Clinical Study Protocol

Title Page

Clinical Study Protocol Title:	A Phase III, Multicenter, Randomized, Parallel Group, Double Blind, Double Dummy, Active Controlled Study of Evobrutinib Compared with Teriflunomide, in Participants with Relapsing Multiple Sclerosis to Evaluate Efficacy and Safety.
Study Number:	MS200527_0082
Amendment Number	Not applicable
Merck Compound Number:	M2951
Study Phase:	III
Short Title:	Phase III Study of Evobrutinib in RMS
Acronym:	Evolution RMS 2
Coordinating Investigator:	Professor Xavier Montalban, MD, PhD Vall d'Hebron University Hospital and Research Institute, Barcelona, Spain.
Sponsor Name and Legal Registered Address:	For all countries, except the US and Canada: Merck Healthcare KGaA an affiliate of Merck KGaA, Darmstadt, Germany Frankfurter Str. 250 Darmstadt, Germany In the US and Canada: EMD Serono Research & Development Institute, Inc. an affiliate of Merck KGaA, Darmstadt, Germany 45A Middlesex Turnpike Billerica, MA, 01821, USA Medical Responsible: Name: Matthew Mandel, MD Address: Global Clinical Development Center – Neurology EMD Serono Research & Development Institute Inc. 45A Middlesex Turnpike, Billerica, MA 01821, USA Phone: +1 781 427 4250 Fax: Not Applicable E-mail: matthew.mandel@emdserono.com

Regulatory Agency Identifying Numbers:	EudraCT: 2019-004980-36 US IND: 129428
Protocol Version:	13 February 2020 / Version 1.0
Replaces Version:	Not applicable
Approval Date:	13 February 2020
Medical Monitor Name and Contact Information:	Kristen Graesser, MD Phone: +1 984 244 8871 E-mail: Kristen.Graesser@iqvia.com

Table of Contents

Table of Tables

Table of Figures

1 Protocol Summary

1.1 Synopsis

Protocol Title: A Phase III, Multicenter, Randomized, Parallel Group, Double Blind, Double Dummy, Active Controlled Study of Evobrutinib Compared with Teriflunomide, in Participants with Relapsing Multiple Sclerosis to Evaluate Efficacy and Safety.

Short Title: Phase III Study of Evobrutinib in RMS.

Rationale: The purpose of this study is to characterize the efficacy and safety of evobrutinib 45 mg administered orally twice daily versus teriflunomide (Aubagio®; 14 mg once a day orally) in participants with relapsing multiple sclerosis (RMS).

Objectives and Endpoints:

Objectives	Endpoints (Outcome Measures)	Further Estimand Attributes
Primary		
To demonstrate superior efficacy with evobrutinib compared to teriflunomide in terms of Annualized Relapse Rate (ARR)	ARR based on qualified relapses at Week 96 in participants with RMS	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death attributable to multiple sclerosis (MS) or treatment: Composite variable. Strategy for handling intercurrent event of death unattributable to MS or treatment: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: relapse rate ratio and confidence interval (CI) (negative binomial [NB] model), with test of treatment effect based on test of relapse rate ratio

Objectives	Endpoints (Outcome Measures)	Further Estimand Attributes
Secondary		
To demonstrate the efficacy of evobrutinib relative to that of teriflunomide on disability progression	• Time to first occurrence of 12-week confirmed disability progression (CDP) as measured by the Expanded Disability Status Scale (EDSS) over 96 weeks • Time to first occurrence of 24-week CDP as measured by the EDSS over 96 weeks	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death attributable to MS or treatment: Composite variable. Strategy for handling intercurrent event of death unattributable to MS or treatment: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: hazard ratio and CI (Cox model), with test of treatment effect based on logrank test
To demonstrate the efficacy of evobrutinib relative to that of teriflunomide on patient reported symptoms and functional status	• Change from Baseline (CFB) in Patient Reported Outcomes Measurement Information System (PROMIS) physical function (PF) score at 96 weeks • CFB in PROMIS Fatigue score at 96 weeks	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: difference of least-squares means of score CFB at 96 weeks and CI (Mixed effect model for repeated measures [MMRM] model), with test of treatment effect based on difference of least-squares means

Objectives	Endpoints (Outcome Measures)	Further Estimand Attributes
To demonstrate the efficacy of evobrutinib relative to that of teriflunomide on magnetic resonance imaging (MRI) lesion parameters	• Total number of T1 Gd+ lesions based on assessments at Week 24, Week 48, and Week 96 • Total number of new or enlarging T2 lesions based on assessments at Week 24, Week 48, and Week 96	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: lesion rate ratio and CI (NB model), with test of treatment effect based on test of lesion rate ratio
To characterize the safety and tolerability of evobrutinib	Safety as assessed by the nature, severity, and occurrence of adverse events (AEs) and adverse events of special interest (AESIs); vital signs; electrocardiograms (ECGs); absolute concentrations and change from Baseline in immunoglobulin (Ig) levels; and clinical laboratory safety parameters up to Week 108	Not applicable

Objectives	Endpoints (Outcome Measures)
OLE Period	
To evaluate the long-term safety, efficacy, and HRQoL of evobrutinib for an additional ~144 weeks	• Efficacy and HRQoL endpoints at Weeks 48, 96, and ~144 o ARR, based on protocol-defined qualified relapses o Change from Baseline in PROMIS PF score o Change from Baseline in PROMIS Fatigue score o Change from Baseline in Medical Outcomes Study 36 Item Short Form Health Survey (SF-36v2) • Efficacy and HRQoL endpoints over ~144 weeks o Time to first occurrence of 12-week confirmed EDSS progression over ~144 weeks o Time to first occurrence of 24-week confirmed EDSS progression over ~144 weeks o Time to first occurrence of 12-week confirmed PF deterioration compared to Baseline over ~144 weeks • Efficacy endpoints at Weeks 24, 48, 96, and ~144 o Total number of new or enlarging T2 lesions o Total number of T1 Gd+ lesions • Safety as assessed by the nature, severity, and occurrence of AEs and AESIs; vital signs; ECGs; absolute concentrations and change from Baseline in Ig levels; clinical laboratory safety parameters up to Week ~144

Overall Design: This is a Phase III, multicenter, randomized, parallel group, double blind, double dummy, active controlled study of evobrutinib with an active control teriflunomide, in participants with RMS.

Eligible participants will be randomized 1:1 to treatment with evobrutinib 45 mg twice daily, or teriflunomide 14 mg once a day (oral), stratified by region and Baseline Expanded Disability Status Scale (EDSS). Blinding will be accomplished using a double dummy design.

Number of Participants: The total sample size is planned to be 930 participants with a randomization ratio of 1:1 (approximately 465 participants per treatment group).

Study Intervention Groups and Duration: The 96-week Treatment Period will be preceded by a 4-week Screening Period (may be extended after discussion with the Medical Monitor but cannot exceed 8 weeks) and followed by a 4-week Safety Follow-up after treatment completion or early discontinuation.

Participants experiencing initial progression of disability between Week 72 and Week 96 will continue participating in the main study for up to 12 additional weeks. The continuation period will increase the duration of participation to a maximum of 108 weeks. Upon completion of the continuation period, participants will have the option of participating in the Open Label Extension (OLE) or ending treatment (and returning for a 4-week Safety Follow-up).

Participants who complete the 96-week double blind, double dummy Treatment Period will be offered participation in the OLE Period of the study, which is expected to last approximately 144 weeks, with a 4-week Safety Follow-up Visit.

Male participants will enter the OLE only after undergoing an accelerated elimination procedure and confirmation of teriflunomide plasma levels < 0.02 mg/L. The teriflunomide level will be reviewed centrally by an independent reviewer in order to maintain the blind of the study. For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE.

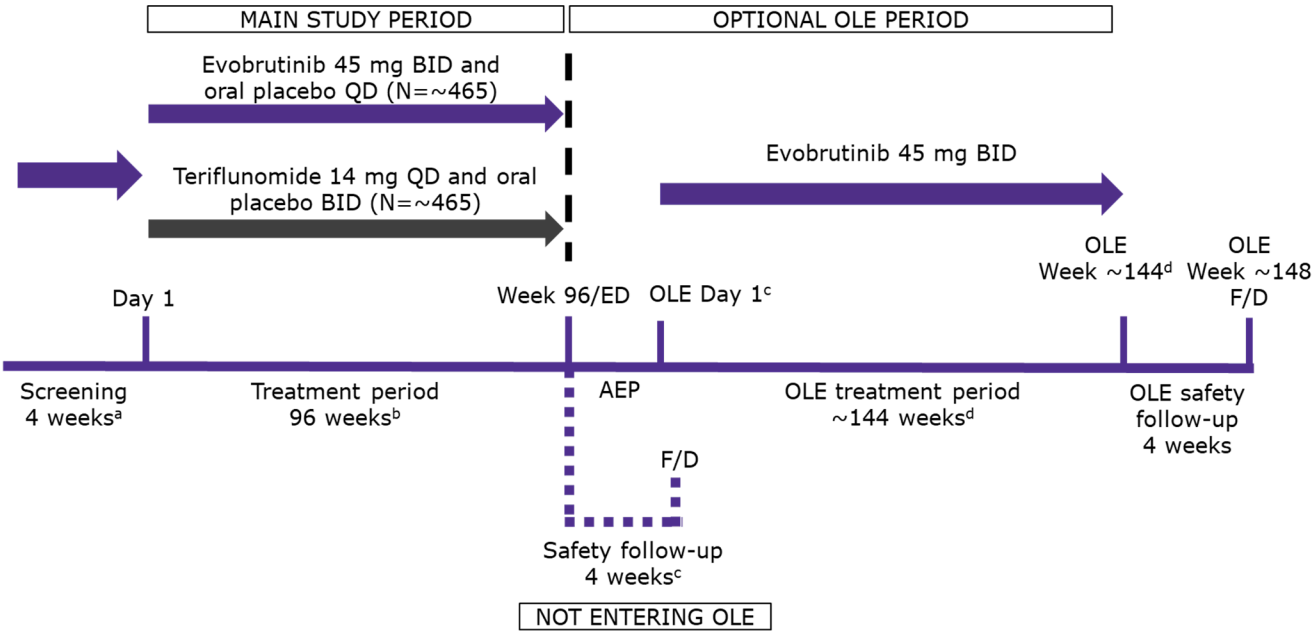
The Safety Follow-up Visit will be deferred until treatment is stopped in the OLE Period, due to either a participant's premature withdrawal/early termination from the OLE, termination of the study by the Sponsor, or completion of the OLE Period.

For participants experiencing a relapse within 4 weeks of eligibility for OLE, these participants will be allowed entry into the OLE after approval from the Sponsor provided that the treatment gap does not exceed 60 days from the last dose of study intervention received in the double blind, double dummy period (Week 96/up to Week 108 visit), and the start of the study intervention treatment in the OLE Period. For any male participants, the accelerated elimination procedure will be followed, as described above.

Involvement of Special Committee(s): Yes. Independent Data Monitoring Committee, Endpoint Adjudication Committee, and Study Steering Committee.

1.2 Schema

Figure 1 Study Schema



AEP = Accelerated Elimination Procedure, BID = Twice Daily, ED = Early Discontinuation, F/D = Follow-up/Discontinuation, OLE = Open Label Extension, QD = Once Daily.

a Screening may be extended to 8 weeks after approval by the Medical Monitor.

b Participants experiencing initial progression of disability between Week 72 and Week 96 will continue participating in the main study for up to 12 additional weeks. The continuation period will increase the duration of participation to a maximum of 108 weeks.

c Participants not entering the OLE will have a F/D visit 4 weeks after the last study intervention administration. All participants entering the OLE will receive evobrutinib after completing the 96-week (or up to 108-week) treatment period of the main study. Male participants will enter the OLE only after undergoing an AEP (described in Appendix 8) and confirmation of teriflunomide plasma levels < 0.02 mg/L. For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE.

d Anticipated duration of the OLE Period is 144 weeks.

1.3 Schedule of Activities

1.3.1 Schedule of Activities: Screening and Treatment Period (All Participants), End of Study (Participants Not Entering Open Label Extension Period)

Main Study		Intervention Period																											Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	Screening[b]	D1/Baseline	W2	W4	W6	W8	W10	W12	W14	W16	W18	W20	W22	W24	W28	W32	W36	W40	W44	W48	W60	W72	W84	W96/ED	W108[c]				
Visit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	-	26	a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Study Day	-28 to -1	1	15	29	43	57	71	85	99	113	127	141	155	169	197	225	253	281	309	337	421	505	589	673	757	-		Visit window is ± 3 days for all visits from W2 to F/D, inclusive.
Informed Consent	X																											
Inclusion and Exclusion Criteria	X	X																										Review Inclusion and Exclusion Criteria before randomization and first dose of study intervention.
Demography	X																											

Main Study		Intervention Period																									Notes
Assessments & Procedures[a]	**Screening**[b]	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108**[c]	**Unscheduled Visit**[d]	**F/D**[e]	a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Full Physical Examination	X	X																		X				X	X	X	X	Additional examinations may be completed at the discretion of the Investigator.
Medical History (includes substance usage)	X																											Substances: drugs, alcohol, tobacco, and caffeine.
MS History	X																											
Serum Pregnancy Test (WOCBP only)	X																											
Highly sensitive urine pregnancy test		X		X		X		X		X		X		X	X	X	X	X	X	X	X	X	X	X	X		X	Monthly urine pregnancy tests to be performed for all WOCBP. In-between visits, urine pregnancy testing will be performed at home.

Main Study		Intervention Period																									Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**			a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Telephone Visit				[----------X----------]																								Telephone visit every 4 weeks from W4 to W96 (if clinic visit not scheduled) to assess for relapses, confirm completion of home pregnancy testing and discuss results, and review concomitant medications, procedures and AE/SAE in between site visits (see Section 8). Participants experiencing new or worsening neurological symptoms, including possible relapse, should be evaluated by the Investigator, if necessary, at an Unscheduled Visit (see Section 8).
QuantiFERON®-TB tuberculosis test	X																			X[f]				X[f]				f Participants in high TB burden settings must have repeated QuantiFERON testing at least annually at the indicated visits, using the assay that was negative at Screening (see Exclusion Criteria 7 and 8). Participants can be tested for TB at any time during the study at the discretion of the Investigator.

