Official Title: **Phase III: UbLiTuximab In Multiple Sclerosis Treatment Effects (ULTIMATE II STUDY).**

NCT Number: NCT03277248

Document Date: SAP Version 3.0: 04 September 2020

SAP Version/Date: Version 3.0 / 04 September 2020

Related to Local Protocol #: Protocol TG1101-RMS302

Protocol Title:

Phase III: UbLiTuximab In Multiple Sclerosis Treatment Effects **(ULTIMATE II STUDY).**

Sponsor: TG Therapeutics, Inc.
2 Gansevoort St; 9th Floor
New York, NY 10014
Tel: (212) 554-4484

IND Number: 127265
EudraCT Number: 2017-000639-15

Biostatistician: [redacted], PhD
[redacted]
TG Therapeutics, Inc.
2 Gansevoort St; 9th Floor
New York, NY 10014

Study Coordination: TG Therapeutics, Inc.
2 Gansevoort St; 9th Floor
New York, NY 10014
Tel: (212) 554-4484

Version: 1.0 **Date:** 14 April 2017
Version: 2.0 **Date:** 7 June 2017
Version: 2.1 **Date:** 26 July 2017
Version: 3.0 **Date:** 04 September 2020

List of Abbreviations and Definitions of Terms

9-HPT	9 Hole Peg Test
ACTH	Adrenocorticotrophic hormone
ADA	Anti-drug antibody
ADME	Absorption, Distribution, Metabolism, Excretion
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
aPTT	Activated partial thromboplastin time
ARR	Annualized relapse rate
AST	Aspartate aminotransferase
ATC	Anatomic Therapeutic Chemical Classification
BART	Blinded Assessment Relapse Team
BMI	Body mass index
BOD	Burden of disease
BP	Blood pressure
BPF	Brain parenchymal fraction
BUN	Blood urea nitrogen
CNS	Central nervous system
CrCl	Creatinine clearance
CRF	Case report form
CSF	Cerebrospinal fluid
CSR	Clinical study report
DILI	Drug-induced liver injury
DMC	Data Monitoring Committee
DSMB	Data Safety Monitoring Board
EAE	Experimental allergic encephalomyelitis
ECG	electrocardiogram
EDSS	Expanded Disability Status Scale
EOPII	End of Phase II
EPTD	Early permanent treatment discontinuation
EQ-5D	EuroQoL (quality of life scale)

FIS	Fatigue Impact Scale
FS	Functional System
FSH	Follicle stimulating hormone
GCP	Good clinical practice
Gd	gadolinium
GEE	Generalized estimating equation
GGT	Gamma glutamyl transferase
Hb	hemoglobin
HCG	Human chorionic gonadotrophin
HIV	Human immunodeficiency virus
HLGT	High Level Group Term
HLT	High Level Term
HMR1726	Teriflunomide
ICH	International Council for Harmonisation
ICU	Intensive care unit
INR	International normalization ratio
IP	Investigational product
IRAP	Independent Relapse Adjudication Committee
ITT	Intent-to-treat
IVRS	Interactive voice response system
LOCF	Last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified Intent-to-Treat
MMRM	Mixed-effect model with repeated measures
MRI	Magnetic resonance imaging or magnetic resonance image
MRI-AC	MRI analysis center
MS	Multiple sclerosis
MSFC	Multiple Sclerosis Functional Composite
MSQoL-54	Multiple Sclerosis Quality of Life-54
NEDA	No evidence of disease activity
PASAT	Paced Auditory Serial Addition Test
PCA	Predefined change abnormal
PK	Pharmacokinetic

PK/PD	Population pharmacokinetic/pharmacodynamic
PML	Progressive Multifocal Leukoencephalopathy
PP	Per Protocol
PR	Pulse rate
PS-MS	Performance Scales–Multiple Sclerosis
PT	Preferred term
RMS	Relapsing forms of Multiple Sclerosis
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SDMT	Symbol Digit Modalities Test
SF-36	Short Form generic health survey (36 items)
SGGT	Serum gamma-glutamyl transferase
SGOT/AST	Serum glutamic oxaloacetic transaminase/aspartate aminotransferase
SGPT/ALT	Serum glutamic pyruvic transaminase/alanine aminotransferase
SMQ	Standardized MedDRA Queries
SOC	System Organ Class
T25-FW	Timed 25-Foot Walk
TEAE	Treatment emergent adverse event
VAS	Visual Analogue Scale
WBC	White blood cells
WHO-DD	World health organization drug dictionary
WPAI	Work productivity and activity impairment

Contents

1 Overview and Investigational Plan

This statistical analysis plan (SAP) provides a detailed description of the statistical strategy for the "TG1101-RMS302" protocol Version 5.0 dated 04 September 2020. The goal of the SAP is to pre-specify the statistical approaches to the analysis of study data prior to study initiation. It has been revised before database lock to cover protocol amendments and all technical aspects of the planned statistical analysis. Significant revisions and the corresponding rationale are summarized in Section 5 of this document.

1.1 Study Design and Randomization

This is a 120 week, Phase 3, randomized, multi-center, double-blinded, double-dummy, active-controlled study that is primarily designed to assess the annualized relapse rate (ARR) and safety/tolerability of ublituximab (TG-1101; UTX)/oral placebo as compared to teriflunomide/IV placebo in subjects with relapsing forms of Multiple Sclerosis (RMS).

After early termination of trial or at the end of study visit at Week 96, subjects will enter a 20 week follow-up program inclusive of the accelerated elimination of teriflunomide.

1.2 Objectives

1.2.1 Primary Objective

To determine the annualized relapse rate (ARR) in subjects with RMS after 96 weeks (approximately 2 years with a year is equal to 365.25 days) treatment with IV infusion of ublituximab/oral placebo compared to 14 mg oral teriflunomide/IV placebo.

1.2.2 Secondary Objectives

The secondary objectives are to examine the effects of ublituximab/oral placebo as compared to teriflunomide/IV placebo as follows:

1. On MRI parameters,
2. Confirmed Disability Progression (CDP),
3. No Evidence of Disease Activity (NEDA),
4. Symbol Digit Modality Test (SDMT)
5. To evaluate the safety of ublituximab/oral placebo, as determined by adverse events (AEs) and serious adverse event (SAEs), including MS worsening.

1.2.3 Tertiary Objectives

1. Change in MSFC score from baseline to Week 96
2. Time to Confirmed Disability Progression for at least 24 weeks
3. Time to Confirmed Disability Improvement for at least 12 weeks
4. Time to Confirmed Disability Improvement for at least 24 weeks
5. Health outcomes (MSQoL-54 (inclusive of SF36); FIS, hospitalization, steroid use, time out of work)
6. Total volume of Gd enhancing T1 lesions per MRI scan over the treatment period
7. Volume of T2 lesions
8. Volume of hypointense T1 lesion component (black holes)
9. Proportion of subjects free of disability progression at 24 weeks, 48 weeks, and 96 weeks
10. Proportion of subjects with a relapse
11. Time to first confirmed relapse

1.3 Determination of Sample Size

[redacted] this required sample size is increased to 220 per group or a total randomized of 440.

1.3.1 Interim Sample Size Reassessment

An independent committee, Blinded Assessment Relapse Team (BART), reporting to TG Therapeutics, and the DSMB will reassess the sample size for the study when 210 of the 220 participants have been randomized. Assuming a uniform rate of recruitment, we need an estimate as to how many person years of exposure will occur before an interim adjustment occurs. If recruitment is going to be accomplished in 8 to 10 months, the average exposure time when 210 subjects per group are recruited would experience half of the recruitment period or 9/2=4.5 months. We can use all of the data available at that time and thus will have approximately 4.5 x 210 = 945 person months or approximately 43 person years of observation (the denominator to compute the ARR).

The Table below provides the sample sizes (total adjusted sample sizes) needed per group per study based on the findings at the interim (Columns of the Table) for reductions in the ARR of teriflunomide by 40%. As an example, if the estimated ARR in the pooled population is 0.20 and we are assuming an hypothesized 40% reduction in the ARR rate by ublituximab, then the ARR in the individual groups (red numbers in the table below) can be derived by solving a simple equation; $(N_U*0.6\ P_T + N_T*P_T)/(N_U + N_T)=0.20$ where P_T is the ARR in the teriflunomide group and N_T and N_U are the sample sizes, assumed to be equal. Thus, if the combined ARR=0.20, then this equation under the hypothesized reduction of 40% yields an estimated ARR of 0.25 in teriflunomide and 0.15 in ublituximab. Calculating the required sample size using a negative binomial comparison with 1.75 years of observation overall, results in an updated sample size of 219 per group. Thus, the original sample size of N=200 would be adjusted upward by adding 19 participants per treatment group, plus the 10% for dropouts yielding a total of 241 subjects per group or 21 additional participants.

In the Table below, which can guide the interim sample size reassessment, we see two color blocks in the table. The green area shows results that have two (2) characteristics - one, they require smaller sample sizes and two, they are unlikely to occur because the ARR rate in the teriflunomide group would have to exceed 0.31 (arbitrarily chosen, but some level is reasonable as an upper limit on what the rate is likely to be). Either way, if the blinded ARR falls in the green area or above, the sample size will not be reduced since we do not plan to reduce the sample size and 210 participants per group will have already been enrolled. The yellow area is the range for the combined total rates. If the overall ARR =0.19 then the estimated teriflunomide rate would be estimated at 0.2375 and ublituximab would be estimated at 0.1425 and the sample size required would be 233 per group. Thus, a rate <= 0.19, the sample size will be increased by 30 per treatment group in each study, but no more. Such observed relapse rates are quite low and unexpected, so that further adjustments are probably not warranted.

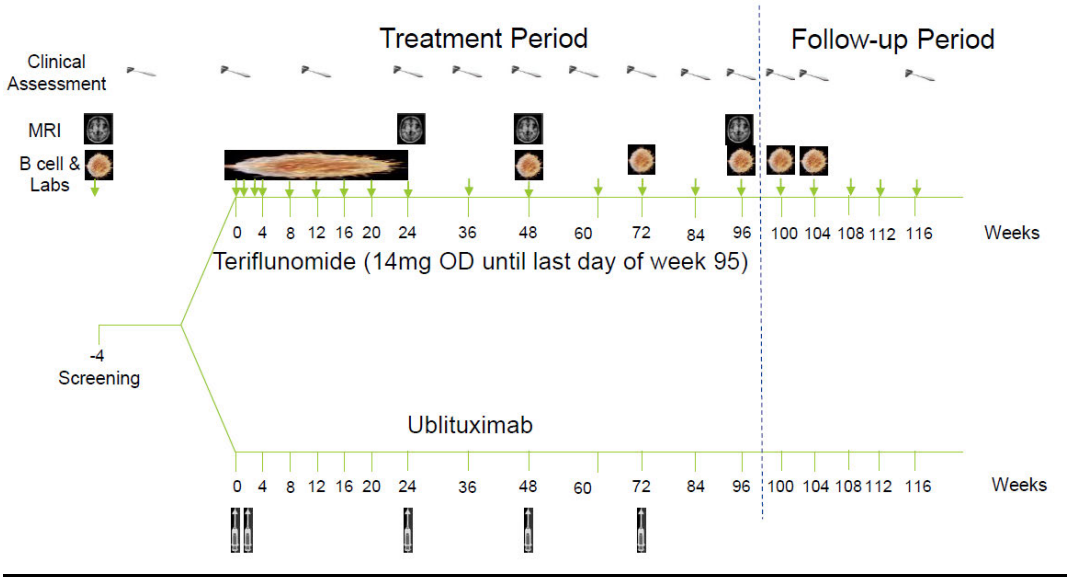
Recruitment will continue while the interim analysis is ongoing and if the participants recruited beyond the 220 per group will be used in the final analyses if the sample size is not increased and will count toward the additional sample size if warranted.

