< 50	> 600	> 999
WBC count (10^9 /L)	NA	< 1.9	> 20.0	> 100.0
Lymphocyte (10^9 /L)	ND	< 0.2	> 4.0	> 8.0
Neutrophils (10^9 /L)	< 1.5	< 1.0	ND	ND
Eosinophils (10^9 /L)	ND	ND	> 5.0	ND
Polymorphonuclear leucocyte/Band cells (%)	ND	ND	> 90%	> 95%
AST (U/L)*	ND	ND	≥ 3 ULN	≥ 5 ULN
ALT (U/L)*	ND	ND	≥ 3 ULN	≥ 5 ULN
Total bilirubin (umol/L)	ND	ND	≥ 2 ULN	≥ 5 ULN
Alkaline Phosphatase (U/L)	ND	ND	> 2.5 ULN	> 5 ULN
INR*	ND	ND	> 1.5 ULN	> 2.5 ULN
Creatinine (umol/L)*	ND	ND	>1.5 ULN or >1.5 × baseline	> 3 ULN or >3 × baseline
Creatinine clearance (mL/min)	< 60	< 30	ND	ND
Urea (mmol/L)	ND	ND	> 2.5 ULN	> 5 ULN
Albumin (g/L)	< 30	< 20	ND	ND
Glucose (mmol/L)	< 3.0	< 2.2	> 8.9	>13.9
Potassium (mmol/L)	< 3.2	< 3.0	> 5.5	> 6.0
Sodium (mmol/L)	ND	< 130	> 150	> 155
Calcium (mmol/L)	< 2.0	< 1.75	> 2.9	> 3.1
Triglyceride (mmol/L)	ND	ND	> 3.42	> 11.4
Cholesterol (mmol/L)	ND	ND	> 7.75	> 12.92

* HH and HHH based on CTCAE 2010 v4.03. ALT = alanine aminotransferase; AST = aspartate aminotransferase; ND = not defined; may be complemented by definitions provided by the central laboratory (see central laboratory manual); ULN = upper limit of normal.
Source: Protocol AC-058B301, appendix 6.

E. Document history

Version	Effective Date	Reason (main)
1.0	11 Dec 2018	New
2.0	20 June 2019	FSIQ-RMS responder threshold added based on results from anchor analysis Minor update for efficacy analyses: estimates per treatment arm (least square means) derived based on observed margins 12-week and 24-week CDA: handling of relapse start date and end date clarified; alternative censoring approach added based on health authority feedback. Safety analyses: Additional analyses added, small updates based on health authority feedback including presentation of AEs from Day 1 to Day 16 and updates to blood pressure and heart rate abnormality thresholds. Editorial updates have been applied.