Main Study		Intervention Period																									Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**			a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Ferritin, and transferrin saturation	X																											See Exclusion Criterion 10.
HIV, HBV and HCV testing	X																											HIV testing will be conducted centrally. Where required by local regulations, testing can be conducted and analyzed locally. HBV and HCV testing will be performed at the central laboratory. Participants positive for anti-HCV antibodies will have reflex testing performed for HCV RNA by PCR. See Exclusion Criterion 34.
Efficacy																												
Neurological Evaluation EDSS, T25-FW, 9-HPT, and SDMT	X	X						X						X			X			X	X	X	X	X	X	X	X	Completed on tablet by the Examining Investigator (assessor) (or Qualified Examining Designee for T25-FW, 9-HPT, and SDMT).
Relapse assessment								X						X			X			X	X	X	X	X	X	X	X	

Main Study		Intervention Period																								Notes
Assessments & Procedures[a]	**Screening**[b]	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108**[c]	**Unscheduled Visit**[d]	**F/D**[e]	a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
MRI scan	X													X						X				X				Participants should meet all nonimaging inclusion/exclusion criteria before the Screening/Baseline MRI is performed. An MRI may be obtained at the Discontinuation Visit if the previous MRI was performed more than 4 weeks prior to the visit.
PRO assessments and health resource utilization																												
PROMIS Fatigue and Physical Functioning, EQ-5D-5L		X						X						X			X			X	X	X	X	X	X		X	PROs should be completed prior to administration of study intervention and prior to any other study assessment(s). The tablet versions will be completed.
SF-36v2		X												X						X		X		X			X	
Health resource utilization		X						X						X			X			X	X	X	X	X	X		X	See Section 8.9.

Main Study		Intervention Period																									Notes
Assessments & Procedures[a]	**Screening**[b]	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108**[c]	**Unscheduled Visit**[d]	**F/D**[e]	a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Safety assessments																											
C-SSRS	X	X						X						X			X			X	X	X	X	X	X		X	The Treating Investigator will complete the tablet version at each assessment.
Reflex testing for HBV DNA	X			X		X		X						X			X			X	X	X	X	X	X		X	For participants who are anti-hepatitis B surface antibody positive and/or anti-hepatitis B core antibody positive at Screening, reflex testing for hepatitis B virus DNA (HBV DNA) by PCR will be performed. See Exclusion Criterion 34.
Evobrutinib concentration assessment			[----------X----------]																									Collect all samples as specified, however analysis will only be conducted if elevated LFT(s) are observed.
Biochemistry and Hematology	X	X						X						X						X		X		X	X	X	X	See Section 7.1 and Appendix 5.
Supplemental LFT: ALP, AST, ALT, GGT, and Total Bilirubin			X	X	X	X	X		X	X	X	X	X		X	X	X	X	X		X		X					If elevated, LFTs will be repeated along with additional testing (see Section 7.1).

Main Study		Intervention Period																									Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**			a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Hepatic/ Autoimmune Panel		Analysis will only be conducted if elevated LFT(s) are observed.																										See Section 7.1 and Appendix 5.
Coagulation	X																			X				X	X	X		
Urinalysis/ microscopy and urine chemistry	X	X												X						X		X		X	X	X	X	See Section 8 and Appendix 5.
12-lead ECG	X	X		X										X										X				A Baseline ECG will be collected on Day 1 prior to dosing. All other ECG assessments must be conducted approximately 45 to 60 minutes after study intervention administration at the site. See PK sampling row for timing of PK collection. Additional ECGs can be done if there are concerns about cardiac signs or symptoms.
Vital Signs, Height, and Weight	X	X						X						X			X			X	X	X	X	X	X	X	X	Collect Height at Screening only.

Main Study		Intervention Period																									Unscheduled Visit[d]		Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**		**F/D[e]**	a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
AE, SAE & AESI Review	[----------X----------]																											AE review to be started after ICF is signed. When there is no clinic visit, AE/SAE review will be conducted as part of the telephone visits.
Concomitant Medication and Procedures Review	[----------X----------]																											When there is no clinic visit, concomitant medication and procedure review will be conducted as part of the telephone visits.
Blinded Immuno-globulin levels		X						X						X						X		X		X			X	
Study Intervention																												
Randomization		X																										Randomization occurs on Day 1, prior to start of study intervention and after participant has met all eligibility criteria and all assessments have been completed.

Main Study		Intervention Period																									Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**			a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Dispense Study Intervention(s)		X						X						X			X			X	X	X	X	X				Dispense as needed per IWRS. Study intervention dispensed at W96 for participants experiencing clinical progression and who are assigned to an additional 12 weeks of study intervention.
Study Intervention(s) Administration		Administration of study interventions																							X[b]			Study intervention must be taken in the fed state, see Section 6.1. On study visit days, first daily dose of study intervention should be administered during the study visit; otherwise, study intervention should be self-administered at home at a set time twice a day.
Study Intervention(s) Compliance		[----------X----------]																										Participant diary to be completed after every study intervention administration. See Section 6.4.

Main Study		Intervention Period																									Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**			a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
PK, PGx, and Biomarker assessments																												
PK sampling		X		X				X						X						X		X		X				Study visits containing PK sampling should be scheduled before first daily dose, and participants should be instructed to wait and take study interventions at the site (see Section 6.1). • Day 1 PK collected at 0 (predose), 0.5, 1, 2, and 4 hours postdose. Prior to the predose blood draw, the baseline ECG assessment should be taken. The postdose ECG assessment on Day 1 should be near the C_{max}, or approximately 45-60 minutes after dosing, with PK sampling within 30 minutes AFTER the ECG assessment.

Main Study		Intervention Period																								Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**			a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
																												• W4, W24, and W96 PK collected at 0 hours (predose) and 1 sample within 30 minutes AFTER the ECG assessment. ECG assessment should be performed approximately 45-60 minutes postdose. • W12, W48, and W72 PK collected at 0 hours (predose) and 1 sample between 1 and 6 hours postdose.
Gene expression		X						X						X						X				X				To be collected prior to first daily dose.
Pharmaco-genetics		X																										Sample to be collected on Day 1 before study intervention starts (optional for participants who sign Pharmacogenetics ICF). If not collected on Day 1 or a redraw is needed, the sample may be obtained at any other point of time during the study.

Main Study		Intervention Period																									Unscheduled Visit[d]	F/D[e]	Notes
Assessments & Procedures[a]	**Screening[b]**	**D1/Baseline**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96/ED**	**W108[c]**			a Home visits (if available) may be provided as an option for some of the scheduled visits. b Screening period is 4 weeks and can be extended to 8 weeks after approval by the Medical Monitor. c The (up to) W108 visit will apply to a subset of participants who have disability progression between 72 and 96 weeks. d For unscheduled visits, assessments not marked below may be performed at the discretion of the Investigator. e F/D visit will be performed 28 ± 3 days after the last study intervention administration.
Biomarkers of disease		X						X						X						X		X		X				To be collected prior to first daily dose.
Novel liver function protein biomarkers		[----------X----------]																										Collect all samples as specified prior to first daily dose, however analysis will only be conducted if elevated LFT(s) are observed.
Novel liver function genomic biomarkers		[----------X----------]																										Collect all samples as specified prior to first daily dose, however analysis will only be conducted if elevated LFT(s) are observed.

9-HPT = 9-Hole Peg Test, AESI = Adverse Event of Special Interest, ALP = Alkaline phosphatase, ALT = Alanine Aminotransferase, AST = Aspartate Aminotransferase, CRF = Case Report Form, C-SSRS = Columbia-Suicide Severity Rating Scale, D = Day, DNA = Deoxyribonucleic Acid, ECG = Electrocardiogram, ED = Early Discontinuation, EDSS = Expanded Disability Status Scale, EQ-5D-5L = EuroQoL 5 Dimension 5 Level, F/D = Follow-up/Discontinuation, GGT = γ-Glutamyl-Transferase, HCV = Hepatitis C Virus, HBV = Hepatitis B Virus, HIV = Human Immunodeficiency Virus, ICF = Informed Consent Form, IWRS = Interactive Web Response System, LFT = Liver Function Test, MRI = Magnetic Resonance Imaging, MS = multiple sclerosis, PCR = polymerase chain reaction, PGx = Pharmacogenetics, PK = Pharmacokinetic, PRO = Patient Reported Outcomes, PROMIS = Patient Reported Outcomes Measurement Information System, RNA = Ribonucleic acid, SAE = Serious Adverse Event, SDMT = Symbol Digit Modalities Test, SF 36v2 = 36-Item Short Form Survey Instrument Version 2, TB = Tuberculosis, T25-FW = Timed 25-Foot Walk, W = Week, WOCBP = Women Of Childbearing Potential.

1.3.2 Schedule of Activities – Optional Open Label Extension Period

Open Label Extension[a]		**Intervention Period**																											**Unscheduled Visit**	**W148F/D[e]**	**Notes**
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**			a OLE Period is expected to last ~144 weeks. b Home visits (if available) may be provided as an option for some of the scheduled visits. c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8). d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1). e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.
Visit		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	-	27	
Study Day		1	15	29	43	57	71	85	99	113	127	141	155	169	197	225	253	281	309	337	421	505	589	673	757	841	1009	-	1037	Visit windows are: • ± 3 days from Day 1 to W44 • ± 7 days from W48 to ~W144/ED
AEP[b]	X																													
Informed Consent	X[f]	X																												f Male participants that undergo AEP have to consent for the OLE period prior to starting the AEP.
Inclusion and Exclusion Criteria		X																												Review Inclusion and Exclusion Criteria before first dose of study intervention.
Blinded teriflunomide levels		X																												For male participants only.
Full Physical Examination		X																						X			X	X	X	Additional examinations may be completed at the discretion of the Investigator.

Open Label Extension[a]		Intervention Period																										Unscheduled Visit	W148F/D[e]	Notes
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**			a OLE Period is expected to last ~144 weeks. b Home visits (if available) may be provided as an option for some of the scheduled visits. c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8). d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1). e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.
Medical History (includes substance usage)		X																												Substances: drugs, alcohol, tobacco, and caffeine.
Serum Pregnancy Test (WOCBP only)		X																												
Highly sensitive urine pregnancy test		X		X		X		X		X		X		[----------X----------]														X	Monthly urine pregnancy tests to be performed for all WOCBP. In-between visits, urine pregnancy testing will be performed at home.

Open Label Extension[a]		Intervention Period																											Unscheduled Visit	W148F/D[e]	Notes
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**				
Telephone Visit				[----------X----------]																											Telephone visit every 4 weeks from W4 to W144 (if clinic visit not scheduled) to assess for relapses, confirm completion of home pregnancy testing and discuss results, and review concomitant medications, procedures and AE/SAE in between site visits (see Section 8). Participants experiencing new or worsening neurological symptoms, including possible relapse, should be evaluated by the Investigator, if necessary, at an Unscheduled Visit (see Section 8).
QuantiFERON-TB tuberculosis test																				X				X	X		X			Participants in high TB burden settings must have repeated QuantiFERON testing at least annually at the indicated visits, using the assay that was negative at Screening (see Exclusion Criteria 7 and 8). Participants can be tested for TB at any time during the study at the discretion of the Investigator.

a OLE Period is expected to last ~144 weeks.

b Home visits (if available) may be provided as an option for some of the scheduled visits.

c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8).

d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1).

e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.

Open Label Extension[a]		Intervention Period																										Unscheduled Visit	W148F/D[e]	Notes
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**			a OLE Period is expected to last ~144 weeks. b Home visits (if available) may be provided as an option for some of the scheduled visits. c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8). d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1). e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.
Efficacy																														
Neurological Evaluation EDSS, T25-FW, 9-HPT, and SDMT		X						X						X			X			X		X		X	X	X	X	X	X	Completed on tablet by the Examining Investigator (assessor) (or Qualified Examining Designee for T25-FW, 9-HPT, and SDMT).
Relapse assessment								X						X			X			X		X		X	X	X	X	X	X	
MRI scan		X												X						X				X			X			Completed at Discontinuation. Participants should meet inclusion/ exclusion criteria before MRI is performed.
PRO assessments and health resource utilization																														
PROMIS Fatigue and Physical Functioning, EQ-5D-5L		X						X						X			X			X		X		X	X	X	X		X	PROs should be completed prior to administration of study intervention and prior to any other study assessment(s). The tablet versions will be completed.
SF-36v2		X												X						X		X		X		X	X		X	

Open Label Extension[a]		Intervention Period																										Unscheduled Visit	W148F/D[e]	Notes
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**			a OLE Period is expected to last ~144 weeks. b Home visits (if available) may be provided as an option for some of the scheduled visits. c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8). d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1). e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.
Health resource utilization		X						X						X			X			X		X		X		X	X		X	See Section 8.9.
Safety assessments																														
C-SSRS		X						X						X			X			X		X		X		X	X		X	The tablet version will be completed by the Treating Investigator at each assessment.
Reflex testing for HBV DNA		X		X		X		X						X		X				X	X	X	X	X	X	X	X			For participants who are anti-hepatitis B surface antibody positive and/or anti-hepatitis B core antibody positive at Screening, reflex testing for hepatitis B virus DNA (HBV DNA) by PCR will be performed. Reflex testing will be serially tested if HBV DNA is within the ranges defined in Exclusion Criterion 34.
Evobrutinib concentration assessment			[----------X----------]																			X		X		X	X			Collect all samples as specified, however analysis will only be conducted if elevated LFT(s) are observed.
Biochemistry and Hematology		X						X						X						X		X		X		X	X	X	X	See Appendix 5.

Open Label Extension[a]		Intervention Period																										Unscheduled Visit	W148F/D[e]	Notes
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**			a OLE Period is expected to last ~144 weeks. b Home visits (if available) may be provided as an option for some of the scheduled visits. c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8). d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1). e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.
Supplemental LFT: ALP, AST, ALT, GGT, and Total Bilirubin			X	X	X	X	X		X	X	X	X	X		X	X	X	X	X		X		X		X					If elevated, LFTs will be repeated along with additional testing (see Section 7.1).
Hepatic/ Autoimmune panel		Analysis will only be conducted if elevated LFT(s) are observed.																												See Section 7.1 and Appendix 5.
Coagulation		X																		X				X				X		
Urinalysis/ microscopy and urine chemistry		X												X						X		X		X		X	X	X	X	See Section 8 and Appendix 5.
Vital Signs and Weight		X						X						X			X			X		X		X		X	X	X	X	
AE, SAE & AESI Review		[----------X----------]																												AE review to be started after ICF is signed. When there is no clinic visit, AE/SAE review will be conducted as part of the telephone visits.