Table 1 – Findings at Interim

% Reduction in ARR		Findings at Interim								
	Treatment	0.160	0.180	0.200	0.220	0.240	0.260	0.280	0.290	0.300
40%	Ublituximab/ Oral Placebo	0.120	0.135	0.150	0.165	0.180	0.195	0.210	0.218	0.225
	Teriflunomide/ Placebo	0.200	0.225	0.250	0.275	0.300	0.325	0.350	0.363	0.375
Avg Follow-Up = 1.75 years	N	268	243	224	208	195	183	174	169	165

1.4 Study Plan

Refer to the study protocol for TG1101-RMS302 for study schedule for analysis variables.



2 Statistical and Analytical Procedures

2.1 Analysis Variables

2.1.1 Demographics and Baseline Characteristics

All baseline safety and efficacy parameters (apart from the one listed below) are presented along with the summary statistics in the safety and efficacy sections (Section 2.4).

Demographic and other baseline characteristics

Demographic variables are gender (Male, Female), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other, or Not Applicable due to country's legislation), age in years (quantitative and categorical variable: <38, and ≥38 years), Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown), country and region.
Regions are defined as follows:

- USA and Western Europe: USA, Spain, United Kingdom
- Eastern Europe: all other countries

Other baseline characteristics:

- Baseline hepatic function (normal, mild, moderate, or severe impairment)
- Baseline renal function (normal, mild, moderate, or severe impairment)

Classification of hepatic impairment per NCI-ODWG Criteria

Category	NCI-ODWG	
	Total Bilirubin	ALT or AST
Mild	B1: ≤ULN	B1: >ULN

	B2: >1-1.5x ULN	B2: Any
Moderate	>1.5-3x ULN	Any
Severe	>3x ULN	Any

Classification of renal impairment per EMA Guidance

Group	Description	GFR (ml/min)
1	Normal renal function	≥ 90
2	Mildly decreased renal function	60- <90
3	Moderately decreased renal function	30- <60
4	Severely decreased renal function	<30 not requiring dialysis
5	End stage renal disease (ESRD)	<15 requiring dialysis treatment

Medical and surgical history

Medical history data will be coded to "Preferred Term (PT)" and associated primary "System Organ Class (SOC)" using the Medical Dictionary for Regulatory Activities (MedDRA) version 20.0. Dictionary upgrades are not planned during the course of the study.

Disease characteristics at baseline

- Number of relapses experienced within past 1 year and 2 years (0, 1, 2, 3, and ≥4), respectively
- Time since first diagnosis of MS (in years) to be derived as
  (Day of randomization – Day of first MS diagnosis) / 365.25
- Time since first symptoms of MS (in years) to be derived as
  (Day of randomization – day of first symptoms of MS) / 365.25
- Time since most recent relapse onset (in months) to be derived as
  (Day of randomization – day of most recent relapse onset) / (365.25 / 12)
- Baseline EDSS score
- Baseline EDSS (≤3.5, >3.5)
- Number of baseline Gadolinium (Gd)-enhancing lesions (0, ≥1)
- Baseline T2 lesion count
- Baseline T2 lesion Volume
- Baseline brain volume

Multiple Sclerosis treatment history

- Received any MS therapy for at least 1 month (30 days) prior to enrollment (Yes, No)
- MS therapy in the previous 5 years by Preferred Name
- Previous corticosteroid treatment for relapse
- Time since last corticosteroid treatment for relapse (months)

Vital signs and height

Vital signs are variables: weight in kg (quantitative variable and categorical variable: <50, ≥50 - <100, ≥100), height (m) and Body Mass Index (BMI) in weight[kg]/height[m²] (<30, ≥30) calculated from weight and height.

Prior or concomitant medications or Follow-Up Medications

All medications taken within 3 months (5 years for previous MS treatments) before randomization and until the end of the study are reported in the case report forms (CRF).

All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD).

- Prior medications are those taken prior to first dose of treatment, either stopped before study treatment start or continuing.
- Concomitant medications are those taken at any time during the study treatment period, from the first study treatment use to the day of last study drug treatment, started before or during study treatment.
- Follow-up medications are those taken at any time between last study treatment and end of follow-up,

Technical details related to calculating dates, imputation for missing dates, etc. are described in Section 2.5.

2.1.2 Duration of Study Treatment Exposure and Compliance

Duration of oral study treatment exposure

Duration of treatment exposure (to oral study treatment teriflunomide/placebo) will be derived as:

(Date of Last oral Study Treatment - Date of First oral Study Treatment + 1).

In addition, duration of treatment exposure will also be categorized using the following intervals:

- 0 < Duration ≤4 weeks
- 4 < Duration ≤12 weeks
- 12 < Duration ≤24 weeks
- 24 < Duration ≤48 weeks
- 48 < Duration ≤72 weeks
- 72 < Duration <96 weeks
- Duration ≥96 weeks.

Treatment compliance

Treatment compliance for ublituximab/ IV placebo is summarized as the total number of infusions received over the 5 expected infusions. Each infusion must be greater than 50% complete.
Furthermore, the overall number of infusions, number and percentage (based on all infusions) of infusions completed with/without interruption and infusions not completed will be calculated per arm.
The number and percentage of patients (based on all patients with at least one infusion) with and without any infusion complication (interruption or stop) per arm will be presented as well.

The percentage of planned total infusion received will also be calculated.

Treatment compliance for teriflunomide/oral placebo is summarized as the percentage of actual drug taken over the exposure duration. It is calculated according to the following formula:

Treatment compliance =100% x (total number tablets taken) / oral treatment exposure duration (days).

2.1.3 Efficacy Variables

Evaluation schedule for efficacy variables are presented in TG1101-RMS302 protocol.

Baseline is defined as the last non-missing value on or before the date of randomization. For subjects who have no value on or before randomization date, the last non-missing value on or before the date of first dose will be used as baseline. Baseline MRIs delayed due to scheduling problems or re-testing, a window of up to 30 days will be allowed for use of the baseline MRI value.

2.1.3.1 Primary Efficacy Variable

The primary efficacy variable (per patient) is the number of IRAP-confirmed relapses (recorded in the eCRF as "Was the relapse adjudicated as protocol defined relapse by IRAP?"=Yes) which started (Start Date of Suspected Relapse Symptoms) on or after the day of randomization and up to the day of last study treatment. If the exact start date of relapse symptoms is unknown, the relapse will be assumed to have started during this period unless sufficient information is available to conclude that the symptoms started before or after (e.g. in case of missing start day, the symptoms start month is after the month of treatment end).

Annualized Relapse Rate will be summarized within treatment group as the ratio of the sum of all patients' relapse counts divided by the sum of all patients' treatment duration (in years). Years under treatment will be derived as:

Treatment years = (date of last treatment intake (oral or IV) – date of randomization + 1) / 365.25.

The statistical testing approach is described in section 2.4.3.1.

Please see below for the assessment process of suspected relapses.

MS Relapses During the Study
Subjects enrolled into the study will be closely monitored through the study course by the Site and Sponsor's personnel as well as by an external Independent Data Safety Monitoring Board (DSMB) to ensure subjects' welfare.

Relapses that occur after study drugs are withdrawn will be assessed over the remainder of the study period and this data will be utilized as part of additional sensitivity analysis (described in the statistical plan) as long as the subject has not withdrawn their consent to be in the trial.

The confirmation of a relapse will be determined by the Independent Relapse Adjudication Committee (IRAP). The steps for confirming a relapse are as follows:

1. **Subject Reported Relapse**

Subjects who experience new or worsening neurologic symptoms are instructed to contact the Treating Neurologist within 48 hours of symptom onset. All new or worsening neurological events, reported at a visit or over the phone, consistent with MS representing a clinical relapse as assessed by the subject will be initially reported to the Treating Neurologist. In parallel, the Examining Neurologist will also be notified to perform additional assessments. Upon the notification to the Treating Neurologist; the time and date of suspected subject reported relapse will be documented in the dedicated page eCRF (subject reported) by the Treating Neurologist or designee.

The following should be recorded in the eCRF (subject reported) by the Treating Neurologist or designee:

- Date and time of symptom onset as reported by the subject
- Symptoms and events reported by the subject (both neurological and non-neurological)
- Date of initial contact by subject
- Time of initial contact by subject

Within 7 days of the subject reporting the suspected relapse to the Treating Neurologist, the subject will be scheduled and assessed by both the Treating and Examining Neurologist independently.

2. **Initial Assessment of Relapse by Treating Neurologist and Examining Neurologist**

The Treating Neurologist will perform a neurological and physical examination and safety assessment. The Treating Neurologist or designee will document in the eCRF the following:

- Time and date of neurological and physical examination
- Clinical findings of neurological and physical examination
- Review of previously documented concomitant medications and medical history
- Safety assessment, including vital signs, CBC, B-cell count, lymphocyte counts, fibrinogen, quantitative immunoglobulin

Subjects with suspected clinical relapses reported to the Treating Neurologist or designee will be referred to the blinded Examining Neurologist who will use Neurostatus to assess the EDSS independently of the Treating Neurologist.

The Examining Neurologist or designee will document in the eCRF the following:

- Date and time of assessment
- EDSS score
- FS Scores

In addition, subjects may not begin IV methylprednisolone (IVMP) treatment of a relapse until the Examining Neurologist has completed his/her examination.

Treatment of relapses may proceed at the discretion of the Treating Neurologist only after the Examining Neurologist has completed his/her exam. The Treating Neurologist will not know the results of the EDSS assessment. When making determination to treat a potential acute relapse

the Treating Neurologist can use IVMP (1.0 g/day for at least 3 days and can be extended to 5 days) and will not affect the subject eligibility to continue in the study.

The findings from either the Treating Neurologist and Examining Neurologist will not be shared. The Examining Neurologist must be blinded to the Treating Neurologist's assessment and vice versa.

Following the completion of assessments by the Treating and Examining Neurologists these results will be sent independently to the IRAP.

3. Confirmation of Relapse by IRAP

All assessments entered into the eCRF from both the Treating Neurologist and the blinded Examining Neurologist, will be sent to an IRAP. The IRAP adjudicates each case based on all available data provided for that case and members are not permitted to contact the site or the sponsor for additional information. The IRAP will, in turn, make the final determination of whether the neurological events meet the criteria for a protocol-defined relapse.

Each episode of relapse must be confirmed by the IRAP, based on the neurological assessments performed by the Treating and/or Examining Neurologist independently by documenting either of the following:

- ≥ 2 points increase on one of the appropriate FS or 1 point on two or more of the appropriate FS. The change must affect the selected FS (i.e., pyramidal, ambulation, cerebellar, brainstem, sensory or visual). Note, the change in FS scores should correspond to the patient's symptoms (e.g., patient reported change in visual acuity should correspond to a change in the vision FS score). Episodic spasm, sexual dysfunction, fatigue, mood change or ladder or bowel urgency or incontinence will not suffice to establish a relapse;
- An increase of ≥0.5 points in the EDSS score (unless EDSS score = 0, then an increase of at least 1.0 points is required) from the previous clinically stable assessment.

MS relapses are defined as a new or worsening neurological symptoms lasting ≥24 hours with the absence of fever, injury, infection or adverse reactions to medications and accompanied by new neurological findings upon examination by the Examining Neurologist. The symptoms are attributed to MS and are preceded by 30 days of stability or improvement in neurological state. The change in EDSS score as assessed by the Examining Neurologist is ≥ 0.5 increase or ≥ 2 points increase on one of the appropriate FS or 1 point on two or more of the appropriate FS. The change must affect the selected FS (i.e. pyramidal, ambulation, cerebellar, brainstem, sensory or visual). Note, the change in FS scale scores should correspond to the patient's symptoms (e.g., patient reported change in visual acuity should correspond to a change in the vision FS score). Episodic spasm, sexual dysfunction, fatigue, mood change or ladder or bowel urgency or incontinence will not suffice to establish a relapse. Please note: sexual dysfunction and fatigue will not be scored. If the symptoms occur less than 30 days following the onset of a protocol defined relapse, it should be considered part of the same relapse and would not be treated with IVMP within the protocol. New or recurrent neurologic symptoms that evolve gradually over months should be considered disability progression, not an acute relapse and should not be treated with steroids.