Open Label Extension[a]		**Intervention Period**																											**Unscheduled Visit**		**Notes**
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**		**W148F/D[e]**	a OLE Period is expected to last ~144 weeks. b Home visits (if available) may be provided as an option for some of the scheduled visits. c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8). d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1). e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.
Concomitant Medication and Procedures Review		[----------X----------]																												When there is no clinic visit, concomitant medication and procedure review will be conducted as part of the telephone visits.
Blinded Immuno-globulin levels		X						X						X						X		X		X		X	X		X	
Study Intervention																														
Dispense Study Intervention(s)		X						X						X			X			X		X				X				Dispense as needed per IWRS.
Study Intervention(s) Administration		X	Administration of study intervention																											Evobrutinib dosing will start the day after OLE Day 1. Evobrutinib must be taken in the fed state, see Section 6.1. On study visit days, first daily dose of study intervention should be administered during the study visit; otherwise, study intervention should be self-administered at home at a set time twice a day.

Open Label Extension[a]		**Intervention Period**																										**Unscheduled Visit**	**W148F/D[e]**	**Notes**
Assessments & Procedures[b]	**AEP Period[c]**	**D1/Baseline[d]**	**W2**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W18**	**W20**	**W22**	**W24**	**W28**	**W32**	**W36**	**W40**	**W44**	**W48**	**W60**	**W72**	**W84**	**W96**	**W108**	**W120**	**W144[a]/ED**			a OLE Period is expected to last ~144 weeks. b Home visits (if available) may be provided as an option for some of the scheduled visits. c Male participants will enter the OLE only after undergoing an AEP (see Appendix 8). d For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE (see Section 4.1). e OLE F/D visit will be performed 28 ± 3 days after last study intervention administration.
Study Intervention(s) Compliance		X	[----------X----------]																											Participant diary to be completed after every study intervention administration. See Section 6.4.
Biomarker assessments																														
Gene expression		X						X						X						X				X						To be collected prior to first daily dose.
Biomarkers of disease		X						X						X						X		X		X		X				To be collected prior to first daily dose.
Novel liver function protein biomarkers		[----------X----------]																				X		X		X				Collect all samples as specified prior to first daily dose, however analysis will only be conducted if elevated LFT(s) are observed.
Novel liver function genomic biomarkers		[----------X----------]																				X		X		X				Collect all samples as specified prior to first daily dose, however analysis will only be conducted if elevated LFT(s) are observed.

9-HPT = 9-Hole Peg Test, AE = Adverse Event, AEP = accelerated elimination procedure, AESI = Adverse Event of Special Interest, ALP = Alkaline phosphatase, ALT = Alanine Aminotransferase, AST = Aspartate Aminotransferase, CRF = Case Report Form, C-SSRS = Columbia-Suicide Severity Rating Scale, D = day, DNA = Deoxyribonucleic Acid, ED = Early Discontinuation, EDSS = Expanded Disability Status Scale, EQ-5D-5L = EuroQoL 5 Dimension 5 Level, F/D = Follow-up/Discontinuation, GGT = γ-Glutamyl-Transferase, HBV = Hepatitis B Virus, ICF = Informed Consent Form, IWRS = Interactive Web Response System, LFT = Liver Function Test, MRI = Magnetic Resonance Imaging, MS = multiple sclerosis, OLE = Open Label Extension, PCR = Polymerase chain reaction, PRO = patient reported outcomes, PROMIS = Patient Reported Outcomes Measurement Information System, SAE = Serious Adverse Event, SDMT = Symbol Digit Modalities Test, SF 36v2 = 36-Item Short Form Survey Instrument Version 2, TB = Tuberculosis, T25-FW = Timed 25-Foot Walk, W = Week, WOCBP = Women Of Childbearing Potential.

2 Introduction

Evobrutinib is a potent, orally administered, highly selective, irreversible inhibitor of Bruton's Tyrosine Kinase (BTK) that is being developed for the treatment of relapsing multiple sclerosis (RMS).

Complete information on the chemistry, pharmacology, efficacy, and safety of evobrutinib is in the current Investigator's Brochure.

2.1 Study Rationale

The purpose of this study is to characterize the efficacy and safety of evobrutinib 45 mg administered orally twice daily versus teriflunomide (Aubagio®; 14 mg once a day orally) in participants with RMS.

Evobrutinib inhibits activation of B cells via the B cell receptor. In addition, evobrutinib inhibits activation of myeloid cells by immune complexes via Fc receptors as well as the differentiation of proinflammatory macrophages. Thus, evobrutinib may be suitable for the treatment of multiple sclerosis (MS).

The main randomized, double blind, placebo-controlled Phase II study of evobrutinib with a parallel, open label, active control group (dimethyl fumarate [Tecfidera®]), in patients with RMS to evaluate efficacy, safety, tolerability, pharmacokinetics (PK), and biological activity (MS200527_0086) is completed; the Open Label Extension (OLE) part of the study is ongoing. The study consisted of first a 24-week period, in which 267 participants were randomized to evobrutinib 25 mg once daily (n = 52), 75 mg once daily (n = 53), 75 mg twice daily (n = 54), dimethyl fumarate 240 mg twice daily (provided open label) (n = 54), or placebo (n = 54). Data from the first 24-weeks were analyzed for the primary analysis (PA). Following the 24-week period, all remaining participants continued to be treated for an additional 24-week period as follows: participants who received previously evobrutinib 25 mg once daily, 75 mg once daily, 75 mg twice daily, or dimethyl fumarate, continued to be treated with their original treatment, and participants who received placebo for the first 24-weeks were switched to evobrutinib 25 mg once daily. Treatment assignment remained blinded with the exception of dimethyl fumarate which was open label throughout the study. The first 48 weeks analysis is referred to as the Blinded Extension Analysis (BEA). For the PA, efficacy analysis consisted of the comparison between the placebo and evobrutinib treatment groups. No formal comparison between the dimethyl fumarate and the evobrutinib treatment groups was performed for either the PA or the BEA.

In the PA, the primary efficacy endpoint of the Phase II study was the total number of gadolinium-enhancing T1 lesions at Weeks 12, 16, 20, and 24. The primary efficacy analysis was a comparison of each evobrutinib dose arm versus placebo based on lesion rate (lesions per scan) ratio, adjusted for Baseline lesion activity. The primary endpoint of the study was met, and both evobrutinib 75 mg once daily (lesion rate ratio 0.30, 95% confidence interval [CI]: 0.14, 0.63; p = 0.0015 [unadjusted], p = 0.046 [adjusted according to Hochberg]) and evobrutinib 75 mg twice daily (lesion rate ratio 0.44, 95% CI: 0.21, 0.93; p = 0.0313 [unadjusted], p = 0.0648 [adjusted

according to Hochberg]) were associated with a reduction in T1 Gd+ lesion rate compared to placebo (Montalban 2019).

The first key secondary endpoint was annualized relapse rate (ARR) at 24-weeks. A trend towards a reduction in ARR (unadjusted [95% CI]) was seen with evobrutinib 75 mg once daily (0.13 [0.03, 0.38]; p = 0.09) and evobrutinib 75 mg twice daily (0.08 [0.01, 0.30]; p = 0.06) versus placebo (0.37 [0.17, 0.70]), with evidence of a dose response (p = 0.014).

In the BEA, results showed an unadjusted ARR at 48 weeks of 0.11 (95% CI: 0.04, 0.25) for evobrutinib 75 mg twice daily and 0.25 (95% CI: 0.12, 0.44) for evobrutinib 75 mg once daily indicating that the efficacy trend observed during the first 24-week period was sustained during the second 24-week period, with greater clinical sustained efficacy in the 75 mg twice daily group. The results of the Phase II OLE (cutoff 26 September 2019) indicate that the participants from the evobrutinib 75 mg twice daily arm had an ARR of 0.11, consistent with what was observed at the end of the main study (Week 48, cutoff 13 July 2018).

Evobrutinib at 45 mg twice daily in the fed state is expected to be similar to the tested 75 mg twice daily dose in the fasted state based on the available exposure/response data (see Section 4.3).

These results provide evidence of evobrutinib's efficacy compared to placebo. Clinical development efforts in the MS indication will focus specifically on RMS, with potential expansion into progressive MS (and secondary progressive multiple sclerosis [SPMS]) with subclinical central nervous system (CNS) inflammation.

Reversible elevations in transaminases (alanine aminotransferase [ALT] and aspartate aminotransferase [AST]) have been observed in participants randomized to evobrutinib in the RMS Phase II study (MS200527_0086). This finding, and the overall safety and risk-benefit assessment of evobrutinib are discussed further in Section 2.3 and Section 4.3.

Taken together, efficacy and safety data from this study support progression into Phase III.

Refer to the Investigator's Brochure for further information about the nonclinical and clinical programs and Guidance for the Investigator.

2.2 Background

Currently there is no cure for MS, but the course of the disease can be altered favorably with disease-modifying drugs (DMDs) with varying levels of efficacy, and distinct safety and tolerability profiles. Most active RMS patients initiate treatment with an interferon-beta or glatiramer acetate therapy. Oral DMDs including dimethyl fumarate and teriflunomide are also used as first-line agents. If responding suboptimally, patients can be treated with an alternative, second-line oral therapy such as cladribine and fingolimod, or infusion agents such as natalizumab and ocrelizumab. Generally, DMDs perceived to be more efficacious have also been shown to be associated with more significant adverse effects, ranging from serious infections (i.e., progressive multifocal leukoencephalopathy [PML]) to autoimmunity and cancer. Switching among these DMDs occurs primarily due to perceived lack of efficacy or the occurrence of adverse events (AEs), as well as individual patient preferences.

Despite the recent approvals of newer therapies for the treatment of MS, there remains an unmet need for highly effective and well-tolerated therapies for patients with MS at all stages of the disease. Early treatment with a highly efficacious and safer DMD could be advantageous for long term quality of life for MS patients and might slow the process of brain atrophy, which accompanies axonal damage and loss of gray and white matter.

Evobrutinib is a highly specific, oral inhibitor of BTK that inhibits B cell activation and B cell/T cell interaction, decreasing plasma cell formation and autoantibody production (Investigator Brochure, Haselmayer 2017). In addition, evobrutinib was shown to inhibit M1 macrophage survival and proinflammatory cytokine release and promotes M2 polarization of reparative human monocytes in vitro (Alankus 2018). In line with the in vitro data, evobrutinib demonstrated pharmacological efficacy in both B cell and T cell dependent mouse models of MS, by reducing CNS inflammation and amelioration of disease severity (Boschert 2017, Torke 2018). Since B cell depletion studies have shown that antibody independent B cell functions play an important role in MS pathogenesis (Bar-Or 2010, Fraussen 2016, Jelcic 2018) and an altered innate immune system contributes to disability progression and repair in MS (Vogel 2013, Rawji 2016), evobrutinib may offer advantages over current approved DMDs.

Clinical efficacy was recently demonstrated with B cell depleting antiCD20 therapies in Phase II and Phase III clinical studies in RMS and progressive MS (Hauser 2008, Hawker 2009, Montalban 2016, Wolinsky 2016). Ocrelizumab (Ocrevus®) inhibited the formation of new inflammatory magnetic resonance imaging (MRI) lesions up to 90% (Hauser 2008) in Phase II RMS studies and high efficacy on MRI (-94%), ARR (-46%), and 24-week disease progression (-40%) was also reached in OPERA Phase I, II, and III studies against interferon-beta. Translational mechanism of action studies in antiCD20 treated RMS patients show diminished proliferation and proinflammatory differentiation of T cells (Bar-Or 2010), pointing towards abrogation of antigen presenting cell function as the primary mechanism. In addition to the role of antiCD20 in B cell antigen presentation, a recent publication of Li et al (Li 2015) describes a diminished proinflammatory myeloid cell response in ocrelizumab treated MS participants. Evobrutinib shows inhibition of myeloid cell activation by immune complexes.

Preclinical proof of concept with evobrutinib has been demonstrated for systemic lupus erythematosus (SLE)/lupus nephritis, experimental autoimmune encephalomyelitis, rheumatoid arthritis (RA) and passive cutaneous anaphylaxis. Oral evobrutinib does not deplete B cells in the studies carried out to date and, upon withdrawal, restoration of immune function can be achieved in days. This is less than treatment with antiCD20 therapies, where restoration of the immune system can take months, and is important should the need to interrupt or stop therapy arise. This suggests that a more favorable benefit to risk balance with respect to infections for evobrutinib versus antiCD20 therapies may be observed. In addition, BTK inhibitors might have broader efficacy than agents that cause B cell depletion, due to the importance of BTK activation downstream of various receptors expressed in myeloid cells, suggesting an additional direct effect of evobrutinib on innate immune cell activation induced by immune complexes, cytokines/chemokines, or toll-like receptor (TLR) activation (Block 2012, López-Herrera 2014, Whang 2014). A direct myeloid inhibition activity also best explains the significant reduction of clinical score, relapse rate, and time to first relapse in T cell dependent experimental autoimmune encephalomyelitis models, in which antiCD20 antibodies do not work.

2.3 Benefit/Risk Assessment

In the PA of the MS200527_0086 study, evobrutinib significantly decreased the number of Gd+ T1 lesions at Weeks 12, 16, 20, and 24 in participants with RMS. Similar to the relationship seen for the MRI endpoints, evobrutinib 75 mg once and twice daily was associated with a lower ARR compared to placebo. In the OLE, the ARR remained low and similar to those observed in the PA and BEA. These results show evobrutinib is efficacious in participants with RMS, and consequently warrant further investigation in Phase III clinical studies.

As of July 2019, approximately 1,214 adult participants in completed and ongoing clinical studies have been exposed to evobrutinib including healthy volunteers, participants with RMS, SLE, or RA, and participants with renal impairment. Evobrutinib was generally well tolerated in all participants. The treatment-related treatment-emergent adverse events (TEAEs) have been primarily mild to moderate in severity.

A confirmed safety finding of increases in liver transaminases has been observed from Study MS200527_0086 in participants with RMS and the ongoing Study MS200527_0018 in participants with SLE and was considered as an important identified risk for evobrutinib. Mild (Grade 1) elevations of liver transaminases were frequent, asymptomatic and reversible; however, more severe events were also rarely reported.

During the first 52 weeks of the MS200527_0086 study, all of the instances of elevated liver enzymes occurred during the first 6 months of exposure to evobrutinib and none of the cases of elevated transaminases had any clinical signs or symptoms, and the transaminase elevations have resolved over time with the withdrawal of evobrutinib. A single participant in the open-label extension phase of the MS200527_0086 study developed contemporaneous elevations of ALT, AST, and bilirubin (all ≥ Grade 3) while on evobrutinib 75 mg once daily. This participant subsequently was suspected to have hemochromatosis, based on a combination of elevated ferritin, elevated transferrin saturation, and homozygosity for a known disease-causing mutation in the human hemochromatosis protein gene.

Given these observations, this protocol excludes participants with hemochromatosis, Wilson's disease, alpha-1-antitrypsin deficiency, or any other chronic liver disease including Gilbert's disease and > 2 × upper limit of normal (ULN) values for ALT, AST, amylase, or lipase elevation at Screening. During the study, liver transaminase will be monitored every 2 weeks, particularly during the first 6 months of exposure to evobrutinib. The protocol has strict stopping criteria for liver transaminase elevations leading to treatment discontinuation and timely medical management. Also, data will be reviewed by an Independent Data Monitoring Committee (IDMC) with appropriate expertise. Beyond this, risk minimization measures proposed are considered standard for this phase of clinical development.

Investigations on embryo-fetal development in toxicological studies showed an increased incidence of malformations (mainly cleft palate) and skeletal variations in mice when compared to the control group, and abortions and/or vaginal bleeding during the last period of gestation in rabbits. In addition, an increase of resorptions, and a lower mean fetal weight were also seen.

Based on these, embryo-fetal toxicity is considered as an important potential risk in participants exposed to evobrutinib. Therefore, female participants of childbearing potential must not be pregnant, must have a negative pregnancy test at the time of enrollment and use highly effective contraception (as specified in the clinical study protocol) during the study period and after the last dose, and should not donate eggs, as risk mitigation measures (for further details, see Section 5.1).

Although no causal relationship has been established, adverse events of special interests (AESIs) including infections (serious and opportunistic infections), lipase and amylase elevation, and seizure, are under close monitoring.

Due to the dominant role of cytochrome P450 3A (CYP3A)4/5 in the metabolism of evobrutinib, the compound may be a victim of drug-drug interactions (DDI) caused by inhibition (competitive/time-dependent) or induction of this enzyme by coadministered perpetrator drugs. The results of the completed clinical CYP DDI study (MS200527_0054) demonstrated that coadministration of evobrutinib with a moderate (e.g., fluconazole) or strong (e.g., itraconazole) CYP3A4 inhibitor resulted in a significant DDI. Evobrutinib peak and total exposures were increased by approximately 2.5- to 3.4-fold when evobrutinib was dosed concurrently with fluconazole or itraconazole. Based on these observations, the coadministration of medications that are moderate or strong inhibitors or inducers of CYP3A4/5 is not permitted during ongoing clinical studies.

No new potential risks emerged from the completed studies: EMR200527_001, MS200527_0019, MS200527_0017, MS200527_0022, EMR200527_002, and MS200527_0081.

Following the review of the totality of the safety data from the clinical studies with evobrutinib, overall, evobrutinib was well tolerated in MS patients up to 75 mg twice daily. The safety profile of evobrutinib has been consistent across doses and indications. No dose-related relationship has been observed for the most frequently reported TEAEs. Furthermore, evobrutinib at 45 mg twice daily in the fed state is expected to be similar to the tested 75 mg twice daily dose in the fasted state based on the available exposure/response data (see Section 4.3).

Overall, considering the unmet medical need in MS patients, reduction of MS activities (decreased in the number of Gd+ T1 lesions and lower ARR compared with placebo), convenience of an oral therapy and the measures put in place to mitigate the important identified and important potential risks, the benefit-risk of evobrutinib 45 mg twice daily fed supports continued clinical development of evobrutinib in this population.

More detailed information about the known and expected benefits and risks and reasonably expected AEs of evobrutinib may be found in Section 4.2 and the Investigator's Brochure. More detailed information about the known risks and benefits of teriflunomide are provided in the locally approved product information (e.g., relevant SmPC or the US Prescribing Information [USPI]).

Based on the available nonclinical and clinical data to date, the conduct of the study, as specified in this protocol, is considered justifiable.

3 Objectives and Endpoints

Table 1 Primary and Secondary Objectives, Endpoints and Estimands

Objectives	Endpoints (Outcome Measures)	Further Estimand Attributes
Primary		
To demonstrate superior efficacy with evobrutinib compared to teriflunomide in terms of ARR	ARR based on qualified relapses at Week 96 in participants with RMS	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death attributable to MS or treatment: Composite variable. Strategy for handling intercurrent event of death unattributable to MS or treatment: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: relapse rate ratio and CI (negative binomial [NB] model), with test of treatment effect based on test of relapse rate ratio

Objectives	Endpoints (Outcome Measures)	Further Estimand Attributes
Secondary		
To demonstrate the efficacy of evobrutinib relative to that of teriflunomide on disability progression	• Time to first occurrence of 12-week confirmed disability progression (CDP) as measured by the Expanded Disability Status Scale (EDSS) over 96 weeks • Time to first occurrence of 24-week CDP as measured by the EDSS over 96 weeks	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death attributable to MS or treatment: Composite variable. Strategy for handling intercurrent event of death unattributable to MS or treatment: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: hazard ratio and CI (Cox model), with test of treatment effect based on logrank test
To demonstrate the efficacy of evobrutinib relative to that of teriflunomide on patient reported symptoms and functional status	• Change from Baseline (CFB) in Patient Reported Outcomes Measurement Information System (PROMIS) physical function (PF) score at 96 weeks • CFB in PROMIS Fatigue score at 96 weeks	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: difference of least-squares means of score CFB at 96 weeks and CI (Mixed effect model for repeated measures [MMRM] model), with test of treatment effect based on difference of least-squares means

Objectives	Endpoints (Outcome Measures)	Further Estimand Attributes
To demonstrate the efficacy of evobrutinib relative to that of teriflunomide on MRI lesion parameters	• Total number of T1 Gd+ lesions based on assessments at Week 24, Week 48, and Week 96. • Total number of new or enlarging T2 lesions based on assessments at Week 24, Week 48, and Week 96	• Strategy for handling intercurrent event of treatment discontinuation: Treatment policy. Strategy for handling intercurrent event of death: While Alive • Population: as defined by inclusion/exclusion criteria • Population level summary: lesion rate ratio and CI (NB model), with test of treatment effect based on test of lesion rate ratio
To characterize the safety and tolerability of evobrutinib.	Safety as assessed by the nature, severity, and occurrence of adverse events (AEs) and adverse events of special interest (AESIs); vital signs; electrocardiograms (ECGs); absolute concentrations and change from Baseline in immunoglobulin (Ig) levels; and clinical laboratory safety parameters up to Week 108	Not applicable

Table 2 Tertiary/Exploratory Objectives and Endpoints

Objectives	Endpoints (Outcome Measures)
Tertiary/Exploratory	
To evaluate the effect of evobrutinib compared to teriflunomide on clinical parameters	• ARR based on qualified relapses at Week 48 • Time to first qualified relapse over 96 weeks • Qualifying relapse-free status at Week 96 • 12-week confirmed EDSS progression free status at Week 96 • 24-week confirmed EDSS progression free status at Week 96 • 12-week confirmed disability improvement status during 96 weeks assessed at Baseline, every 12 weeks up to Week 96 • 24-week confirmed disability improvement status during 96 weeks assessed at Baseline, every 12 weeks up to Week 96
To evaluate the efficacy of evobrutinib relative to that of teriflunomide on MRI parameters	• New Gd+ T1 lesion free status at Week 96 based on assessments up to Week 96 • Mean number of T1 Gd+ lesions per scan based on assessments up to Week 96 • Change in volume of T1 Gd+ lesions from Baseline to Week 96 based on assessments up to Week 96 • New or enlarging T2 lesion free status at Week 96 based on assessments up to Week 96 • Mean number of new or enlarging T2 lesions per scan based on assessments up to Week 96

Objectives	Endpoints (Outcome Measures)
	• Change in volume of T2 lesions from Baseline to Week 96 based on assessments up to Week 96 • Combined unique active (CUA) lesion free status at Week 96 based on assessments up to Week 96 • Mean number of CUA lesions per scan based on assessments up to Week 96 • Total number of CUA lesions based on assessments up to Week 96 • Total number of new T1 hypo-intense lesions based on assessments up to Week 96 • Percentage change in brain volume (BV) from Week 24 to Week 96 based on assessments up to Week 96 • Percentage change in thalamic volume from Week 24 to Week 96 based on assessments up to Week 96 • Percentage change in cortical grey matter volume from Week 24 to Week 96 based on assessments up to Week 96 • Change in normalized T1 intensity within pre-existing nonenhancing T2 weighted lesion volume from Baseline to Week 96 based on assessments up to Week 96 • Volume of slowly evolving lesions (SELs) at Weeks 24, 48, and 96
To evaluate the effect of evobrutinib compared to teriflunomide on functional parameters	• Time to $\geq$ 20% increase (confirmed at 12 weeks) in Timed 25-foot walk (T25-FW) during 96 weeks assessed at Baseline, every 12 weeks up to Week 96 • Time to $\geq$ 20% increase (confirmed at 12 weeks) in 9-hole Peg Test (9-HPT) during 96 weeks assessed at Baseline, every 12 weeks up to Week 96

Objectives	Endpoints (Outcome Measures)
To evaluate the effect of evobrutinib compared to teriflunomide on composite parameters	• Disease activity free status at Week 48 and Week 96 defined by: o Qualifying relapse-free status o New Gd+ T1 lesions free status and new or enlarging T2 lesion free status o 12-week confirmed EDSS progression free status • Time to first occurrence of 12-week CDP during 96 weeks based on a composite score defined by: o 12-week confirmed EDSS progression (at least 0.5- or 1.0-point change, depending on the Baseline EDSS) or; o 12-week confirmed worsening (≥ 20%) in T25-FW versus Baseline or; o 12-week confirmed worsening (≥ 20%) in 9-HPT versus Baseline. • No evidence of progression or active disease (NEPAD) at Week 48 and 96 as defined by: o No protocol-defined relapses on treatment o No 12-week disability progression on EDSS o No CDP of 20% or more in 9-HPT score o No CDP of 20% or more in T25-FW times o No new or enlarging T2 MRI lesions and no gadolinium-enhancing T1 lesions
To evaluate the efficacy of evobrutinib relative to that of teriflunomide on cognitive function	• Change from Baseline in SDMT score at Week 48 and at Week 96 based on assessments up to Week 96

Objectives	**Endpoints (Outcome Measures)**
To evaluate the efficacy of evobrutinib relative to that of teriflunomide on patient reported symptoms and functional status	• Change from Baseline in PROMIS PF score at Week 48 based on assessments at Baseline, every 12 weeks up to Week 48 • Change from Baseline in PROMIS Fatigue score at Week 48 based on assessments at Baseline, every 12 weeks up to Week 48 • Time to first occurrence of 12-week confirmed PF deterioration compared to Baseline (decrease of at least 5 points on PROMIS PF score) over 96 weeks • Change from Baseline in Patient Reported Outcome scores at Week 48 and Week 96: o Medical Outcomes Study 36 Item Short Form Health Survey (SF-36v2) o EuroQoL 5 Dimension 5 Levels (EQ-5D-5L)
To evaluate the effect of evobrutinib on health resource utilization (HRU) relative to that of teriflunomide over 96 weeks	• Absolute values of HRU, including but not limited to doctor/home/emergency visits, hospitalizations, paid assistance, and missed work
To characterize the PK profile of evobrutinib in participants with MS and to describe the exposure-response relationship between evobrutinib and efficacy endpoints and safety endpoints	• PK parameters: $CL_{/F}$, $V_{z/F}$, C_{max}, AUC, and T_{max} after a single dose (Day 1 data), and at steady state (Weeks 2, 4, 8, 12, 24, 48, 72 and 96 data)
To assess relationship between candidate disease biomarker and disease activity or treatment response	Level of biomarkers of disease with disease activity/treatment response at Baseline, Weeks 12, 24, 48, 72, 96; and upon relapse/disease progression (unscheduled)

Objectives	Endpoints (Outcome Measures)
To evaluate the relationship of the novel biomarkers of liver function and protein/genomic biomarkers of liver function compared to standard clinical chemistry endpoints (ALT)	Levels of novel biomarkers of hepatic function compared to ALT at Baseline, and on treatment
To assess the effect of evobrutinib on gene expression in whole blood	Gene expression at Baseline, Weeks 12, 24, 48, and 96

Table 3 Objectives and Endpoints for the Open Label Extension Period

Objectives	Endpoints (Outcome Measures)
OLE Period	
To evaluate the long-term safety, efficacy, and HRQoL of evobrutinib for an additional up to ~144 weeks	• Efficacy and HRQoL endpoints at Weeks 48, 96, and ~144 o ARR, based on protocol-defined qualified relapses o Change from Baseline in PROMIS PF score o Change from Baseline in PROMIS Fatigue score o Change from Baseline in Medical Outcomes Study 36 Item Short Form Health Survey (SF-36v2) • Efficacy and HRQoL endpoints over ~144 weeks o Time to first occurrence of 12-week confirmed EDSS progression over ~144 weeks o Time to first occurrence of 24-week confirmed EDSS progression over ~144 weeks o Time to first occurrence of 12-week confirmed PF deterioration compared to Baseline over ~144 weeks

Objectives	Endpoints (Outcome Measures)
	• Efficacy endpoints at Weeks 24, 48, 96, and ~144 o Total number of new or enlarging T2 lesions o Total number of T1 Gd+ lesions • Safety as assessed by the nature, severity, and occurrence of AEs and AESIs; vital signs; ECGs; absolute concentrations and change from Baseline in Ig levels; clinical laboratory safety parameters up to Week ~144

4 Study Design

4.1 Overall Design

This is a Phase III, multicenter, randomized, parallel group, double blind, double dummy, active controlled study of evobrutinib with an active control group teriflunomide, in participants with RMS.