The suspected relapse must be reviewed and confirmed by the IRAP; Appendix D of TG1101-RMS302. The IRAP will then notify the CRO and Treating Neurologist of the results of its review.

Treating Neurologist/designee will enter into the eCRF whether the neurological symptoms identified were a result of an IRAP-confirmed relapse.

The subject must re-consent at the time of each IRAP-confirmed or Treating Neurologist medically confirmed relapse to continue.

Signing the informed consent form for re-consent (if the subject has a confirmed relapse), to proceed with the study at the next scheduled office visit which should be within 6 weeks after first reporting of relapse.

Re-consent criteria

In case of an IRAP-confirmed diagnosis of MS relapse (as defined in the protocol, Appendix C) or Treating Neurologist medically confirmed relapse during the study, the following actions will be taken:

- The subject will be reminded of the current approved MS medications/treatments and the opportunity to terminate the study and be treated with an approved MS medication.
- The subject will be requested to re-sign an informed consent form within 6 weeks after first reporting of relapse if he/she chooses to continue to participate in the study, in the same treatment assignment.

2.1.3.2 Secondary Efficacy Variables

Key secondary efficacy variables are as follows:

1. Total number of gadolinium enhancing (Gd-enhancing) T1-lesions per MRI scan by Week 96.
2. Total number of new and enlarging T2 hyperintense lesions (NELs) per MRI scan by Week 96
3. Time to confirmed disability progression (CDP) for at least 12 weeks occurring during the 96-week, double-blind treatment period. *
4. Proportion of subjects with no evidence of disease activity (NEDA) from Week 24 to Week 96.
5. Proportion of subjects reaching impaired SDMT (Symbol Digit Modalities Test) from baseline to Week 96.
6. Percentage change in Brain Volume from baseline to Week 96.

* Confirmed Disability Progression for at least 12 weeks during the 96-week treatment period will be analyzed using pooled data from the two identical studies TG1101-RMS301 and TG1101-RMS302.

2.1.3.3 Tertiary Efficacy Variables

All tertiary variables will be compared between treatment groups at a Type I error of 0.05 with no adjustment for multiplicity. [redacted]

1. Change in MSFC score from baseline to Week 96
2. Time to Confirmed Disability Progression (CDP) for at least 24 weeks
3. Time to Confirmed Disability Improvement (CDI) for at least 12 weeks
4. Time to Confirmed Disability Improvement (CDI) for at least 24 weeks
5. Health outcomes (MSQoL-54 (inclusive of SF36); FIS, hospitalization, steroid use, time out of work)
6. Total volume of Gd enhancing T1 lesions per MRI scan over the treatment period
7. Volume of T2 lesions
8. Volume of hypointense T1 lesion component (black holes)
9. Proportion of subjects free of disability progression at 24 weeks, 48 weeks, and 96 weeks
10. The Proportion of subjects with a relapse
11. Time to first confirmed relapse

2.1.4 Safety Variables

The safety analysis will be based on the reported AEs and other safety information, such as clinical laboratory data, electrocardiogram (ECG), physical examination, and vital signs.

Baseline for non-efficacy variables is defined as last non-missing value before the first dose of study medication (teriflunomide/IV placebo or ublituximab/oral placebo). If the time of first dose intake or safety variables is missing and only date part is present, then the baseline will be the last non-missing value on or before the date of first dose date, unless the safety variable value is collected clearly after the administration of first dose.

2.1.4.1 Adverse Events

Adverse events will be coded and classified according to the System Organ Class (SOC) and Preferred Term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA).

The occurrence of AEs (including SAEs and adverse events of special interest [AESIs]) will be captured from the time the informed consent is obtained until the participant terminates participation in the study.

AESI will be categorized using a grouped set of MedDRA terms based on Standardized MedDRA Queries (SMQs) and clinical concept, as determined by the Sponsor prior to database lock. The MedDRA terms selected for AESI analysis will be provided in the Clinical Study Report (CSR).

Suicidal ideation and behavior will be identified by the AE preferred terms included in MedDRA SMQ "Suicide/self-injury". SMQ version 20 includes the following preferred terms: Assisted suicide, Columbia suicide severity rating scale abnormal, Completed suicide, Depression suicidal, Intentional overdose, Intentional self-injury, Poisoning deliberate, Self-injurious ideation, Suicidal behavior, Suicidal ideation, Suicide attempt, Suicide threat.

2.1.4.2 Laboratory Safety Variables

Clinical laboratory data consist of hematology, serum chemistry and urinalysis. Clinical laboratory values will be analyzed after conversion into standard international units. International units will be used in all listings and tables.

The laboratory parameters will be classified as follows:

Hematology

Hematology and differential panel: hemoglobin, hematocrit, red blood cell count, mean corpuscular hemoglobin, white blood cell count, neutrophils, lymphocytes, monocytes, eosinophils, basophils, and platelets

Serum chemistry/ Pancreatic enzyme panel
The serum chemistry to be analyzed are presented in the table below.

Serum Chemistry		
Albumin	Glucose	SGPT [ALT]
Alkaline phosphatase	LDH	Sodium
Bicarbonate	Magnesium	Total bilirubin
BUN	Phosphorus	Total Protein
Calcium	Potassium	Uric acid
Chloride	SGGT	Indirect/direct bilirubin
Creatinine	SGOT [AST]	

Pancreatic enzyme panel: serum amylase and lipase

Urinalysis

Urinalysis panel for routine safety: pH, ketones, protein, glucose, blood, urobilinogen, bilirubin, microscopic sediment, specific gravity and leucocyte esterase.

Coagulation
Fibrinogen

All observations of these tests including unscheduled measurements will be assigned to the appropriate safety analysis visit window, as defined in Section 2.5.3. In the presence of multiple measurements of the same laboratory test in the same window, the one closest to the targeted visit date will be used for the by-visit summaries. In case of equal distance, the later observation will be selected. All other lab results will only be listed.

2.1.4.3 Vital Signs

The following vital sign parameters will be analyzed:

- Systolic and diastolic blood pressures (mmHg)

- Heart Rate (beats/min)
- Weight (kg)
- Body Temperature (All results will be converted to Celsius)

2.1.4.4 Physical Examination

Physical examination is assessed per visit and body system, with possible outcomes normal and abnormal (plus free text for specification if abnormal).

2.1.4.5 ECG

ECG interpretation (normal/abnormal) will be collected by visit and - where applicable - time point (pre-/post-dose).

2.1.5 Pharmacokinetic Variables

To characterize population PK of ublituximab, serum samples will be drawn from all subjects pre-infusion any time on infusion days of ublituximab or IV placebo and at approximately 30 minutes +/- 15 minutes post infusion on infusion days (Week 1 Day 1, Week 3 Day 15, Weeks 24, 48, and 72). One additional serum sample will be taken on Week 96. [redacted]

2.1.6 Other Variables

- Immunogenicity-absence or presence of ADA to ublituximab (Anti-Drug Antibody; ADA)
- Pharmacodynamic parameter (% CD19+ B cells)

2.2 Analysis Populations

Randomized subjects are all subjects allocated based on the randomization process and recorded in the database. Subjects will be randomly assigned in a 1:1 ratio to either active teriflunomide or active ublituximab. Randomization will be generated by the IWRS.

2.2.1 Safety Population

The Safety Population will include all subjects who receive at least one dose of study drug (ublituximab or teriflunomide, with corresponding placebos). All safety assessments including toxicity and anti-drug antibodies will be performed on the Safety Population by the first treatment actually received.

2.2.2 Intention-to-Treat (ITT)

The Intention-to-Treat (ITT) population will consist of all randomized subjects. Subjects will be analyzed by randomized treatment group. Sensitivity analyses of key efficacy endpoints will be performed based on ITT population.

2.2.3 Modified Intention-to-Treat (mITT) Population

The modified Intention-to-Treat (mITT) population will consist of all subjects in the ITT population who received at least one dose of study medication and have at least one baseline and post baseline efficacy assessment. The primary efficacy analyses (of primary, secondary and tertiary efficacy endpoints, except MRI related) will be performed based on the mITT population.

The MRI related efficacy analyses will be performed on the subjects in mITT population who have baseline and post-baseline MRI efficacy assessments (mITT-MRI).

Subjects will be analyzed by randomized treatment group.

2.2.4 Per-Protocol (PP) Population

The Per-Protocol (PP) population will consist of all subject in the mITT group who were treated for at least 1.75 years and do not have a major protocol deviation that might impact efficacy analysis. The PP population will only be used for sensitivity analysis of the primary efficacy and key secondary variables. Subjects will be analyzed by randomized treatment group.

MRI related efficacy analyses will be repeated on the subjects in the PP population who have baseline and post-baseline MRI efficacy assessments (PP-MRI).

All reported protocol deviations will be reviewed and assessed for their potential impact on the efficacy analyses before database lock and unblinding. Subjects with one or more major protocol deviation that might impact the efficacy analysis will be excluded from the per-protocol population. Details will also be given in the protocol deviation plan.

Major efficacy-related protocol deviations include, but not limited to, the following:

- EDSS score >5.5 at screening
- Subjects with less than 2 clinical relapses in the last 2 years and less than 1 clinical relapse in last 1 year prior to screening and no Gd-enhancing lesion in the screening MRI
- Subject randomized more than once
- Compliance <80% (see Section 2.4.2)
- Subjects taking prohibited medication which are approved disease modified MS drugs during the study
- Non-randomized subject received double-blind study medication
- Screen failure subject randomized

- Randomized subject received incorrect treatment

The final list of subjects who are to be excluded from the per-protocol population will be documented prior to the data unblinding, with reason of exclusion and related protocol deviations as applicable.

All major protocol deviations will be summarized and listed.

[redacted]

[redacted]

2.3 Disposition of Subjects

Subject disposition will be tabulated as follows:

- Screened subjects defined as subjects who signed the informed consent;
- Screen failure subjects and reasons for screen failure;
- Not randomized but exposed to study medication
- Randomized subjects;
- Randomized but not exposed subjects;
- Randomized and exposed subjects;
- Subjects who completed the 96-week study treatment period;
- Subjects who discontinued study treatment and reasons for discontinuation, but continue to allow assessments in the study after stopping study medication;
- Subjects who terminate follow-up after treatment discontinuation/completion and the reasons for termination. These subjects stopped treatment with study medication and are no longer available for assessment during the study or willing to participate in any further assessments.

For all categories of subjects (except for the screened and non-randomized categories) percentages will be calculated using the number of randomized subjects per treatment group as the denominator.

Additionally, the analysis populations for safety and the three efficacy populations (ITT, mITT and PP) will be summarized in a table by subject counts on the randomized population.

2.4 Statistical Methods

The statistical analyses will be done once all subjects have completed Week 96 (Week 96 visit or earlier if withdrawn).

All statistical analyses will be performed at the 2-sided 0.05 level of significance, except as otherwise noted.

2.4.1 Demographics and Baseline Characteristics

Continuous data will be summarized using number with available data: mean, standard deviation (SD), median, minimum, and maximum, and 1st and 3rd quartile for each treatment group. Categorical or ordinal data will be summarized using the number and percentage of subjects within each treatment group. Summaries will be presented for the ITT, mITT and Safety population.

Demographics and Baseline Characteristics

Demographic and baseline characteristics as described in Section 2.1.1 will be summarized by treatment group and overall using descriptive statistics.

Medical and surgical history

Medical history and concurrent medical conditions (ongoing at screening) will be summarized by primary system organ class (SOC) and preferred term for each treatment group. The table will be sorted alphabetically by SOC and by decreasing frequency in the ublituximab column.