Eligible participants will be randomized 1:1 to treatment with evobrutinib 45 mg twice daily, or teriflunomide 14 mg once daily (oral), stratified by region and Baseline Expanded Disability Status Scale (EDSS). Blinding will be accomplished using a double dummy design. The total sample size is planned to be 930 participants (approximately 465 participants per treatment group).

The 96-week Treatment Period will be preceded by a 4-week Screening Period (may be extended after discussion with the Medical Monitor but cannot exceed 8 weeks) and followed by a 4-week Safety Follow-up after treatment completion or early discontinuation.

Participants experiencing initial progression of disability between Week 72 and Week 96 will continue participating in the main study for up to 12 additional weeks. The continuation period will increase the duration of participation to a maximum of 108 weeks. Upon completion of the continuation period, participants will have the option of participating in the OLE or ending treatment (and returning for a 4-week Safety Follow-up).

Participants who complete the 96-week double blind, double dummy Treatment Period will be offered participation in the OLE Period of the study, which is expected to last approximately 144 weeks, with a 4-week Safety Follow-up Visit. The purpose of the OLE Period is to allow all the participants the opportunity to receive active treatment with evobrutinib and to collect long term safety and efficacy data. The Investigator should review the optional OLE Period with the participant prior to the double blind, double dummy Week 96 visit. Signed consent will be obtained prior to participation in the OLE Period.

Male participants will enter the OLE only after undergoing an accelerated elimination procedure (Appendix 8) and confirmation of teriflunomide plasma levels < 0.02 mg/L. The teriflunomide level will be reviewed centrally by an independent reviewer in order to maintain the blind of the study. For female participants entering the OLE Period, their last visit in the main study is also their first visit in the OLE.

The Safety Follow-up Visit will be deferred until treatment is stopped in the OLE Period, due to either a participant's premature withdrawal/early termination from the OLE, termination of the study by the Sponsor, or completion of the OLE Period.

In some cases, due to relapse, Institutional Review Board (IRB)/Independent Ethics Committee (IEC) approval, regulatory-specific, or other administrative delays, a participant may experience a treatment gap between the study intervention last dose received in the double blind, double dummy period (Week 96/up to Week 108 visit) and the start of the OLE study intervention treatment. Upon Principal Investigator request, these participants may still

be able to enroll in the OLE with approval from the Sponsor, on a case-by-case basis, provided that the treatment gap does not exceed 60 days from the last dose of study intervention received in the double blind, double dummy period (Week 96/up to Week 108 visit), and the start of the study intervention treatment in the OLE Period. For any male participants, the accelerated elimination procedure will be followed, as described above. If the day of rollover to the OLE occurs after the double blind, double dummy Safety Follow-up Visit, all assessments noted at the OLE Day 1 visit will need to be completed. For participants that rollover after the Week 96/up to Week 108 visit but prior to their scheduled Safety Follow-up Visit, concomitant medications and AEs will need to be reviewed and updated, and the Principal Investigator will need to ensure that the participant remains eligible for the study (see Section 5.3). No other additional assessments other than dispensing of study intervention will need to be completed.

An interim analysis (IA) for sample size re-estimation, based on 12-week confirmed disability progression (CDP; as measured by EDSS) data pooled from the present study and second Phase III study (MS200527_0080), will occur, triggered when approximately 35% to 45% of planned 12-week CDP events have been observed, as specified in the Interim Analysis Plan. The IA will evaluate conditional power (CP; probability of rejecting the null hypothesis at the PA of pooled 12-week CDP conditional on observed data) associated with the 12-week CDP endpoint based on pooled data. The IDMC will use a prespecified rule, defined in the IDMC Charter or related document, to determine whether CP is in the promising zone $0.3 \leq CP < 0.8$, and if so, by how much to increase enrollment up to a maximum of 35% (here $CP < 0.3$ represents the unfavorable/futility zone, while $CP \geq 0.8$ represents the favorable/efficacy zone).

4.2 Scientific Rationale for Study Design

This study was designed to determine the efficacy and safety of evobrutinib in participants with RMS. The primary objective will focus on reduction of relapses relative to teriflunomide over 96 weeks in adult participants with RMS based on the ARR at Week 96. The OLE will allow for assessment of long term safety and efficacy of evobrutinib.

The findings in Section 2 support the pathogenic contribution of B cells to MS damage. In contrast, a failed clinical study with another B cell targeting agent, atacicept, supports the notion that certain B cell subtypes may mediate beneficial anti-inflammatory effects (Kappos 2014). Novel nondepleting B cell therapies may deliver a more favorable benefit risk profile than current B cell directed therapeutic approaches. In the Phase II study of evobrutinib, the PA showed a clinically significant reduction in MRI activity in participants treated with evobrutinib compared to placebo. Furthermore, the results indicate that the ARR in the OLE remained low and similar to those observed in the PA and BEA.

There is consensus in the MS community, that the use of placebo in Phase III studies with RMS participants is no longer ethical, due to the availability of established and effective therapies (Polman 2008). Teriflunomide is an acknowledged therapy that has recently been used in pivotal studies as a comparator and is well-suited for the objectives of this study. It is currently widely used in the treatment of MS and has a well-established efficacy and safety

profile. Teriflunomide is administered as a tablet once daily and eliminates the burden of injections specific to other standard of care therapies such as interferons. For further information on teriflunomide, refer to the locally approved product information (e.g., relevant SmPC or USPI).

The study will have a double dummy design, to minimize the potential for bias and maintain the integrity of the clinical data generated from this study. It also reduces the risk of concluding that superiority to the active comparator was driven by participant and assessor bias.

This study plans to enroll RMS participants according to the McDonald MS 2017 criteria (Thompson 2018) with an EDSS score of 0 to 5.5 at Screening and Baseline (Day 1) who had at least one or more documented relapses within the 2 years before Screening with either:

- one relapse which occurred within the last year prior to randomization, OR
- the presence of at least 1 T1 Gd+ lesion within 6 months prior to randomization

These criteria have been implemented to further characterize the benefits of treatment with evobrutinib in a wide range of RMS participants with varying degrees of disease activity and severity. The age range will be limited to ≤ 55 years with the aim to avoid confounding by neurological conditions prevalent in older participants.

The proposed study endpoints are widely accepted as clinically relevant, and have been used in numerous pivotal clinical studies in RMS. The primary endpoint for the study will be ARR over 96 weeks, based on qualified relapses. Secondary efficacy endpoints will include time to 12 or 24-week CDP (as measured by EDSS), and total number of new or enlarging T2 lesions at Week 96 (based on postbaseline assessments at Weeks 24, 48, and 96), total number of T1 Gd+ lesions at Week 96 (based on postbaseline assessments at Weeks 24, 48, and 96). Prevention of relapses, prevention/delay of accumulation of sustained neurological disability, as well as effect on MRI are meaningful goals in the treatment of participants with RMS.

Patient Reported Outcomes (PROs) including those assessing fatigue and physical function (PF), are included as a secondary endpoint and several exploratory endpoints in the current study. Fatigue has been reported as the most bothersome symptom for MS patients among other concerns, such as bladder and bowel problems, cognitive impairment, visual disorders, musculoskeletal issues e.g., stiffness, spasm, walking difficulty and balance problems (Martin 2017, Patti 2011, Brañas 2000). In turn, these are associated with impairments in various functional areas, including instrumental activities of daily living (IADLs), limitations with physical activities such as participating in sports, limitations related to ability to work or study, and social interactions (Patti 2014, LaRocca 2011). The current data from randomized controlled studies support positive treatment effects of currently available DMDs (including, teriflunomide, dimethyl fumarate, natalizumab, and ocrelizumab) on Health-related Quality of Life (HRQoL), i.e., improving or preventing the worsening of HRQoL (Jongen 2017).

Refer to the Investigator's Brochure for further information about the nonclinical and clinical programs and Guidance for the Investigator.

4.3 Justification for Dose

Selection of the Phase III dose was based on pharmacodynamic (PD), efficacy, and safety data from the Phase IIb study MS200527_0086, and supported by clinical pharmacology studies where dosing of the tablet was included (Studies MS200527_0017 and MS200527_0019). The models describing the exposure-response relationship between PK and BTK occupancy (target engagement), ARR, number of T1 Gd+ lesions and number of new or enlarging (active) T2 lesions (efficacy endpoints) were used in the selection of the dose. The results of the modeling of evobrutinib concentrations, BTK occupancy, ARR, T1 Gd+ and active T2 lesions were used to select the dose of 45 mg twice daily (fed) that will deliver efficacy anticipated to be better than 75 mg once daily (fasted) or equal to 75 mg twice daily (fasted) observed in the RMS Phase IIb study.

In the exposure-response analysis of Study MS200527_0086 for participants with RMS, a steep exposure-response curve was observed with a threshold evobrutinib exposure ($AUC_{0\text{-}24,ss}$) of 355 ng•h/mL associated with the onset of a reduction in ARR, and a significant reduction in ARR observed at evobrutinib exposure of ≥ 400 ng•h/mL. Participants with exposure ≤ 355 ng•h/mL had mean ARR values between 0.42 to 0.63, whereas participants with exposure > 355 ng•h/mL had mean ARR values between 0.00 to 0.14. The significant reduction in ARR was observed when trough BTK occupancy was > 95%, indicating high target engagement over the entire dosing interval. In addition, for the participants with RMS and > 0 Gd+ T1 lesions at Baseline or > 13 cc T2 lesion volume at Baseline, a significant reduction in the number of T1 Gd+ lesions, and the number of active T2 lesions was observed when evobrutinib exposure was greater than 468 ng•h/mL. This significant reduction in the number of active T2 lesions was observed when trough BTK occupancy was > 96%.

Thus, the criteria for selection of an efficacious dose to be used in Phase III is based on targeting exposure of ≥ 468 ng•h/mL and with a high percentage of participants expected to have > 95% trough BTK occupancy.

One dose of evobrutinib will be utilized in this study. A dose of 45 mg twice daily, when given with a meal, is expected to provide a (geometric) mean evobrutinib daily exposure ($AUC_{0\text{-}24,ss}$) of 840 ng•h/mL, with only 7.6% of the participants expected to have a daily exposure of < 400 ng•h/mL. The 45 mg twice daily dose will be administered to participants with a meal, where food (either a high-fat or low/moderate-fat meal) has been shown to increase evobrutinib exposure by 50%. Thus, the daily exposure with 45 mg twice daily fed is similar to the exposure observed in the Phase IIb study for 75 mg twice daily fasted (835 ng•h/mL), which provided a significant reduction in ARR (see Section 2.1). Mean peak evobrutinib exposure ($C_{max,ss}$) for 45 mg twice daily fed is expected to be 24% less (128 ng/mL) than the observed mean peak exposure (168 ng/mL) for 75 mg twice daily fasted.

In addition, high BTK occupancy (at least 95% occupancy in 93% of the participants) is expected at the end of a dosing interval; therefore, similar efficacy to that of 75 mg twice daily (BTK occupancy of at least 95% in 92% of the participants).

Teriflunomide will be administered orally at the highest approved dose of 14 mg once a day. Refer to the locally approved product information (e.g., relevant SmPC or USPI) for further details.

Additional information about the justification for dose for evobrutinib may be found in the Investigator's Brochure.

4.4 End of Study Definition

A participant has completed the study if he/she has completed all study parts, including the Treatment Period and the last visit (Safety Follow-up Visit).

The end of the study is defined as the date of the last visit of the last participant.

5 Study Population

The criteria in Sections 5.1 and 5.2 are designed to enroll only participants, who are appropriate for the study, thereby ensuring the study fulfills its objectives. All relevant medical and nonmedical conditions should be taken into consideration when deciding whether a participant is suitable for this study.

Prospective approval of protocol deviations to inclusion and exclusion criteria, also known as protocol waivers or exemptions, is not permitted.

Before performing any study assessments that are not part of the participant's routine medical care, the Investigator will confirm that the participant or the participant's legal representative has provided written informed consent, as indicated in Appendix 2 Study Governance.

5.1 Inclusion Criteria

Participants are eligible to be included in the study only if all the following criteria apply:

Age

1. Are 18 to 55 years of age at the time of signing the informed consent.

Type of Participant and Disease Characteristics

2. Are diagnosed with RMS (relapsing-remitting multiple sclerosis [RRMS] or SPMS with relapses) in accordance with 2017 Revised McDonald criteria (Thompson 2018).
3. One or more documented relapses within the 2 years before Screening with either:
    a. one relapse which occurred within the last year prior to randomization, OR
    b. the presence of at least 1 Gd+ T1 lesion within 6 months prior to randomization.
4. Have an EDSS score of 0 to 5.5 at Screening and Baseline (Day 1)
    a. Participants with an EDSS score ≤ 2 at Screening and Baseline (Day 1) are only eligible for participation if their disease duration (time since onset of symptoms) is no more than 10 years.

5. Are neurologically stable for ≥ 30 days prior to both Screening and Baseline.

Sex

6. Are female or male
   a. Male Participants:

      Agree to the following during the study intervention period and for at least 2 years after study intervention due to the long elimination period for teriflunomide of 2 years, unless the participant undergoes an accelerated elimination procedure (see Appendix 8) with a confirmed teriflunomide level of < 0.02 mg/L after the last dose of study intervention:

      • Refrain from donating sperm

      PLUS, either:

      • Abstain from intercourse with a Woman of Childbearing Potential (WOCBP).

      OR

      • Use a male condom:
        - When having sexual intercourse with a WOCBP, who is **not** currently pregnant, and advise her to use a highly effective contraceptive method with a failure rate of < 1% per year, as described in Appendix 3, since a condom may break or leak.

   b. Female participants

      • Are **not** pregnant or breastfeeding, and at least one of the following conditions applies:
        - Not a WOCBP.

      OR

        - If a WOCBP, use a highly effective contraceptive method (i.e., with a failure rate of < 1% per year), preferably with low user dependency, as described in Appendix 3 for the following time periods:
          - Before the first dose of the study intervention(s), if using hormonal contraception:
            - Has completed at least one 4-week cycle of an oral contraception pill and either had or has begun her menses

          OR

            - Has used a depot contraceptive or extended-cycle oral contraceptive for at least 28 days and has a documented negative pregnancy test using a highly sensitive assay.

AND

- A barrier method, as described in Appendix 3
  - Have a negative serum or highly sensitive urine pregnancy test, as required by local regulations, within 4 to 8 weeks and a highly sensitive urine pregnancy test at Baseline before the first dose of study intervention. If a urine test cannot be confirmed as negative (e.g., an ambiguous result), a serum pregnancy test is required.
  - During the Intervention Period

    For at least 2 years after study intervention due to the long elimination period for teriflunomide of (up to) 2 years, unless the participant undergoes an accelerated elimination procedure (see Appendix 8) with a confirmed teriflunomide level of < 0.02 mg/L after the last dose of study intervention and agree not to donate eggs (ova, oocytes) for reproduction during this period (see Appendix 9). The Investigator evaluates the effectiveness of the contraceptive method in relationship to the first dose of study intervention.

- Additional requirements for pregnancy testing during and after study intervention are in Sections 8.2.4 and 8.2.5.
- The Investigator reviews the medical history, menstrual history, and recent sexual activity to decrease the risk for inclusion of a female with an early undetected pregnancy.

Informed Consent

7. Capable of giving signed informed consent, as indicated in Appendix 2, which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and this protocol.
8. Participants must be contactable by email or telephone throughout the study.

5.2 Exclusion Criteria

Participants are excluded from the study if any of the following criteria apply:

Medical Conditions

1. Participants diagnosed with progressive MS, in accordance with the 2017 Revised McDonald criteria, as follows:

   a. Participants with primary progressive MS.

   b. Participants with SPMS without evidence of relapse.

2. Disease duration > 10 years in participants with an EDSS ≤ 2.0 at Screening and Baseline (Day 1).

3. Immunologic disorder other than MS or any other condition requiring oral, intravenous (IV), intramuscular, or intra-articular corticosteroid therapy, with the exception of well-controlled Type 2 diabetes mellitus or well controlled thyroid disease.
4. History or current diagnosis of other neurological disorders that may mimic MS, including but not limited to: neuromyelitis optica, transverse myelitis, bilateral optic neuritis of simultaneous onset, Lyme disease, HTLV-1-associated myelopathy, untreated vitamin B12 deficiency, neurosarcoidosis, cerebrovascular disorders, documented peripheral neuropathy (including polyneuropathy or mononeuropathy).
5. History or current diagnosis of PML. If a brain MRI has findings suggestive of PML, cerebrospinal fluid JC virus polymerase chain reaction (CSF JCV PCR) should be performed to rule out PML (see Appendix 7).
6. Active, clinically significant viral, bacterial, or fungal infection, or any major episode of infection requiring hospitalization or treatment with parenteral anti-infectives within 4 weeks of Screening, or completion of oral anti-infectives within 2 weeks before or during Screening, or a history of recurrent infections (i.e., 3 or more of the same type of infection in a 12-month rolling period). Vaginal candidiasis, onychomycosis, and genital or oral herpes simplex virus considered by the Investigator to be sufficiently controlled would not be exclusionary.
7. The participant:
    - Has a history of or current diagnosis of active tuberculosis (TB)

      OR
    - Is currently undergoing treatment for latent TB infection (LTBI)

      OR
    - Has an untreated LTBI as determined by documented results within 3 months of the Screening Visit of a positive TB skin test with purified protein derivative (PPD) with induration ≥ 5 mm.

      OR
    - Has current household contacts with active TB, unless prophylaxis treatment has been completed and documented evidence that household contacts have completed treatment.

      OR
    - Has a positive QuantiFERON-TB test at Screening, unless the participant has completed chemoprophylaxis for LTBI (as per applicable local guidelines) prior to the Screening Visit.

Study participants in high TB burden settings (> 100 cases / 100,000 individuals [World Health Organization, 2019], based on World Health Organization [WHO] TB database [WHO TB Burden Estimates]) must repeat QuantiFERON testing at least annually at the visits indicated (see SoA Sections 1.3.1 and 1.3.2), using the assay that was negative at Screening (see Exclusion Criterion 8).

In addition, participants can be tested for TB at any time during the study, at the discretion of the Investigator.

Participants with documented completed appropriate LTBI treatment would not be excluded and are not required to be tested.

Note: TB skin test with PPD will not be performed at Screening. Reference to TB skin test results above is in reference to a potential participant's past results.

8. If the QuantiFERON-TB test results are indeterminate, then the individuals will be evaluated with T-SPOT.TB at the request of the Investigator. In this case, if the T-SPOT.TB is negative, the individual may be enrolled (see Section 8 for exceptions to tests analyzed by a central laboratory). If T-SPOT.TB is not available, the individual is excluded from participation in the study.
9. Individuals with a diagnosis of hemochromatosis, Wilson's disease, alpha-1-antitrypsin deficiency, or any other chronic liver disease including Gilbert's disease will be excluded from the study.
10. Individuals with elevated transferrin saturation (> 50% transferrin saturation in males; and > 40% transferrin saturation in females) and/or with elevated ferritin levels > 500 µg/L will be excluded.
11. Individuals with sickle cell anemia, thalassemia and/or any chronic blood disorder requiring blood transfusions will be excluded from the study.
12. History of splenectomy at any time, or any major surgery within 2 months prior to Screening.
13. History of myocardial infarction or cerebrovascular event within 6 months prior to Screening, or current active angina pectoris, history of or current congestive heart failure New York Heart Association (NYHA) Class III or Class IV, uncontrolled seizures (remote infantile febrile seizures are not exclusionary), prolonged untreated hypertension (systolic ≥ 160 mm Hg and/or diastolic ≥ 100 mm Hg), active gastrointestinal bleeding, or any other significant active medical condition in the Investigator's opinion or Sponsor's/designee's opinion.
14. A history of attempted suicide within 6 months prior to Screening or a positive response to items 4 or 5 of the Columbia-Suicide Severity Rating Scale (C-SSRS) at Screening.
15. An episode of major depression within the last 6 months prior to Screening (clinically stable minor depression is not exclusionary).

16. History of cancer with the following exceptions:
    - A confirmed history of nonmelanoma skin cancer Stage 0 (in situ) or Stage 1, considered cured > 5 years is not exclusionary.
    - A history of in situ cervical cancer, considered cured > 5 years, is not exclusionary.
    - A history of Stage I prostate cancer with normal Prostate-Specific Antigen (PSA), considered cured for > 5 years, is not exclusionary.

    Any history of cancer not meeting these exceptions is exclusionary.
17. On Screening electrocardiogram (ECG), any abnormality (e.g., uncontrolled second or third degree atrioventricular conduction block, ventricular tachyarrhythmias) that in the Investigator's opinion may impact participation in the study.
18. Any other clinically significant abnormality per Investigator opinion.

Prior/Concomitant Therapy

19. Contraindication to teriflunomide or leflunomide or incompatibility with teriflunomide or leflunomide use, including;
    a. Hypersensitivity to teriflunomide, or to any excipients.
    b. Hypersensitivity to leflunomide, or to any excipients.
    c. Cessation of teriflunomide therapy due to poor tolerability or safety concerns, or suboptimal response.
20. Injectable (e.g., IV, intramuscular, intra-articular) or oral glucocorticoids, or ACTH (e.g., Acthar gel) within 4 weeks prior to randomization (inhaled and topical corticosteroids are allowed) (see Section 8).
21. Treatment with monthly IV methylprednisolone (see Section 6.5.1).
22. Treatment with beta-interferons or glatiramer acetate within 4 weeks prior to randomization.
23. Treatment with dimethyl fumarate, diroximel fumarate, or other approved fumaric acid esters within 4 weeks prior to randomization provided lymphocyte count is > 1,000 cells/µL prior to randomization.
24. Treatment with teriflunomide within 4 weeks with an accelerated elimination procedure or 14 weeks without the completion of an accelerated elimination procedure prior to randomization (see Appendix 8 for accelerated elimination procedure).
25. Use of lymphocyte trafficking blockers (e.g., natalizumab, fingolimod, or siponimod) within 48 weeks prior to randomization.
26. Use of IV immunoglobulin (Ig) or plasmapheresis within 12 weeks prior to randomization.

27. Treatment with rituximab and/or ocrelizumab. Participants who have received 1 dose of rituximab or ocrelizumab, and reason for treatment discontinuation was not treatment failure, will be eligible to enter the study if the last dose of rituximab or ocrelizumab was at least 48 weeks prior to randomization.
28. Treatment with any other B cell depleting therapy, BTK inhibitors (including evobrutinib), mitoxantrone, or lymphocyte-depleting therapies (e.g., alemtuzumab, antiCD4, cladribine, cyclophosphamide, azathioprine, total body irradiation, bone marrow transplantation).
29. Concomitant treatment with medications commonly used for symptom management of MS patients will be exclusionary as follows:
    a. Participants taking dantrolene are to be excluded. Participants on other antispasticity agents can be included if they have been on a stable dose over the 3 months prior to randomization.
    b. Participants on dalfampridine (Ampyra) or fampridine can be included only if they have been on a stable dose 3 months prior to randomization.
    c. Medications known to lower the seizure threshold are not permitted unless reviewed and the eligibility of the participant is confirmed by the Medical Monitor.
30. Treatment with medical marijuana for MS symptoms, unless it is consistent with local MS treatment guidelines and local regulations. Dosage, formulation, and route of administration should be recorded as a concomitant medication.
31. On anticoagulation, or antiplatelet therapy other than daily aspirin for cardioprotection. Fish oil supplements must be stopped 4 weeks prior to randomization.
32. Participants currently receiving (or unable to stop using prior to receiving the first dose of study intervention) potent (strong to moderate) inducers of CYP3A (must stop at least 3 weeks prior), medications or herbal supplements known to be potent (strong to moderate) inhibitors of CYP3A (must stop at least 1 week prior), or drugs mainly metabolized by CYP3A with a narrow therapeutic index (must be stopped at least 1 day prior; see Section 6.5). Diabetes medications such as tolbutamide, and pioglitazone, repaglinide and rosiglitazone or other CYP2C8 substrates are also exclusionary.

Prior/Concurrent Clinical Study Experience

33. Participation in any investigational drug study within 6 months **or** 5 half-lives of the investigational drug, **whichever is longest**, prior to Screening.

Diagnostic Assessments

34. Any of the following:

    a. History of or positive for human immunodeficiency virus (HIV) at Screening.

    b. History of or positive for hepatitis C virus (HCV) antibody and/or HCV RNA by PCR at Screening. However, if a participant has a history of HCV infection and has completed and documented appropriate treatment at least 1 year prior to Screening AND is negative for HCV RNA by PCR at Screening, participants will not be excluded from the study.

       Note: All participants found to be positive for anti-HCV antibody at Screening will have reflex testing performed for HCV RNA by PCR to assess study eligibility.

    c. Positive for hepatitis B surface antigen (HBsAg) at Screening.

    d. For participants who are negative for HBsAg at Screening but are anti-hepatitis B surface antibody positive without history of vaccination for Hepatitis B and/or anti-hepatitis B core antibody positive with or without history of vaccination for Hepatitis B at Screening, reflex testing for hepatitis B virus DNA (HBV DNA) by PCR will be performed:

       i. Hepatitis B antibody positive participants who have detectable HBV DNA are excluded.

       ii. Hepatitis B antibody positive participants who are HBV DNA negative are not excluded from the study. However, these participants will have HBV DNA monitoring by PCR at visits noted in the Schedule of Activities (SoA) (see Section 1.3).

35. Estimated glomerular filtration rate (eGFR) < 60 mL/min/1.73 m^2 as calculated by the 4-variable Modification of Diet in Renal-Disease equation by the central laboratory or any renal condition that would preclude the administration of gadolinium (e.g., acute kidney injury).

36. ALT, AST, amylase, or lipase > 2 × ULN of laboratory reference range, total bilirubin > 1.5 × ULN, or any other clinically significant laboratory abnormality.

37. Significant cytopenia, including neutrophil count < 1,500 / mm^3, platelet count < 75,000 / mm^3, absolute lymphocyte count < 1,000 / mm^3, or a white blood cell count < 3,500 / mm^3.

Other Exclusions

38. Any allergy, contraindication, or inability to tolerate teriflunomide or evobrutinib or any of their excipients, including lactose, which is an excipient in the oral Study Intervention (e.g., evobrutinib tablets, placebo tablets).

    Note: Individuals with acquired lactose intolerance are not excluded but should be aware that the oral study intervention contains lactose and should be monitored for gastrointestinal symptoms related to the increased consumption of lactose in the study intervention, and made aware of the risks.

39. Inability to comply with MRI scanning, including contraindications to MRI such as known allergy or other contraindications to gadolinium contrast media, claustrophobia, presence of a pacemaker, cochlear implants, ferromagnetic devices or clips, intracranial vascular clips, insulin pumps, nerve stimulators.

40. Vaccination with live or live-attenuated virus vaccine within 1 month prior to Screening.

41. Regular alcohol consumption within 6 months prior to the study defined as: an average weekly intake of > 14 units for males or > 7 units for females. One unit is equivalent to 8 g of alcohol: a half pint (~240 mL) of beer, 1 glass (125 mL) of wine or 1 (25 mL) measure of spirits.

5.3 Criteria for Entry into Open Label Extension Period

5.3.1 Inclusion Criteria for Open Label Extension Period

Participants who meet the following entry criteria may participate in the OLE Period:

1. Complete the 96-week, double blind, double dummy Treatment Period, and who, in the opinion of the Investigator, may benefit from treatment with evobrutinib.
2. Are able and willing to provide written informed consent for the OLE Phase (e.g., before the first administration on OLE Day 1) and to comply with the study protocol.
3. WOCBP are willing to continue to use the contraceptive methods as described in Section 5.1 (Inclusion Criterion 6) and Appendix 3.
4. Males are willing to undergo an accelerated elimination procedure prior to entering OLE (see Appendix 8 for accelerated elimination procedure).