Prior/concomitant/follow-up medications

The prior, concomitant and follow-up medications will be presented on the safety population.

Medications will be summarized by treatment group, according to the international classification of medicines (WHO-DD dictionary), using the terms corresponding to the ATC class level 4 and Preferred Name. All ATC codes corresponding to a medication will be summarized. Subjects will be counted once in each ATC categories (anatomic or therapeutic) linked to the medication.

The tables for prior, concomitant and follow-up medications will be sorted by decreasing frequency of anatomic category followed by all other therapeutic classes based on the incidence in the Ublituximab group. In case of equal frequency regarding anatomic categories or therapeutic categories, alphabetical order will be used.

In addition, number and percentage of confirmed relapses requiring steroid treatment will be presented.

Previous MS treatment

Variables related to previous MS treatment will be summarized per treatment group.

No statistical tests will be performed on demographic and baseline characteristic data.

2.4.2 Primary Efficacy Variable

Main statistical model and adjustment for covariates

The primary analysis will be performed using the mITT with a Negative Binomial regression model to accommodate the potential over-dispersed data appropriately. The model will include the total number of confirmed relapses with onset between randomization date and the day of last treatment as response variable, treatment group, EDSS strata (baseline EDSS score ≤3.5 versus >3.5) and clinic region (defined in section 2.1.1) as covariates. There will be two treatment groups (teriflunomide or ublituximab). In order to account for different treatment durations among subjects, the log-transformed standardized treatment duration will be included in the model as an "offset" variable for appropriate computation of the annualized relapse rate. [redacted] will be used to assess the overall model with subjects in a repeated statement using a Generalized estimating equation (GEE) approach. The standardized treatment duration is defined as (date of last treatment – randomization date +1)/365.25.

Two-sided 95% confidence intervals of the rate ratio will be provided for the comparisons of ublituximab versus teriflunomide. The estimated relapse rate for each treatment group and the difference between them, together with the corresponding 2-sided 95% confidence intervals will be provided.

[redacted]

The "overall" estimated annualized relapse rates (number of relapses divided by the sum of years under treatment) will also be presented by treatment group.

Other analysis population for the primary efficacy variable

Analysis of annualized relapse rate will also be performed using the ITT and PP populations as a supportive analyses.

Handling of dropouts

Dropouts are handled by using the offset variable in the model which adjusts for the duration of time on treatment. This is different than multiple imputation, which seeks to estimate the information that would be available after dropping out. The primary analysis includes confirmed relapses with onset from randomization to date of last treatment. Subjects who withdraw from study drug treatment can still be followed and participate in study visits. However, the primary analysis will only include the time from randomization until the date of last treatment.

Subgroup analysis

Subgroup analyses will be performed for the primary efficacy variable by:

- Gender
- Race
- Age category (<38, ≥38)
    - Age of 38 is expected to be close to the median age of all randomized subjects
- Regions as defined in 2.1.1
- Baseline EDSS strata (<3.5, ≥3.5)
- Number of relapses experienced in the 2 years prior to study start (≤1, 2, 3, and ≥4)
- Received approved disease modifying MS drug prior to enrollment (Yes, No)
- Number of baseline Gd-enhancing lesions (0, ≥1)

The GEE method will be used to assess consistency of treatment effects across the subgroup levels, utilizing a treatment-by-subgroup interactions test at significance level of 0.05, and forest plots of the group comparisons of ARR and their confidence intervals. Each subgroup factor will be assessed separately. The model will include treatment, EDSS strata, region, subgroup, and treatment-by-subgroup interaction as fixed effects, log-transformed standardized treatment duration as an offset variable. In addition, summary statistics of relapse rates will be provided for each subgroup, based on the model for each subgroup, including treatment, EDSS strata and region as covariates.

Sensitivity Analysis

1. The primary analysis will be repeated based on all reported MS relapses (rather than just confirmed ones).

2. A time to relapses analysis using Cox proportional hazards model with first IRAP-confirmed relapse as the primary event will be conducted to compare with the primary analysis in a negative binomial model. The Cox proportional hazards model will be specified with treatment, region, number of relapses in previous year, baseline EDSS score strata (<3.5, ≥3.5), baseline number of T1 Gd-enhancing lesions, sex, and the subject's age at baseline as covariates.

3. To assess the effects of treatment withdrawals, the primary analysis will be repeated using the information on relapses which occur during the period following drug withdrawal to the end of the study period regardless of whether additional treatment were utilized. Relapses which occurred after permanent discontinuation of study medication will be included and natural log (time on study in years) rather than the natural log (time on study drug in years) will be used as the offset variable in the negative binomial model, which is the primary analysis.

4. We will examine the dropouts from the trial using all participants who provided withdrawal of consent excluding deaths. Multiple imputation will be used to impute the expected number of relapses a participant would have had if they had continued to participate in the trial. To perform the imputation, 10 replicates will be used since a relapse on treatment is relatively low frequency event. The covariates used to model and predict the imputations will include: treatment, region, number of relapses in previous year, baseline EDSS score strata (<3.5, ≥3.5), baseline number of T1 Gd-enhancing lesions, sex and the subject's age at baseline as covariates.

5. For secondary outcomes T1 Gd enhancing, T2 NELs, and 12 week CDP, the effects of dropouts and study discontinuations will be similarly evaluated as number 4 above.

6. Additional sensitivity analyses may be conducted as needed

2.4.3 Secondary Endpoints

2.4.3.1 Statistical testing Strategy

Once the null hypothesis of the primary analysis has been rejected, the key secondary outcomes will be tested using a hierarchical approach with the order specified below using a step-down procedure where each test is at a Type I error 0.05. If any endpoint fails to reach significance, then formal testing of significance of the subsequent secondary outcomes will not be performed. All p-values will be reported and will be interpreted as either confirmatory or non-confirmatory (i.e., descriptive only).

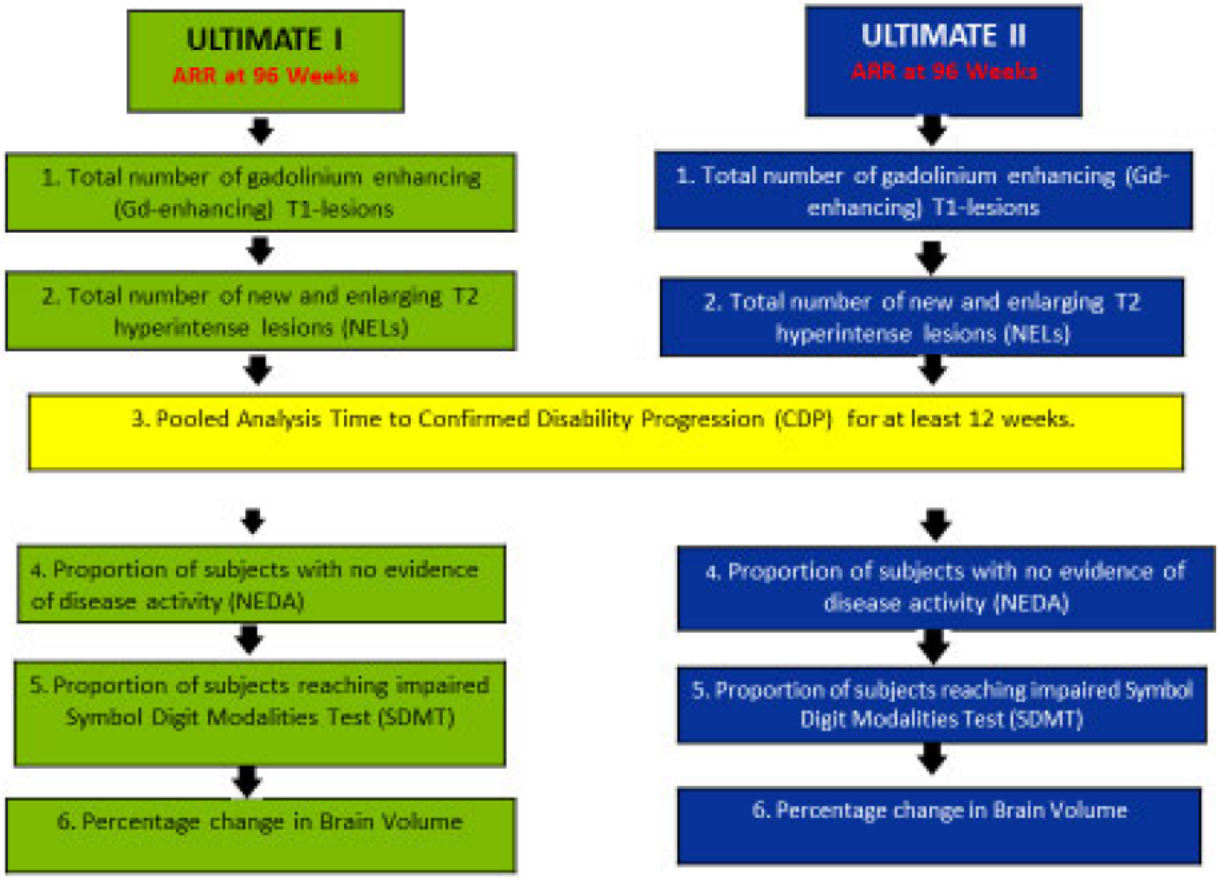
1. Total number of gadolinium enhancing (Gd-enhancing) T1-lesions per MRI scan by Week 96.
2. Total number of new and enlarging T2 hyperintense lesions (NELs) per MRI scan by Week 96
3. Time to confirmed disability progression (CDP) for at least 12 weeks occurring during the 96-week, double-blind, treatment period (pooled analysis with Study TG1101-RMS302). *
4. Proportion of subjects with no evidence of disease activity (NEDA) from Week 24 to Week 96.
5. Proportion of subjects reaching impaired SDMT (Symbol Digit Modalities Test) from baseline to Week 96.
6. Percentage change in Brain Volume from baseline to Week 96.

* Confirmed Disability Progression for at least 12 weeks during the 96-week treatment period will be analyzed using pooled data from the two identical studies TG1101-RMS301 and TG1101-RMS302.

For all tests other than #3 (12-week CDP), the testing will be based on Study ULTIMATE II individually, at alpha=0.05.

For #3, 12-week CDP endpoint, the testing will be based on pooled data from both Studies ULTIMATE I and ULTIMATE II achieve sufficient statistical power to detect relevant treatment differences. Therefore, this test will not be performed until data from both studies are available. Specifically, this test is considered confirmatory only if all tests for the primary endpoint and all secondary endpoints before this test (i.e. #1 & #2) are statistically significant at alpha=0.05 level in both studies. Once this test is statistically significant, all subsequent tests (#4-6) will continue at alpha=0.05 level within each individual study.

The statistical testing sequence is illustrated below.