5.3.2 Exclusion Criteria for Open Label Extension Period

Participants will be excluded from the OLE if they meet any of the following exclusion criteria at the OLE screening (OLE Day 1):

1. Participants who did not complete study intervention in main study/Week 96 Visit or Week 108 Visit.
2. Treatment with injectable (e.g., IV, intramuscular, intra-articular) or oral glucocorticoids, or ACTH (e.g., Acthar gel) within 30 days before OLE Day 1.

3. Contraindications to MRI, for example, presence of pacemakers or other implanted metal devices (excluding dental braces), an allergy to gadolinium based MRI contrast, renal impairment, or claustrophobia that cannot be medically managed.
4. History of suicidal ideation or an episode of clinically severe depression (as determined by the Investigator) within 12 weeks prior to OLE Day 1.
5. History of abnormal laboratory results that, in the opinion of the Investigator, are indicative of a significant cardiac, endocrine, hematologic, immunologic, metabolic urologic, pulmonary, gastrointestinal, dermatologic, psychiatric, renal, neurologic (other than MS), and/or other major diseases.
6. Any of the following abnormal blood tests during the Treatment Period requiring discontinuation of study intervention, and/or at End of Study visit or at OLE Day 1: alanine aminotransferase/serum glutamate pyruvate transaminase (ALT/SGPT), aspartate aminotransferase/serum glutamic oxaloacetic transaminase (AST/SGOT), gamma glutamyl-transferase, amylase, or lipase $\geq 2 \times$ ULN, eGFR < 60 mL/min/1.73 m^2 as calculated by the 4-variable Modification of Diet in Renal-Disease equation.
7. Female participants who have a positive pregnancy test result, are pregnant, or are currently breast feeding.
8. Inability to comply with study requirements.

5.4 Lifestyle Considerations

5.4.1 Meals and Dietary Restrictions

Refrain from consumption of red wine, Seville oranges, grapefruit or grapefruit juice, pomelos, grapefruit hybrids, exotic citrus fruits, cranberries, or their juices from 7 days before the start of study intervention until after the final dose (see Section 6.5.4).

5.4.2 Caffeine, Alcohol, and Tobacco

During each dosing session, participants will abstain from alcohol for 24 hours before the start of dosing until after collection of the final PK and/or PD sample. See Exclusion Criterion 41 for alcohol consumption while participating in the study.

There are no restrictions on caffeine or tobacco intake.

5.5 Screen Failures

Individuals who do not meet the criteria for participation in this study (screen failure) may be rescreened once with approval by the Medical Monitor. The second Screening Period is a new 28-day Screening Period. This Screening Period may be extended after approval by the Medical Monitor (but cannot exceed 8 weeks) for the reasons described in Section 8. Rescreened participants will be assigned a new identification number. See Section 8 for required testing to be redone at rescreening.

6 Study Intervention(s)

Study intervention is any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to a study participant per the study protocol.

6.1 Study Intervention(s) Administration

Study Intervention Name:	Evobrutinib	Teriflunomide
Dose Formulation:	Tablets	Tablets
Unit Dose Strength(s)/ Dosage Level(s):	45 mg (active or placebo)	14 mg (active or placebo)
Route of Administration:	Oral	Oral
Dosing Instructions:	Evobrutinib to be taken with a meal twice per day	Teriflunomide to be taken with a meal once per day
	First oral dose of study intervention on the day of the in-clinic study visits be administered onsite to allow predose assessments as indicated in SoA (see Section 1.3.1).	
Supplier/ Manufacturer:	Evobrutinib and placebo will be supplied by the Sponsor.	Teriflunomide and placebo will be supplied by the Sponsor.
Packaging and Labeling	Evobrutinib and placebo will be packed in blister wallets/kits. Teriflunomide and placebo will be packed in blister wallets/kits. Each kit will be packaged and labeled per all applicable regulatory requirements and Good Manufacturing Practice Guidelines.	

6.2 Study Intervention(s) Preparation, Handling, Storage, and Accountability

Storage

All study intervention supplied to each site must be stored in their original containers carefully and safely. The storage facility at the study site must be locked and temperature controlled. Evobrutinib and its placebo must be stored at no more than 30°C (86°F). Teriflunomide and its placebo must be stored at 15°C to 25°C (59°F to 77°F).

In case there has been a temperature deviation at the clinical site, the site must contact the clinical research associate without delay for further evaluation and assessment by the designated quality assurance personnel at Merck Healthcare KGaA or delegated personnel at the packaging and distribution provider. The medication with the temperature excursion

should still be stored at the required temperature, but quarantined during the investigations and must be appropriately labeled as "quarantine storage".

Detailed recommendations for the use of teriflunomide are described in the locally approved product information (e.g., relevant SmPC or USPI), as appropriate.

The preparation, handling and storage of the study interventions will be documented in a separate Pharmacy Manual.

The Investigator, institution, or the head of the medical institution (where applicable) is responsible for study intervention accountability, reconciliation, and record maintenance (i.e., receipt, reconciliation, and final disposition records).

- Upon receipt of the study intervention(s), the Investigator or designee must confirm appropriate temperature conditions have been maintained during transit and any discrepancies are reported and resolved before use. Also, the responsible person will check for accurate delivery. Further guidance and information for study intervention accountability are provided in the Pharmacy Manual.
- Only participants enrolled in the study may receive study intervention(s) and only authorized site staff may supply it. All study intervention(s) must be stored in a secure, environmentally-controlled, and monitored (manual or automated) area, in accordance with the labeled storage conditions, and with access limited to the Investigator and authorized site staff.
- Dispensing will be recorded on the appropriate accountability forms so that accurate records will be available for verification at each monitoring visit.
- Study intervention(s) accountability records at the study site will include the following:
  - Confirmation of receipt, in good condition and in the defined temperature range.
  - The inventory provided for the clinical study and prepared at the site.
  - The dose(s) each participant used during the study.
  - The disposition (including return, if applicable) of any unused study intervention(s).
  - Dates, quantities, batch numbers, container numbers, expiry dates, and the participant numbers.
- The Investigator site will maintain records, which adequately documents that participants were provided the doses specified in this protocol, and all study intervention(s) provided were fully reconciled.
- Unused study intervention(s) must not be discarded or used for any purpose other than the present study. No study intervention that is dispensed to a participant may be redispensed to a different participant.
- A Study Monitor will periodically collect the study intervention(s) accountability forms.
- Further guidance and information for the final disposition of unused study intervention(s) are provided in the Pharmacy Manual.

6.3 Measures to Minimize Bias: Study Intervention Assignment and Blinding

6.3.1 Study Intervention Assignment

Participants will be randomized in a blinded fashion to either evobrutinib or teriflunomide treatment in a 1:1 ratio. The randomization will be stratified by 2 factors: region (4 levels: North America, Western Europe, Eastern Europe, Rest of World) and Baseline (Day 1) EDSS (2 levels: < 4.0, ≥ 4.0). A unique participant identification number, assigned according to the Pharmacy Manual, will be used throughout the study.

Study using IWRS	After confirmation of participant's eligibility and at the last practical moment prior to study intervention administration, participants will be centrally allocated to either evobrutinib or teriflunomide in a 1:1 ratio using an Interactive Web Response System (IWRS) and per a computer-generated randomization list. The IWRS will be used to assign unique participant numbers, allocate participants to study intervention group at the randomization visit, and study intervention to participants at each study intervention visit. Before the study is initiated, the log in information and directions for the IWRS will be provided to each site. The site will contact the IWRS prior to starting study intervention administration for each participant.

6.3.2 Blinding

Blinding Method

- Study intervention assignment will be randomized and blinded. Blinding will be accomplished using a double dummy design given the differences in tablet characteristics between the study intervention and the active control.
- Packaging and labeling will be prepared to protect the blinded nature of the study. Study interventions will be provided in a manner that participants or Investigator will not be able to distinguish active study intervention from placebo. Each kit will be labeled by the manufacturer with a unique kit number; labeling will not indicate whether the medication is study intervention (evobrutinib/teriflunomide) or placebo. Blinded treatment kit numbers will be obtained through the IWRS.
- The independent Examining Investigator (assessor) (or their Qualified Examining Designee) will perform neurological assessments only, will remain blinded to all study interventions (evobrutinib and teriflunomide), all other study-related assessments, and to all laboratory results throughout the study. The central MRI reader will be blinded to all study interventions (evobrutinib and teriflunomide) throughout the study. The participants,

site staff, and the Investigator will be blinded to all study interventions throughout the Treatment Period. The Contract Research Organization (CRO) and Sponsor will be blinded to all study interventions until the database is locked for the PA (see Section 9.4).

- Study participants will be instructed to not discuss AEs, signs, and/or symptoms with the independent Examining Investigator (assessor).
- The Treating Investigator is a neurologist and will be the only one to discuss signs and symptoms with the participant to prevent unblinding of study personnel, based on associated clinical signs and symptoms related to study intervention. The Treating Investigator will be blinded to study intervention and will be responsible for:
  - Management of the routine neurological care of the participant, including the management of background treatment and associated safety monitoring
  - Assessment (including assignment of causality) and treatment of AEs and MS relapses
  - Review of hematology and blood chemistry results from the central laboratory to assess whether the participant's study intervention should be discontinued as per the criteria detailed in Section 7.1
  - The Treating Investigator may designate the backup Treating Investigator or the Treating Nurse at the investigational site to perform some of the tests and evaluations listed under "Treating Investigator"
- The participants, site staff, and the Investigators will remain blinded to all study interventions during the Safety Follow-up.
- The IDMC and supporting independent statistician will be unblinded to treatment, as described in the IDMC charter.
- The bioanalytical monitors and analytical laboratories will be unblinded to study treatment codes to enable sample testing of study intervention prior to database lock. The process by which masked participant identifiers will be used to prevent linking PK data with other clinical data will be documented.
- Study intervention information that would unblind the study participants will not be reported with participant identifiers to investigative sites or blinded personnel until the study has been unblinded.
- All other staff, other than those identified above, will remain blinded to the evobrutinib and teriflunomide study interventions.
- Only when the last participant reaches 96 weeks of study intervention (up to 108 weeks if participant experiences initial progression between Week 72 and Week 96 and is thereby assigned to an additional 12 weeks of study intervention) or discontinues study intervention prematurely, the protocol deviations are determined, and the database is locked for the PA will the drug codes be broken and made available for the data analysis. At that point, the CRO and Sponsor study teams will be unblinded to study intervention. For participants entering the OLE at the Week 96 (or Week 108) visit, the data will no longer be blinded. If the sample size is increased when the IA is performed, the PA will be triggered based on recalculated event count (see Section 9.4.4.2), but there will still be a

requirement that protocol deviations be determined and database locked prior to breaking the drug codes for the data analysis.

- All breaks of the study blind must be adequately documented.

Assignment Method Retention

- The IWRS will give the Investigator the ability to break the blind with respect to study intervention for any participant, removing the need for physical retention of the intervention assignment at the site.

6.3.3 Emergency Unblinding

In an emergency, the Investigator is solely responsible for determining if unblinding of a participant's study intervention assignment is warranted. Participant safety must always be the first consideration in this decision. The Sponsor must be notified within 24 hours after unblinding. The Investigator must provide the Sponsor the reason for unblinding without revealing the study intervention, except to the designated drug safety representative via the Emergency Unblinding Notification Form. The date of and reason for unblinding must be recorded in the source documents and Case Report Form (CRF). Contact information for unblinding in an emergency is given on the participant emergency card provided to each participant, as noted in Appendix 2 (Study Governance).

The Sponsor's drug safety department will submit any Suspected Unexpected Serious Adverse Reactions (SUSAR) reports to regulatory authorities and ethics committees with unblinded information, per applicable regulations. Only blinded information will be provided to the study team.

6.4 Study Intervention Compliance

The study intervention will be administered at the site on study visit days as defined in the SoA (see Section 1.3). All other administrations of study intervention will be done by the participant or participant's caregiver at home throughout the rest of the study. Participants or participant's caregiver will be asked to record the date and time of dosing and food intake around dosing in a participant diary.

Participants will be instructed to bring all study intervention, including the used packaging/empty boxes and all blisters, to each study visit indicated in the SoA (see Section 1.3), and to allow for the assessment of compliance with study intervention. Prior to discharge from each scheduled visit, participants will be given sufficient study intervention for at-home administration until the next scheduled visit during the Treatment Period. On study visit days indicated in the SoA (see Section 1.3), the previous week(s) study intervention adherence will be documented using pill counts.

Insufficient compliance with the protocol-specified dosing regimen is defined as receiving < 80% of the required number of scheduled doses of study medication. Noncompliance is further addressed in Sections 7.1 and 8.4.

6.5 Concomitant Therapy

All concomitant therapies (e.g., medicines or nondrug interventions) used from the time the participant signs the informed consent until the participant completed the study/until the participant's last visit, including any changes, should be recorded in the CRF. For prescription and over-the-counter medicines, vaccines, vitamins, and herbal supplements, record the name, reason for use, dates administered, and dosing information.

Contact the Medical Monitor for any questions on concomitant or prior therapy.

Treatment for Symptoms of Multiple Sclerosis

The Treating Investigator should attempt to maintain therapies or treatments for symptoms related to MS (e.g., walking ability, spasticity, incontinence, pain, fatigue) reasonably constant throughout the study. However, if such therapies need to be added or modified during the study due to changes in the participant's clinical presentation, then this could be done at the discretion of the Investigator, with the exception of those treatments listed in Section 6.5.3 (Prohibited Medications). Nonetheless, any medications that are considered necessary for the participant's well-being may be given at the discretion of the Investigator.

During the OLE, initiation of therapy with dalfampridine (Ampyra) is allowed, if indicated by the Treating Investigator.

6.5.1 Rescue Medicine

Participants who experience a MS relapse during study intervention may receive rescue medication pursuant to the following restrictions:

1. Up to 1 g daily of methylprednisolone administered intravenous for up to 5 consecutive days. Where possible, the use of corticosteroids should be avoided in the 3 weeks prior to a scheduled MRI scan. If participants receive corticosteroids for a relapse, every effort will be made to obtain the scan prior to the first steroid dose if the presteroid scan is within 1 week of the scheduled visit. In all instances, the treatment of the relapse per the Investigator's clinical judgement takes priority over the timing of the MRI scan.
2. Oral tapering of corticosteroid rescue medication is permitted, with a maximum of 15 days of tapering allowed.
3. In the treatment of acute exacerbations of MS, daily intramuscular or subcutaneous doses of Acthar gel, 80-120 units and not exceeding 3 weeks may be administered. An additional taper of up to 2 weeks with less frequent dosing frequency will be permitted as indicated.