2.4.3.2 MRI

The following MRI results will be provided by the central lab:

Measurement/Visit	SCREENING	WEEK 12 DAY 84	WEEK 24 DAY 168	WEEK 48 DAY 336	WEEK 96 DAY 672	EARLY WITHDRAWAL FROM TREATMENT
T1 Unenhancing Lesion Count	X					
T2 Lesion Count	X					
New Enlarging T2 Lesion Count			X (Screening)	X (Previous)	X (Previous)	X (Previous)
New Enlarging T2 Lesion Volume			X (Screening)	X (Previous)	X (Previous)	X (Previous)
T2 Lesion Volume	X		X	X	X	X
New T1 Unenhancing Lesion Count			X (Screening)	X (Previous)	X (Previous)	X (Previous)
T1 Unenhancing Lesion Volume	X		X	X	X	X
Normalized Brain Volume	X					
Brain Volume			X	X	X	X
Gad Enhancing Lesion Count	X	X	X	X	X	X

Gad Enhancing Lesion Volume	X		X	X	X	X

The analysis endpoints will be derived as follows:

Protocol endpoint	MRI test result to be used	Derivation
total number of gadolinium enhancing (Gd-enhancing) T1-lesions per MRI scan by Week 96	Gad Enhancing Lesion Count (GADCNT)	[GADCNT(w12)+GADCNT(w24)+GADCNT(w48)+GADCNT(w96)]/[number of MRIs available (w12/w24/w48/w96)] Note: the division by the number of MRIs will be accomplished by an offset in the negative binomial regression
Total number of new and enlarging T2 hyperintense lesions (NELs) per MRI scan by Week 96	New Enlarging T2 Lesion Count (NEWT2CNT)	[NEWT2CNT(w24)+ NEWT2CNT(w48)+ NEWT2CNT(w96)]/ [number of MRIs available (w24/w48/w96)] Note: the division by the number of MRIs will be accomplished by an offset in the negative binomial regression
Volume of hypointense T1 lesion component (black holes)	T1 Unenhancing Lesion Volume (T1UNVOL)	T1UNVOL (w24) , T1UNVOL (w48) ,T1UNVOL (w96)
Volume of T2 lesion	T2 Lesion Volume (T2VOL)	T2VOL(w24), T2VOL(w48),T2VOL(w96)
MRI part of NEDA definition (no T1 Gd+ lesions and no new/enlarging T2 lesions)	Gad Enhancing Lesion Count (GADCNT), New Enlarging T2 Lesion Count (NEWT2CNT)	All values , GADCNT(w24), GADCNT(w48), GADCNT(w96), NEWT2CNT(w24), NEWT2CNT(w48), NEWT2CNT(w96) need to be zero
Percent change in brain volume as detected by MRI compare to baseline	Brain Volume (BV)	(BV(w24) – NBV(baseline))/NBV(baseline)*100 (BV(w48) – NBV(baseline))/NBV(baseline)*100 (BV(w96) – NBV(baseline))/NBV(baseline)*100
Total volume of Gd-enhancing T1 lesions per MRI scan over the treatment period	Gad Enhancing Lesion Volume (GADVOL)	[Sum of all post-baseline GADVOL(w24/w48/w96)] / [number of MRIs available (w24/w48/w96)]

MRI data will only be used for analysis if not rejected or invalid, e.g. due to change of scanner, indicated as comments from the central lab.

MRIs will also be excluded from analysis if corticosteroids (ATC level 2 = "CORTICOSTEROIDS FOR SYSTEMIC USE") were dosed within 30 days prior to the scan.

The MRI count variables (Total Number of gadolinium enhancing (Gd-enhancing) T1-lesions per MRI scan at Weeks 12 (where applicable), 24, 48, and 96, Total number of new and enlarging T2 hyperintense lesions (NELs) per MRI scan at Weeks 24, 48, and 96) will be assessed for differences between the treatment groups using negative binomial regression with an offset based on the log-transformed number of post-baseline MRI scans and covariates region, baseline EDSS strata, and corresponding baseline lesion count.

For analysis of percent brain volume change from baseline, a Mixed Model Repeated Measures (MMRM) analysis will be implemented via [redacted] by fitting percent change from baseline values on the cube root transformed volume from all visits after baseline in the treatment period with scheduled assessments of brain volume. The models will include factors (fixed effects) for treatment, region, baseline EDSS strata, visit, treatment-by-visit interaction, and baseline value of

the brain volume (cubic root transformed). The factor, visit, with nominal visits to avoid undue reliance on a linear in time effect will have 3 levels (Week 24, Week 48 and Week 96). An unstructured correlation matrix will be used to model the within-patient errors. Parameters will be estimated using restricted maximum likelihood method with the Newton-Raphson algorithm. Denominator degrees of freedom will be estimated using Satterthwaite's approximation.

2.4.3.3 NEDA

The proportion of subjects with no evidence of disease activity (NEDA) will be calculated at Week 96. A subject with NEDA is defined as a subject without evidence of disease activity, which is defined as a subject without relapses confirmed by the IRAP, without MRI activities (no T1 Gd+ lesions and no new/enlarging T2 lesions), and no 12-week Confirmed Disability Progression.

1. Any evidence of disease activity from Week 24 to Week 96 will be counted as not reaching NEDA. Any evidence of disease activity before Week 24 will not count.
2. In the case of early termination at any time point (including before Week 24), even if an event was not reported before early discontinuation, the subject will be considered as not reaching NEDA.

NEDA rates will be compared using logistic regression [redacted] with baseline adjustments the same as used in the primary endpoint analysis, without treatment duration offset, but including log-transformed baseline MRI counts (T1 unenhancing, T2, Gad enhancing).
To avoid zero values for the log transformation of MRI counts, 1 will be added to each observation before transforming.

2.4.3.4 SDMT

The total score at one visit is defined as the total number of correct answers reported in the CRF. Impaired SDMT is defined as a decrease from baseline of at least four points at any post-baseline SDMT assessment up to the Week 96 visit. The proportion of impaired SDMT will be analyzed in all subjects within the modified intention to treat (mITT) population.

SDMT rates will be compared using logistic regression [redacted] with baseline adjustments the same as used in the primary endpoint analysis, without treatment duration offset, but including log-transformed baseline MRI counts (T1 unenhancing, T2, Gad enhancing). To avoid zero values for the log transformation of MRI counts, 1 will be added to each observation before transforming
.

2.4.3.5 EDSS/Disability

12 week Confirmed Disability Progression (CDP) is defined as an increase in EDSS

- at least 1 point higher than the baseline EDSS if the baseline EDSS is ≤5.5
- at least 0.5 higher than the baseline EDSS if the baseline EDSS is >5.5

Disability progression is considered confirmed (CDP) when the increase in the EDSS is confirmed (same criteria) at a regularly scheduled visit 12 weeks after the initial documentation of neurological worsening (does not include unscheduled visits).

The time to onset of 12 week confirmed disability progression (CDP) will be the time to progression to the EDSS change defined above. For each of these events, the time from randomization to the date of the first measurement of the increased EDSS as required is the "time to event". If no event occurs, the time to CDP is regarded as censored at the last scheduled EDSS assessment. Summary tables of progression to CDP will present the proportion of subjects who suffered CDP during the study at 2 years and associated 95% CIs for each treatment group. The median time-to-event with 2-sided 95% CIs as well as the proportion of subjects remaining event-free at times of interest will be estimated using Kaplan-Meier methods implemented with [redacted]. The proportion of subjects with 12-week confirmed disability progression will be summarized.

EDSS results will be summarized as continuous and categorical variables by scheduled assessment time point and treatment group. All data collected for assessment of EDSS will be listed.

2.4.4 Tertiary Endpoints

All tertiary analyses will be assessed at a Type I error of 0.05 with no adjustment for multiplicity[redacted].

1. Change in MSFC score from baseline to Week 96
2. Time to Confirmed Disability Progression (CDP) for at least 24 weeks
3. Time to Confirmed Disability Improvement (CDI) for at least 12 weeks
4. Time to Confirmed Disability Improvement (CDI) for at least 24 weeks
5. Health outcomes (MSQoL-54 (inclusive of SF36); FIS, hospitalization, steroid use, time out of work)
6. Total volume of Gd enhancing T1 lesions per MRI scan over the treatment period
7. Volume of T2 lesions
8. Volume of hypointense T1 lesion component (black holes)
9. Proportion of subjects free of disability progression at 24 weeks, 48 weeks, and 96 weeks
10. Proportion of subjects with a relapse
11. Time to first confirmed relapse

2.4.4.1 MSFC

Items collected per visit		
Timed 25-Foot Walk	Time for Trial 1 (sec)	TW1(visit)
	Time for Trial 2 (sec)	TW2(visit)
9-Hole Peg Test	Dominant Hand: Time for Trial 1 (sec)	PTD1(visit)
	Dominant Hand: Time for Trial 2 (sec)	PTD2(visit)
	Non-Dominant Hand: Time for Trial 1 (sec)	PTN1(visit)
	Non-Dominant Hand: Time for Trial 2 (sec)	PTN2(visit)
PASAT 3 seconds	PASAT 3" Total Score	PAS(visit)

Summarization steps

Average 25-Foot Walk (summarizing the two trials of one assessment)	TW(visit) = Mean(TW1(visit), TW2(visit))	If one of the two (TW1, TW2) is missing at a visit, use the available value for TW If both are missing, use instructions in the next row.
Z-score leg	$Z_{leg,average}$(visit) = - (TW(visit) – mean(TW(baseline, all subjects in analysis population))/SD(TW(baseline, all subjects in the analysis population)))	Note: The leftmost minus sign is intended to assign the same direction (higher=better) as for PASAT If subject's input TW(visit) is not available, the Z-score is set to -13.7.
Average 9-Hole Peg Test (summarizing the 4 trials – two of each hand)	PT(visit) = mean((1/mean(PTD1(visit), PTD2(visit)), 1/mean(PTN1(visit), PTN2(visit)))	Note: means are inverted to assign the same direction (higher=better) as for PASAT. If one of the two (PTx1, PTx2) is missing at a visit, use the available value. If both are missing for a hand, use 777 for this hand.
Z-score arm	$Z_{arm,average}$(visit) = (PT(visit) – mean(PT(baseline, all subjects in analysis population))/SD(PT(baseline, all subject in analysis population)))	
Z-score cognitive	$Z_{cognitive}$ = (PAS(visit) – mean(PAS(baseline, all subjects in analysis population)))/SD(PAS(baseline, all subjects in analysis population))	
MSFC score	MSFC(visit) = $Z_{leg,average}$(visit) + $Z_{arm,average}$(visit) + $Z_{cognitive}$	

Change in MSFC will be tested using linear Mixed Models including all scheduled assessment time points with baseline as a covariate along with the other covariates as used in the primary analysis.

2.4.4.2 Disability endpoints

24 week Confirmed Disability Progression (24 Week CDP) is defined as an increase of ≥ 1.0 point from the baseline EDSS score that is not attributable to another etiology (e.g., fever, concurrent illness, or concomitant medication – this information will be reflected in the assessments documented by the investigator. In the statistical analysis no further checks are necessary.) when the baseline score is 5.5 or less, and ≥ 0.5 when the baseline score is above 5.5. Disability progression is considered confirmed when the increase in the EDSS is confirmed at a regularly scheduled visit at least 24 weeks after the initial documentation of neurological worsening.

12 Week Confirmed Disability Improvement (12 Week CDI) is defined as a reduction from the baseline EDSS score of at least 1.0 point (or 0.5 points if the baseline EDSS score was >5.5) that was sustained and confirmed at the next scheduled visit at least 12 weeks after the initial documentation of neurological improvement. Similarly, 24-week Confirmed Disability Improvement (24 Week CDI) requires an initial reduction from baseline EDSS score and a subsequent confirmation of the reduction at all regular scheduled visits for at least 24 Weeks after the initial documentation of neurological worsening.

The analysis of the time to 24-week CDP and CDI at 12 and 24 weeks will utilize the same approach as that used for 12-week CDP (including subject proportions).

2.4.4.3 Fatigue Impact Scale

The FIS is composed of 40 items that assess the individual's attribution of functional limitations to their subjective experience of fatigue in 3 dimensions.

- 0 - "No problem"
- 1 - "Small problem"
- 2 - "Moderate problem"
- 3 - "Big problem"
- 4 - "Extreme problem"

(Lower score = Less fatigue impact)

Cognitive Dimension Score = sum of items 1; 5; 6; 11; 18; 21; 26; 30; 34; 35

Physical Dimension Score = sum of items 10; 13; 14; 17; 23; 24; 31; 32; 37; 38

Social Dimension Score = sum of items 2; 3; 4; 7; 8; 9, 12; 15; 16, 19; 20; 22; 25; 27; 28; 29; 33; 36; 39; 40

Total Score = sum of all items (1-40)

If up to four items are missing within one questionnaire, these items will be imputed with the median of the available observations of the same questionnaire for the Total Score derivation. If more than four items are missing, the Total Score will be regarded as missing. The dimension scores will only be calculated if all included items are available.

Change in FIS will be tested using linear Mixed Models including all scheduled assessment time points with baseline as a covariate along with the other covariates as used in the primary analysis

2.4.4.4 MSQOL-54

The available software will be used to summarize the questionnaire data. In a first step, SDTM datasets containing response items will be transformed into a file format appropriate to input to the scoring software. The output data of the scoring software will be transformed into ADaM format for statistical analysis.

Change in MSQoL-54 will be tested using linear Mixed Models including all scheduled assessment time points with baseline as a covariate along with the other covariates as used in the primary analysis

2.4.4.5 Time out of work

Diaries will be summarized for periods of approximately 48 weeks and the overall treatment period, calculating, per period:

- Sum of hours worked (not tabulated)

- Sum of hours missed
- Total work hours (worked+missed) (not tabulated)
- Percentage of hours missed (missed / worked+missed).

The percentages of missed work hours will be compared between arms using Wilcoxon rank sum tests.

2.4.4.6 Steroid use

- Number of IRAP confirmed relapses treated with steroid
- Percentage of IRAP confirmed relapses treated with steroid

The number of IRAP confirmed relapses treated with steroid will be analyzed the same way as the primary endpoint.

2.4.4.7 Hospitalization due to (suspected) MS relapse

- Days in hospital due to suspected MS relapse
- Percentage of days in hospital due to suspected MS relapse (with all days in hospital as the denominator)
- Percentage of subjects hospitalized due to suspected MS relapse

[redacted]

2.4.4.9 Immunogenicity

Blood samples will be collected per study protocol for determination of ADA.

f

2.4.4.10 Analysis of other variables

The proportion of subjects hospitalized due to suspected MS relapse and the proportion of subjects receiving steroid as treatment for IRAP confirmed relapse will be analyzed the same way as NEDA.

For analysis of lesion volume related variables (Total volume of Gd enhancing T1 lesions per MRI scan over the treatment period, Volume of T2 lesions, Volume of hypointense T1 lesion component

(black holes)) Mixed Model Repeated Measures (MMRM) analyses will be implemented via [redacted]

For every valid MRI, cubic root transformed volumes will be used for analysis. For each time point, the sum of these values for all scheduled post-baseline MRIs, divided by the number of MRIs contributing to this sum will be used as the analysis value. Change from baseline in these cubic root transformed, cumulative values will be analyzed. The descriptive analysis will use the untransformed data.

The models will include factors (fixed effects) for treatment, region, baseline EDSS strata, visit, treatment-by-visit interaction, and baseline value (cubic root transformed). The factor, visit, with nominal visits to avoid undue reliance on a linear in time effect will have 3/4 levels (Week 12 – where applicable - , Week 24, Week 48, and Week 96). An unstructured correlation matrix will be used to model the within-patient errors. Parameters will be estimated using restricted maximum likelihood method with the Newton-Raphson algorithm. Denominator degrees of freedom will be estimated using Satterthwaite's approximation.

Time to first confirmed relapse is defined as (date of relapse onset – date of randomization + 1) and will be regarded as censored at the end of treatment. The analysis will be similar to the one of time to confirmed disability progression. The Nelson–Aalen estimator is a non-parametric estimator of the cumulative hazard rate function in case of censored data or incomplete data and will be provided for descriptive purposes.

The proportions of subjects with a relapse and subjects free of disability progression at different time points will be estimated using the Kaplan-Meier method.

2.4.5 Analyses of Safety Data

Safety results will be summarized and presented by treatment group.

Observation period

The observation period is defined as the time from first dose of study medication up to 20 weeks after the last dose of study medication or until the final subject completes the Week 96 assessment.

General common rules

All safety analyses will be performed on the Safety Population as defined in Section 2.2.1, unless otherwise specified, using the following common rules:

1. For quantitative safety parameters based on central laboratory/ reading measurements, visit and treatment group will use descriptive statistics to summarize results and change from baseline values. Summaries will include the value and/or the worst value of the observation period.
2. All of the values including unscheduled measurements will be assigned to the appropriate safety analysis visit window. The safety analysis visit window is defined in Section 2.5.2. In the presence of multiple measurements of the same test in the same window, the one closest to the targeted visit date will be used for the by-visit summaries.

3. The analysis of the safety variables will in general be descriptive and no systematic statistical testing is planned.

2.4.5.1 Analysis of Adverse Events

Adverse event classification:

1. Pre-existing AEs are defined as adverse events that started prior to treatment start and are ongoing at the time of treatment start. After treatment start, if these adverse events worsened during the treatment period these will be considered as treatment-emergent adverse events (TEAEs).
2. TEAEs are defined as AEs with an onset on, or after the first dose of study medication and through 30 days after the last dose of study medication or with an onset prior to the first dose of study medication that increases in severity on, or after the first dose of study medication and through 30 days after the last dose of study medication.
3. Post-treatment AEs are defined as AEs that developed or worsened during the post treatment period after stopping study medication. As an example, AEs occurring during the 20-week follow-up involving rapid teriflunomide/oral placebo elimination would be considered a post-treatment AE.

Data will be analyzed as available in the database as not all subjects are expected to be followed up for the full period e.g. due to switch to the open-label extension.
Adverse event incidence tables will be presented by treatment group, the number (n) and percentage (%) of subjects experiencing an AE. The incidence of events will be presented: A subject will only be counted once in the tables within a treatment phase. In a subset of tables, also all occurrences will be counted providing a severity of occurrence of the AE. The denominator for computation of percentages is the Safety Population within each treatment group.

Not treatment emergent AE(s) summary:

- Listing of (not treatment emergent) AEs showing treatment group, subject identifier, SOC, PT, verbatim term, date and day of onset, end date/time, duration, outcome, severity, date of death (if relevant), relationship to study treatment, action taken, AE seriousness and criteria.

TEAE(s) summaries:

The following TEAE summaries will be generated for the Safety Population.

Overview of TEAEs, summarizing number (%) of subjects with any

- (related/all) TEAE
- (related/all) Serious TEAE
- (related/all) TEAE leading to death
- (related/all) TEAE of grade 3 or higher
- (related/all) TEAE leading to drug interruption
- (related/all) TEAE leading to study discontinuation
- (related/all) TEAE of special interest (AESI)
- (related/all) TEAE leading to permanent discontinuation

- TEAEs by primary SOC and PT, showing number (%) of subjects with at least one TEAE, sorted alphabetically by SOC and decreasing incidence of PTs within SOC for the ublituximab arm. This sorting order will be applied to all other tables, unless otherwise specified.
- TEAEs by PT, sorted by decreasing frequency in the ublituximab arm
- Common TEAEs (incidence ≥ 2% in any treatment group for PT) by primary SOC and PT, showing number (%) of subjects with at least one common TEAE.
- All at least possibly drug related (to IV and/or oral study treatment) TEAEs by primary SOC and PT, showing number (%) of subjects with at least one possibly related TEAE, sorted by sorting order defined above.
- All TEAEs by maximal severity, presented by primary SOC and PT, showing number (%) of subjects with at least one TEAE by severity (ie, mild, moderate, or severe), sorted by sorting order defined above.
- Listing of TEAEs showing treatment group, subject identifier, SOC, PT, verbatim term, date and day of onset, end date/time, duration, outcome, severity, date of death (if relevant), relationship to study treatment, action taken, AE seriousness and criteria.

AESI summaries:

- Treatment-emergent AESI by primary SOC and PT, showing number (%) of subjects with at least one TEAE, sorted alphabetically by SOC and decreasing incidence of PTs within SOC for the Ublituximab arm. This sorting order will be applied to all other tables, unless otherwise specified.
- Listing of treatment-emergent AESI showing treatment group, subject identifier, SOC, PT, verbatim term, date and day of onset, end date/time, duration, outcome, severity, date of death (if relevant), relationship to study treatment, action taken, AE seriousness and criteria.

Treatment-Emergent Suicidal Ideation and Behavior (SIB)

- Treatment-emergent SIB events will be summarized by treatment group, showing number (%) of subjects with at least one TEAE of suicidal ideation and behavior, and number (%) of subjects for each PT.
- A listing of suicidal ideations and behaviors will also be presented.
- Additional SIB analyses may be conducted as needed, such as follow-up time adjusted incidence rates of SIB event.

Serious TEAE(s) summaries:

- All serious TEAEs by primary SOC and PT, showing number (%) of subjects with at least one serious TEAE, sorted by sorting order defined above.
- All at least possibly related (to IV and/or oral study treatment) serious TEAEs by primary SOC and PT, showing number (%) of subjects with at least one possibly related serious TEAE, sorted by sorting order defined above.
- All serious TEAEs by seriousness criteria, presented by primary SOC and PT, showing number (%) of subjects with at least one serious TEAE by seriousness criteria.

- Listings of SAEs showing treatment group, subject identifier, SOC, PT, verbatim term, date and day of onset, end date/time, duration, outcome, severity, date of death (if relevant), relationship to study treatment, action taken, AE seriousness and criteria.

TEAE(s) leading to treatment discontinuation summaries:

These events will be selected by the eCRF variables "Action taken with IV study treatment" and/or "Action taken with Oral study treatment" = "Permanently discontinued".

- All TEAEs leading to treatment discontinuation, by primary SOC and PT, showing number (%) of subjects with at least one TEAE sorted by sorting order defined above.
-
- Listings of TEAEs leading to treatment discontinuation showing treatment group, subject identifier, SOC, PT, verbatim term, date and day of onset, end date/time, duration, outcome, severity, date of death (if relevant), relationship to study treatment, action taken, AE seriousness and criteria.

Deaths summaries:

- Number (%) of subjects who died during the study period will be shown by treatment group.
- Death in non-randomized subjects or randomized but not treated subjects will be listed.
- All TEAEs leading to death, by primary SOC and PT, showing number (%) of subjects sorted by sorting order defined above.
- Listings of deaths showing treatment group, subject identifier, gender, race, age, duration of exposure, date of death (if relevant), causes and circumstances surrounding the death

2.4.5.2 Analysis of Laboratory Variables

Clinical laboratory variables that will be analyzed include hematology tests, serum chemistry tests, pancreatic enzymes, a coagulation test, and urinalysis (see Section 2.1.4.2 for the complete list of parameters).

Summary statistics (including mean, SD, median, Q1, Q3, minimum and maximum) of all laboratory variables (raw data and changes from baseline) will be calculated for each visit. Mean changes from baseline with the corresponding standard deviation will be plotted over time (at same time points) in each treatment group for selected parameters using box plots, which include hemoglobin, lymphocytes, neutrophils, platelets, ALT, AST, TBL, and serum creatinine (and CrCL).

Subjects with laboratory values outside of the normal reference range at any post-baseline assessment will be summarized and graded per NCI CTCAE Version 4.03 when applicable. Subject incidence of laboratory toxicity will be summarized by treatment group and maximum grade for each laboratory test. Shift tables between baseline and worst post-baseline results will also be presented.

A listing of subjects with at least one post-baseline abnormal lab test will be provided and will display the complete profile over time of all values of corresponding test. In this listing, baseline, endpoint value and individual values will be flagged when lower or higher than the lower or upper laboratory limits and/or when reaching the absolute limit of abnormality criteria.

Drug-Induced Liver Injury (DILI)

The liver function tests, namely AST, ALT, total bilirubin (TBL), and alkaline phosphatase are used to assess possible drug induced liver injury. The proportion of subjects with abnormal values at any post-baseline visit by baseline status will be displayed by treatment group for each parameter. The proportion of subjects with abnormal values at any post baseline visit will also be displayed by duration of exposure for each treatment group.

Time to onset of the initial ALT and AST elevation (>3x ULN) and total Bilirubin elevation (>2x ULN) will be analyzed using Kaplan-Meier estimates, presented by treatment group. Evaluations of drug-induced serious hepatotoxicity (eDISH) plot (Hy's Law graphs) of distribution of peak values of ALT, AST and the maximum of ALT and AST versus peak values of total bilirubin will also be presented. Note that the ALT/AST and Total bilirubin values will be presented on a logarithmic scale. The eDISH plots will be divided into 4 quadrants with a vertical line corresponding to 3x ULN for ALT and/or AST and a horizontal line corresponding to 2x ULN for total bilirubin. Similarly, a graph for ALT vs. TBL and a graph for AST vs. TBL will be provided.

The normalization (to ≤1 ULN) of elevated liver function tests will be summarized by categories of elevation (≥3x, ≥5x, ≥8x, ≥10x, ≥20x ULN for ALT, AST, ALT or AST, ≥1.5x, ≥2x ULN for alkaline phosphatase, and ≥1.5x, ≥2x ULN for total bilirubin), with following categories of normalization: not normalized, normalized after temporary discontinuation of study medications, normalized after permanent discontinuation of study medications. Note that a subject will be counted only under the maximum elevation category.

2.4.5.3 Analyses of Vital Sign Variables

The summary statistics (including mean, SD, median, Q1, Q3, minimum and maximum) of all vital signs variables (raw data and changes from baseline) will be calculated for each visit by treatment group and presented. Mean changes from baseline with the corresponding standard deviations will be plotted over time (at same time points) in each treatment group.

The incidence of abnormal vital signs at any time during the observation period will be summarized by treatment group. A shift table from baseline to worst vital sign status during the observation period will be presented.
A listing of subjects with at least one post-baseline abnormal vital signs observation will be provided and will display the whole profile over time of all parameters.

2.4.5.4 Analysis of Physical Examination Variables

The incidence of abnormal findings per body system will be tabulated by visit and for the whole post-baseline observation period. Free text specifications will be listed.

2.4.5.5 Analysis of Electrocardiogram Variables

ECG interpretation will be tabulated as frequencies by visit and time point, and in a shift table of baseline vs. worst post-baseline observation. Results per time point and free-text specifications of abnormalities will be listed.

2.4.6 Duration of Study Treatment Exposure and Compliance

The duration of study treatment exposure and compliance will be summarized by actual treatment within the safety population as continuous and categorical variables.

Duration of oral study treatment exposure

Duration of exposure (to oral study treatment/placebo) will be summarized descriptively as a quantitative variable (N, mean, SD, median, Q1, Q3, minimum, and maximum).

The cumulative duration of treatment exposure, defined as the sum of subjects' duration of treatment exposure in subject years, will be summarized by treatment group.

Treatment compliance / Exposure to IV study treatment

Treatment compliance for both ublituximab/IV placebo and teriflunomide/oral placebo will be summarized as quantitative variables (N, mean, SD, median, Q1, Q3, minimum, and maximum). The percentage of subjects who have a compliance of <80%, 80% to <100%, and >=100% will be tabulated by treatment group.

Furthermore, for teriflunomide or oral placebo, the proportion of months on at least 80 percent of prescribed oral medication will be calculated with the following algorithm:

1. For every period between a dispensation of tablets and the succeeding dispensation (or the last return), the ratio #tablets taken/#duration or the period in days is calculated

2. This calculated ratio is assigned to each day of the containing period

3. For each full 28 day period starting on the first day of oral treatment (or the potentially shorter last period) the fractions are added and divided by the number of days in this period. This result is compared with the defined threshold (0.8 / 80%), and for each subject the percentage of periods reaching the threshold will be calculated and summarized.

2.4.7 Analyses of Pharmacokinetics

Pharmacokinetics for ublituximab

Population PK (nonlinear mixed effect modeling) analyses will be performed by an independent vendor and reported separately. This analysis will assess dose- and time-dependent kinetics as well

as covariates effects on the PK of ublituximab. [redacted] Exposure-response relationships for efficacy and safety will be evaluated to determine the recommended therapeutic dose.

The population PK and exposure-response analyses will be conducted using the PK population that will consist of subjects who have at least one baseline and one post-baseline PK assessment. Subjects will be excluded from the PK analysis if data are incomplete which may influence the analysis. Excluded cases will be documented together with the reason for exclusion.

Pharmacokinetics data will be reported based on all subjects treated with ublituximab with any PK samples. [redacted]

Plasma Concentration for teriflunomide

In order to assess teriflunomide compliance, four plasma samples will be taken from all subjects. The first sample will be taken prior to oral administration of teriflunomide or oral placebo on Week 1 Day 1. The second sample will be taken prior to infusion of ublituximab or IV placebo on Week 48. The third sample will be taken at Week 96, prior to first dose of elimination therapy. The last plasma sample for teriflunomide will be taken at Week 116 and will be used to determine that teriflunomide has been eliminated from all subjects following the accelerated elimination procedure with either active charcoal or cholestyramine. For subjects who withdraw from treatment early, a blood sample will be collected at the Early Withdrawal from Treatment visit prior to first dose of elimination therapy.

Only plasma samples from subjects treated with teriflunomide will be analyzed for plasma teriflunomide concentrations. See the Laboratory Manual for PK collection and processing instructions.

The plasma concentration of teriflunomide for each subject will be presented for each time point in a data listing. Descriptive statistics will be used to describe the plasma concentration for each timepoint. A timeplot of the average plasma concentration over time will be constructed.

2.4.8 Analyses of Anti-Drug Antibody (ADA)

Immunogenicity to determine the absence or presence of ADA will be done for only the subjects randomized to ublituximab. Immunogenicity results, specifically the impact of ADA on the PK of ublituximab, will be reported in conjunction with the population PK analysis. . The percentage of

subjects developing ADA will be summarized by actual treatment group. A listing will be provided for ADA data.

2.4.9 Analysis of Pharmacodynamics

B lymphocyte cell counts (% CD19+ B cells) will be tabulated per scheduled time point together with absolute and percentage changes from baseline.

2.5 Data Handling Conventions

Data handling discussions for the primary efficacy is included in Section 1.2.1. The further discussion for the data handling conventions for other variables will be included in this section.

General conventions

The following formulas will be used for computations of parameters.

Demographic formulas:

The age variable recorded in the eCRF will be used for analysis, no derived age variable is planned as only the year of birth is reported.

$$\text{BMI (in kg/m}^2) = \frac{\text{weight in kg}}{\text{height (in cm)} \times \text{height (in cm)}} \times 10000$$

Renal function formulas:

Creatinine clearance value will be derived using the equation of Cockroft and Gault. The calculation requires body weight and age. The last weight measurement on or before the visit of the creatinine measurement will be used. If available, the age used by central lab at the visit will also be used for this derivation, otherwise the age at informed consent will be used.

$$crcl\ (mL/\min) = \frac{(140 - age) \times weight(kg)}{0.814 \times creatinine(umol/L)} \quad \text{(For males)}$$

$$crcl\ (mL/\min) = \frac{(140 - age) \times weight(kg)}{0.814 \times creatinine(umol/L)} \times 0.85 \quad \text{(For females)}$$

2.5.1 Missing Data

For categorical variables, subjects with missing data are not included in calculations of percentages unless otherwise specified. The number of subjects with missing data will be presented.

For data listings, the character date will always be used to present the date collected in the eCRF, which allows the actual recorded date to be shown if elements are unknown/missing.

2.5.1.1 Multiple sclerosis medical history variables

If the date of first MS diagnosis, first symptoms of MS, most recent relapse onset or start and end of previous MS treatments are incomplete, (i.e., month is unknown or day is unknown), unknown month/day will be set to July 1st, an unknown day will be set to the 15th.

2.5.1.2 Date of first and last treatment intake

For all treated subjects, first treatment date is the date recorded on the baseline CRF.

If first dose day is missing and first dose month and year are not missing:
Impute first dose date using the first day of the month. If this leads to a date before first drug dispense date, use the first drug dispense date instead. Imputation flag is 'D'.

If first dose month is missing and first dose year is not missing:
Impute first dose date using 1 January as the day and month. If this leads to a date before first drug dispense date, use the first drug dispense date instead. Imputation flag is 'M'.

If first dose year is missing:
Impute first dose date using the first drug dispensing date if the subject did not return all medication. Imputation flag is 'Y';

Last treatment intake date is collected on eCRF page. For treated subjects, if the date is incomplete,

If last dose day is missing, and last dose month and year are not missing:
Impute last dose date using the last day of the month. If this leads to a date after study completion/discontinuation date, use study completion/discontinuation date instead. Imputation flag is 'D'.

If last dose month is missing, and last dose year is not missing:
Impute last dose date using 31 December as the day and month. If this leads to a date after study completion/discontinuation date, use study completion/discontinuation date instead. Imputation flag is 'M'

If last dose year is missing:
Impute last dose date using study completion/discontinuation date. Imputation flag is 'Y'.

2.5.1.3 Adverse event information

Adverse event start date
AE start date will be used for AE classification and analysis. If AE start date is not complete, then the character variable will keep the original incomplete date, the numerical date variable will be imputed and an imputation flag will indicate which date component is missing.

If AE start day is missing, and AE start month and year are not missing:
If AE start year is the same as first dose year and the AE start month is the same as the first dose month) then impute AE start day using the day of first dose. If this leads to a date after the AE end date, use AE end date instead. Otherwise (AE start year is not the same as first dose year and/or AE start month is not the same as the first dose month) impute the AE start day using the first day of the

month. If this leads to a date before informed consent, the informed consent date will be used. Imputation flag is 'D'.

If AE start month is missing, and AE start year is not missing:
If AE start year is less than the first dose year, use the informed consent day and month. If AE start year is equal to the first dose year, use the first dose day and month. If this leads to a date after the AE end date, use AE end date instead. If AE start year is after the first dose year, use 1 January. Imputation flag is 'M'.

If AE start year is missing:
Impute AE start date using the day of first dose. If this leads to a date after the AE end date, use AE end date instead. Imputation flag is 'Y'.

Adverse event end date
The general recommendation is not to impute AE end date. However, since AE end date will be used for AE starting date imputation, In order to carry through the logic for programming, the following intermediate step will be used. Afterwards, the original character/numeric date recorded in CRF and supportive imputed end date will be kept in the final analysis dataset.

If AE end day is missing, and AE end month and year are not missing:
Impute AE end date using the last day of the month. If this leads to a date after end of follow-up date, use end of follow-up date instead.

If AE end month is missing, and AE end year is not missing:
Impute AE end date using 31 December as the day and month. If this leads to a date after end of follow-up date, use the end of follow-up date instead.

If AE end year is missing:
Impute AE end date using the end of follow-up date.

Adverse event relationship and severity
Missing AE relationship and severity will be imputed according to the rules below at the data level using "analysis" variables. The original variables will also be included in the analysis datasets.

Severity of adverse event

If the severity is missing for one of the treatment emergent occurrences of an AE (the same SOC/PT), the maximal severity on the remaining occurrences will be considered. If the severity is missing for all the occurrences a "missing" category will be added in summary table.

Relationship to investigational product

If the assessment of the relationship to investigational product is missing, then the relationship to investigational product has to be assumed as related and the AE considered as such in the derived analysis variable.

Adverse event Listings

For date variables, only character date variables will be presented, and no imputed dates will be printed in the listings. If the AE start date is imputed, then the onset relative day will be empty; if the AE start date or end date is imputed, then AE duration will be empty.

22.5.1.4 Prior and concomitant medication (other than MS treatment) date

Medication start and end date will be used for prior, concomitant and follow up medication classification and potential analysis for specific medications. In addition, the start and stop date for concomitant systemic corticosteroid will also be used to identify valid MRI scans (after a minimum of 30 days following the completion of a course of corticosteroid) for analysis, therefore different imputation rules will be used for end date of systemic corticosteroid to avoid excluding too many MRI scans (details can be found below). To keep the data integrity, if the date is not complete, then the character variable will keep the original incomplete date, the numerical date variable will be imputed (impute the end date first) and an imputation flag will indicate which date component is missing. Only the character variable will be used in data listing.

As prior and concomitant medications are flagged using the recorded start and end dates, it is not possible to apply different imputation rules based on the prior/concomitant status of a record. So the same algorithm will be used for all records made in the "prior/concomitant medication" section of the eCRF. The only distinction is made between systemic corticosteroid treatments and all other treatments. Systemic corticosteroid treatment is defined as ATC level 2 = "CORTICOSTEROIDS FOR SYSTEMIC USE".

Prior/Concomitant medication start date

The imputation rule for concomitant medication start date is the same as AE start date.

Prior/Concomitant medication end date

Imputation of non-systemic corticosteroid medication end date

<u>If end day is missing, and end month and year are not missing</u>:
Impute end date using the last day of the month. If this leads to a date after end of study follow up date, use end of follow up date instead. Imputation flag is 'D'.

<u>If end month is missing, and end year is not missing</u>:
Impute end date using 31 December as the day and month. If this leads to a date after end of study follow up date, use the end of follow up date instead. Imputation flag is 'M'.

<u>If end year is missing</u>:
Impute date using the end of follow up date. Imputation flag is 'Y'.

Systemic corticosteroid medication end date

According to the protocol, the preferred corticosteroids treatment in protocol is methylprednisolone sodium succinate 1 g, intravenously daily for 3-5 days. To avoid excluding too many MRI scans due

to the incomplete date, the end date of corticosteroid will be imputed by the date closest to 5 days after the start. Details are described as below:

If end day is missing and end month and year are not missing:
If the end month and year are the same as start date + 5, then impute the end date using start date +5; If the month and year are earlier than start date + 5, then impute the day using the last day of the month; If the month and year are later than start date + 5, then impute the day using the first day of the month. If this leads to a date after end of study follow up date, use the end of follow up date instead. Imputation flag is 'D'.

If end month is missing and year is not missing:
If the end year are the same as start date + 5, then impute the end date using start date +5; If the year are earlier than start date + 5, then impute the month and day using December 31; If the year are later than start date + 5, then impute the month and day using June 15. If this leads to a date after end of study follow up date, use the end of follow up date instead. Imputation flag is 'M'.

If end year is missing:
Impute the end date using start date +5. If this leads to a date after end of study follow up date, use the end of follow up date instead. Imputation flag is 'Y'.

A special note, for subjects with incomplete systemic corticosteroid start date, the rule of 5 days after the start would not apply for the initial imputation. Thus for these cases, after imputation of start date, the end date will be re-imputed by repeating the above rules.

2.5.2 Windows for Time Points

MRIs (including repeat imaging in case of QC rejections) are linked to specific visits, so it is not necessary to assign visits based on relative timing.
Other efficacy assessments will be summarized by the nominal visit as well and no visit window will be defined.

For safety assessments, the reference date for the derivation of relative days of events or findings will be the date of first drug intake as documented in the study medications administration page of the eCRF. All available values obtained between 2 visits including unscheduled measurements will be assigned to the appropriate visit window as specified in section 2.5.3. In the event of multiple measurements of the same test in the same window, the one closest to the targeted visit date will be used for the by-visit summary, in case of equal distance the later observation will be selected. For further details please see the section "Unscheduled Visits" as well.

2.5.3 Unscheduled Visits

Only scheduled visit efficacy measurements will be used for analyses visit to exclude the temporary fluctuations in the clinical status that may occur with a relapse.

Safety data from unscheduled visits will be used in in all safety analyses. The unscheduled visits will be assigned to the appropriate analysis time window. The one closest to the targeted visit date will be used in the presence of multiple measurements within the same time window. If two observations

within a window have the same target distance, the later one will be used for analysis. The following windows will be used:

Week of study	Target day	Minimal days	Maximum days
1 (day 1)	1	1	1
1 (day 2)	2	2	4
2	8	5	11
3	15	12	21
4	28	22	42
8	56	43	70
12	84	71	98
16	112	99	126
20	140	127	154
24	168	155	210
36	252	211	294
48	336	295	378
60	420	379	462
72	504	463	546
84	588	547	630
96	672	631	686
100	700	687	714
104	728	715	770
116	812	771	

2.5.4 Pooling of Centers for Statistical Analyses

Due the small sample size in some centers/countries, the centers will be pooled to regions for statistical analysis. Please refer to section 2.1.1 for the definition of regions.

3 Interim Analysis

There are no interim efficacy analyses planned for this study.

A preplanned interim analysis for sample size reassessment will be performed by BART.

4 Software Documentation

All summaries and statistical analyses will be generated using SAS® version 9.4 or higher (Cary, NC).

5 Revisions of the SAP

Version 3.0 (dated 04 September 2020) contains the following updates

Section	Modification	Rationale
Cover letter	Updated personnel	
Terms	Added abbreviations and definitions of terms	Terms added to SAP
1.1	Clarified follow-up period after week 96 or early termination	To correspond with protocol
1.2	Updated study objectives	To correspond with protocol
1.3	Additional sample size details provided	To correspond with the protocol
1.4	Schedule of Assessments refer to protocol	To correspond with protocol
2.1.1	Demography categories updated	To correspond to CRF
2.1.1	Disease characteristics variables updated	To use a similar derivation approach as for other duration variables
2.1.2	Treatment exposure and compliance updated	To correspond with protocol
2.1.3	Updated primary, secondary and tertiary efficacy variables	To correspond with protocol
2.1.3	Relapse re-consent criteria	Includes Treating Neurologist medically confirmed relapse
2.1.4.	ECG and physical examination added	To mention all relevant safety assessments
2.1.4.1	AESI and Suicidal ideation language updated	To mention all relevant safety assessments
2.1.5	PK variables added	
2.1.4.2	Lab tests updated	To correspond with parameters provided by the central lab
2.2	Definitions of analysis populations	To reflect the study protocol
	Procedures for protocol deviations	To reflect Protocol Deviation Plan
2.3	Clarified disposition of subjects	To correspond with protocol
2.4	Update various statistical methods including demographics and baseline characteristics, duration of oral study treatment exposure, primary efficacy, secondary endpoints, tertiary endpoints, analyses of safety data, analyses of pharmacokinetics, analyses of ADA, and analysis of pharmacodynamics	To correspond with protocol
2.5	Data handling conventions for general conventions, missing data, date of first and last treatment, adverse event information, windows for time points, unscheduled visits, pooling of centers	To be consistent with protocol

Appendix B	Re-consent can occur with Treating Neurologist medically confirmed relapse	To be consistent with protocol

6 References

1 Li D, Held U, Petkau J, Daumer M, Barkhof F, Fazekas F et al. MRI T2 lesion burden in multiple sclerosis: a plateauing relationship with clinical disability Neurology 2006; 66: 1384 – 1389.

2 Hochberg (1988) A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75(4):800-802

3 Jill Fisher, et al. Multiple sclerosis functional composite (MSFC) administration and scoring manual. Revised, October 2001.

4 The EuroQol Group. EuroQol- a new facility for the measurement of health-related quality oflife. Health Policy 1990; 16(3): 199-208.

5 Dolan P. Modeling valuations for EuroQol health states. Medical Care Vol 35, No. 11, pp.1095-1108. 1997

6. Liang, K-Y, and Zeger, S. L., Longitudinal data analysis using generalized linear models. Biometrika 1986; 73: 13-2

7. Hauser, SL. et al. NEJM. 2017 ; 366(3): 221-234

8. O'Connor P et al. NEJM. 2011; 365: 1293 – 1303

9. Vermersch et al. MSJ. 2014; 6(6): 705 – 716

Appendix A:

Independent Relapse Adjudication Panel (IRAP)

The Phase III study includes an objective definition of relapse and a standardized procedure for assessment of all suspected relapses as described in Section 4.4.1 of the protocol (TG1101-RMS302). As outlined in this section, the Treating Neurologist will, within these guidelines, exercise his or her best clinical judgement in treating the suspected relapse. However, neither the Treating Neurologist nor the blinded rater/Examining Neurologist will make the final determination of whether a suspected event is counted as a relapse for the purpose of calculating the primary study outcome of Annualized Relapse Rate (ARR). Instead, an independent, blinded IRAP will make all relapse determinations in the Phase III study. The purpose of the IRAP is to ensure that relapses are scored in a manner that is consistent and objective as possible, and to minimize the variability between sites.

The IRAP will review all subject records pertaining to suspected on-study relapses (i.e., new or worsening neurological symptoms of acute onset) that occur outside the scheduled study visits and are reported by the subject within 48 hours. Upon receiving the report by the subject, of a suspected relapse/acute neurological event, the Treating Neurologist and Examining Neurologists will in an independent manner, and blinded to the results of their separate examinations follow the procedures and timelines as described below and outlined in Appendix B (Appendix E of TG1101-RMS302). All suspected relapses initially reported by the subject are evaluated by the IRAP regardless of whether the Treating Neurologist considered the subject's symptoms to be the result of a relapse.

The IRAP comprises of 6 independent neurologists with expertise in MS clinical research who are not investigators in any other TG Therapeutics-sponsored studies with ublituximab. The IRAP members are trained on study procedures by the Sponsor and an independent Contract Research Organization (CRO) who manages all IRAP activities throughout the study. IRAP members are required to review trial data and perform adjudications in accordance with the protocol- specified definition (section 4.4.1 of the protocol (TG1101-RMS302).

The CRO transfers blinded data that are relevant to assessment of relapses (entered into the eCRF by both the Treating Neurologist and the Examining Neurologist) and prepares blinded case dossiers to be reviewed by the IRAP via a secure electronic portal. The dossiers include a standardized listing for each suspected relapse that specified onset date, type and severity of reported symptoms, whether the event had disqualifying characteristics (e.g., lasted ≥24 hours, or symptoms not MS-related), and whether event has been treated with methylprednisolone. The dossiers also include vital signs, AEs related to the potential relapse, physical examinations, and EDSS scores with all supporting data from the standardized Neurostatus exam; values from baseline and the most recent prior quarterly assessment are provided as context for values from the relapse evaluation visit. Medical/surgical history and clinical episode history are also provided. In order to minimize the potential for inclusion of data which can potentially unblind an individual case under review (e.g. inclusion of AE with details that could suggest a treatment assignment),

the CRO will perform additional programmatic checks on text fields pertaining to AEs and Physical Examination findings for data points that are part of the relapse case listing.

Each potential relapse is reviewed by 2 IRAP members who have worked and entered their evaluations independently, on cases randomly assigned to them by the CRO. The IRAP adjudicates each case based on all available data provided for that case and members are not permitted to contact the site or the sponsor for additional information. If 2 IRAP reviewers reached the same decision as to whether an episode constitutes a relapse, the decision is recorded by the CRO. If the 2 reviewers provide a conflicting assessment, a third IRAP member will review the case so that a majority vote can be obtained. The majority vote of up to 3 IRAP members serves as the final determination as to whether an event meets criteria for an on-study protocol defined relapse.

Relapse determinations by the IRAP are communicated to the Treating Neurologist, who will in turn, notify the subject and update the eCRF accordingly. The IRAP has no involvement in subsequent subject management decisions. In order to maintain independence and blinding, the Examining Neurologist will not receive this report. Furthermore, neither the Treating Neurologist nor the subject may disclose the IRAP decision to the Examining Neurologist. The Treating Neurologist, upon receiving the IRAP decision, will counsel the subject per Section 4.4.1 of the protocol (TG1101-RMS302). The subject will need to give her/his consent to continue the treatment phase of the trial.

For the Phase III study, the results of IRAP adjudications are transferred to the Sponsor only after the last subject completes the study. The Sponsor is blinded to the outcome of the IRAP adjudications while the study is ongoing.

Appendix B:

Evaluation Procedure for Subject Reported Relapses