6.5.2 Permitted Medicines

The only permitted medications are the following:

- Medications required per the medical history that:
  - Are not specifically prohibited by the protocol during the study,
  - Are considered necessary for the participants' welfare, and
  - Will not interfere with the study intervention

Medications under the conditions described above may be given at the Investigator's discretion.

Treatment with medical marijuana for MS symptoms is permitted, if it is consistent with local MS treatment guidelines and local regulations.

Any medicines that are considered necessary to protect the participant's welfare in emergencies may be given at the Investigator's discretion, regardless if it results in a protocol deviation. However, as this may impact continued participation in the study, the Medical Monitor must be informed that prohibited medication was used.

6.5.3 Prohibited Medicines

Medications prohibited before the study are listed in the exclusion criteria (see Section 5.2).

The following medications and therapies are not permitted during the study and would require discontinuation of the study intervention:

- Rituximab, ocrelizumab, and any other B cell depleting therapy, BTK inhibitors (including evobrutinib), mitoxantrone, or lymphocyte-depleting therapies (e.g., alemtuzumab, anti CD4, cladribine, cyclophosphamide, azathioprine, total body irradiation, bone marrow transplantation).
- Lymphocyte trafficking blockers (e.g., natalizumab, fingolimod, or siponimod).
- Intravenous Ig therapy and/or plasmapheresis and immunosuppressive treatments.
- Beta-interferons or glatiramer acetate.
- Dimethyl fumarate, diroximel fumarate, other approved fumaric acid esters, or leflunomide.
- Dantrolene. Other antispasticity agents are permitted at study entry if the participant has been on a stable dose over the 3 months prior to randomization. During the study, antispasticity agents can be added and/or the regimen modified based on the Investigator's clinical judgement in order to manage the participant's symptoms (see Section 6.5).
- Medications known to lower the seizure threshold should be avoided. If treatment with these medications is required, the Investigator must inform the Medical Monitor.

  - Dalfampridine (Ampyra) or fampridine are not permitted unless the participant has been on a stable dose for 3 months prior to randomization due to the possibility of confounding effects on key study measures.

- Anticoagulation (e.g., warfarin), fish oil supplements, or antiplatelet therapy other than daily aspirin for cardioprotection.
- Diabetes medications such as tolbutamide, and pioglitazone, repaglinide and rosiglitazone or other CYP2C8 substrates.
- Medications known to be potent (strong to moderate) inhibitors of CYP3A, potent inducers of CYP3A, or drugs mainly metabolized by CYP3A with a narrow therapeutic index.
- Cholestyramine and activated charcoal, for use other than for accelerated elimination procedure after discontinuation (see Appendix 8).
- Live-attenuated vaccines.
- Biologic therapies for MS are prohibited. Biologic therapies for other indications must be discussed with the Medical Monitor as their use may impact continued participation in the study, with the exception of insulin and antibodies used for bone density (e.g., denosumab), which are permitted.

For further information on teriflunomide-related DDI, see Appendix 9 and refer to the locally approved product information (e.g., relevant SmPC or USPI).

6.5.4 Other Interventions

Herbal or nutritional supplements (including, but not limited to, St. John's wort, grapefruit, Seville oranges, cranberries, or juices of these fruits) known to be potent inhibitors of CYP3A must be stopped at least 1 week prior to randomization.

Teriflunomide has additional potential for drug interactions and due to the blinded treatment assignment, these should be considered when treating participants. See Appendix 9 for potential teriflunomide DDI, as defined in the locally approved product information (e.g., relevant SmPC or USPI).

6.6 Dose Selection and Modification

Not applicable.

6.7 Study Intervention after the End of the Study

The Sponsor will not provide any additional care to participants after they leave the study because such care would not differ from what is normally expected for participants with relapsing forms of MS.

6.8 Special Precautions

See Section 7.1 for precautions related to abnormal liver function.

If any of the MRIs of the brain have findings suggestive of PML, CSF JCV PCR should be performed to rule out PML, see Appendix 7 for further details. See Section 7.1 for precautions related to suspected PML.

6.9 Management of Adverse Events of Interest

Adverse events of special interests are liver AEs (possible drug-induced, non-infectious, non-alcoholic, and immune-mediated), infections (serious and opportunistic infections), lipase and amylase elevation, and seizure. All serious and nonserious AESIs must be additionally documented and reported using the appropriate Report Form as specified in Appendix 4.

Liver adverse events

The elevations of transaminases observed in participants treated with evobrutinib were frequent, asymptomatic, and reversible on discontinuation of evobrutinib. The mechanism is unknown.

Evobrutinib liver AESIs will include transaminases and bilirubin elevations, biological Hy's Law cases based on laboratory data; any type of acute or chronic hepatitis (any grade), suspected drug-induced liver injury, acute or chronic hepatic failure, fibrosis, cirrhosis and other liver damage-related conditions.

Infections

As per its mechanism of action, evobrutinib may impair B cell function which might lead to a decreased humoral immunity and consequently an increased risk of infection. Overall in completed studies in participants the Medical Dictionary for Regulatory Activities (MedDRA) system organ class (SOC) Infection was one of the most reported SOCs (e.g., in the MS200527_0086 RMS study approximately 18% to 32% of participants treated with evobrutinib reported infection; a similar rate was reported in the placebo group in the 0 to 24 week period), the individual events were of low grade, mainly Grade 1, nonserious and did not lead to study intervention discontinuation. Treatment of infections must be prompt and done in accordance with local standard of care depending on considerations such as the nature and severity of the infection and participant's overall health status. Any Common Terminology Criteria for Adverse Events (CTCAE) Grade ≥ 3 or serious adverse events (SAEs) of infection and opportunistic infection are considered as an AESI.

Amylase and lipase elevations

Asymptomatic elevations in amylase or lipase or both in participants treated with evobrutinib have been observed at a variety of time points and reported as TEAEs or noted as laboratory abnormalities. In RMS Study MS200527_0086, the incidence of TEAEs of lipase increased was slightly higher in evobrutinib 75 mg once daily and 75 mg twice daily arms (5 [9.4%] and 5 [9.3%], respectively) when compared to other arms (approximately between 4% to 6%).

However, shifts from Baseline to highest grade on treatment were similar across all treatment arms for both amylase and lipase. In evobrutinib studies in other indications and in healthy participants the incidence of TEAEs of increased amylase or lipase, or both was infrequent and no clinically meaningful differences were observed across treatment arms. Nonetheless, any CTCAE Grade ≥ 3 elevation of lipase and amylase and any type of pancreatitis are classified as AESIs.

Seizures

Seizures are more common in patients with MS than in the general population, occurring in 2% to 3% of MS patients (Poser 2003). Convulsions were observed in early studies of evobrutinib in dogs, however the plasma concentration of evobrutinib was approximately 140-fold greater than it is predicted for the dose used in this study. One participant with RMS with significant brain lesion load reported seizure of unclear clinical picture. The PK data for this participant did not exceed the expected values and was similar to other participants in the study. Anticonvulsant therapy was started and the participant continued treatment with evobrutinib with no reoccurrence. The Investigator did not consider the event to be related to evobrutinib. No event of convulsion/seizure was reported in other indications. Evobrutinib has been administered to approximately 800 patients with MS, RA and SLE. Moreover, an electroencephalogram study in healthy volunteers did not show an epileptogenic potential for evobrutinib. Any type of seizures/epilepsy of any grade or its consequences are classified as AESIs.

7 Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal

7.1 Discontinuation of Study Intervention

The Investigator should discontinue study intervention when a participant meets one of the conditions outlined below or if the Investigator believes that it is in the best interest of the participant. Due to the prolonged half-life of teriflunomide, and the possibility that the participant received teriflunomide (main study), accelerated elimination with cholestyramine or activated charcoal may be considered (for more details see Appendix 8). In cases for which permanent discontinuation of study intervention is required, no rechallenge will be allowed.

Criteria for Permanent Discontinuation of Study Intervention:

The Investigator should **permanently discontinue study intervention** for the criteria outlined below.

For laboratory or assessment related criteria:

- a neutrophil count < 500 / mm^3
- a neutrophil count 500–999 / mm^3 with fever
- platelet count < 25,000 / mm^3
- platelet count 25,000–49,999 / mm^3 with bleeding
- an increase in lipase to > 5 × ULN

- an increase in amylase to > 5 × ULN
- an increase in serum creatinine to > 3 × ULN
- any other laboratory abnormality of clinical significance, with the exception of lymphopenia (see below)
- QTcF > 500 msec OR an increase in QTcF > 60 msec relative to the participant's Baseline ECG (Day 1) is observed and confirmed (with a second ECG).
- Detectable HBV DNA. Should this occur, consultations with specialists, such as a hepatologist, can be performed at the discretion of the Investigator and the Medical Monitor should be informed. In addition, a comprehensive hepatic/autoimmune panel is required (see below).

  Note: HBV DNA is assessed only in Hepatitis B antibody-positive participants during the study (see SoA and Exclusion Criterion 34).
- For an increase in AST or ALT to > 3 × ULN in combination with an increase in bilirubin to > 2 × ULN, the study intervention should be permanently discontinued and the Medical Monitor informed. In addition, a comprehensive hepatic/autoimmune panel is required (see below).

For other reasons:

- Pregnancy. As a pregnant participant may have received teriflunomide and due to the prolonged half-life of teriflunomide, an accelerated elimination procedure must be performed and a teriflunomide level < 0.02 mg/L has to be reached (for more details see Appendix 8).
- Any events that endanger the safety of the participant.
- Sponsor decision to end clinical study.
- Adverse events, if discontinuation of study intervention is desired or considered necessary by the Investigator and/or participant.
- Noncompliance regarding study intervention (see Section 6.4)
- Use of prohibited medications, as defined in Section 6.5.3.
  - Note: Any medications that are considered necessary for the participant's well-being may be given at the discretion of the Investigator. However, as this may impact continued participation in the study, the Medical Monitor must be informed that prohibited medication was used.
- Lack of efficacy and/or progression of MS as defined by Investigator judgement or when a medication other than permitted medications (as defined in Section 6.5.2) is needed for treatment (see above regarding prohibited medications).

Criteria for Temporary Discontinuation of Study Intervention:

The Investigator should temporarily discontinue study intervention for the discontinuation criteria outlined below, inform the Medical Monitor, and perform confirmatory testing as instructed. Depending on the result of the confirmatory test, the Investigator should follow instructions as outlined below, including **permanently discontinuing study intervention** when indicated.

For any study intervention discontinuation related to laboratory or assessment results, the participant should be followed with additional testing as needed until a return to within normal limits or an acceptable value. In cases for which permanent discontinuation of study intervention is required, no rechallenge will be allowed.

Liver Function Testing criteria:

- For an increase in AST or ALT to > 3 × ULN, temporarily discontinue the study intervention and recheck the value within 72 hours (and no later than 1 week).
  - If the value is still > 3 × ULN upon retest, the study intervention should be permanently discontinued and the Medical Monitor informed.
  - If the value has decreased to ≤ 3 × ULN, the Investigator may reinitiate study intervention (see Section 7.1.1).
- For an increase in bilirubin of > 1.5 × ULN, temporarily discontinue the study intervention and recheck the value within 72 hours (and no later than 1 week).
  - If the value is still > 1.5 × ULN upon retest, the study intervention should be permanently discontinued and the Medical Monitor informed.
  - If the value has decreased to ≤ 1.5 × ULN, the Investigator may reinitiate study intervention.

For participants who temporarily or permanently discontinue study intervention because of abnormal liver function or detection of HBV DNA (see Exclusion Criterion 34), consultations with specialists, such as a hepatologist, can be performed at the discretion of the Investigator and the Medical Monitor should be informed. In addition, a comprehensive hepatic/autoimmune panel is required, including the following (performed by central laboratory, unless otherwise stated):

- International normalized ratio (INR), partial thromboplastin time, fibrinogen, high sensitivity c reactive protein
- Anti-HAV IgM, HBsAg, anti-HBc, anti-HBsAg, anti-HCV, anti-HEV IgG and IgM, anti-VCA IgG and IgM, anti-EA IgG, anti-EBNA IgG, anti-CMV IgG, IgM, EBV PCR, and CMV PCR
- Erythrocyte sedimentation rate (analyzed locally)
- Antinuclear antibody, antismooth muscle antibody, antibody to liver-kidney microsomes
- Albumin
- Alkaline phosphatase

- Ferritin and transferrin saturation
- Focused genetic testing for variants that confer risk for liver diseases and/or drug-related liver injury, including but not limited to: testing for variants in the High Iron Fe (human hemochromatosis protein) (HFE) gene (C282Y, H63D) in the setting of abnormal ferritin/transferrin saturation values as defined in Exclusion Criterion 10.

Other Laboratory Criteria:

- For a decrease in neutrophil count to 500 to 999 / mm^3 without fever, temporarily discontinue study intervention and recheck the value within 1 week.
    - **If the value is still < 1,000 / mm^3 upon retest, permanently discontinue study intervention.**
    - For an improvement to 1,000 to 1,499 / mm^3 upon retest, continue to hold the study intervention and recheck the value within 1 week.
        - If a further downward trend is observed, **permanently discontinue study intervention**.
        - If no further downward trend is observed, the Investigator may reinitiate study intervention.
    - If the neutrophil count returns to ≥ 1,500 / mm^3 upon retest, the Investigator may reinitiate study intervention.
- For a decrease in neutrophil count to 1,000 to 1,499 / mm^3, temporarily discontinue study intervention and recheck the value within 1 week.
    - If a downward trend is observed upon retest, **permanently discontinue study intervention**.
    - If no downward trend is observed, the Investigator may reinitiate study intervention.
- For a decrease in platelet count to 25,000 to 49,999 / mm^3 without bleeding, temporarily discontinue study intervention and recheck the value within 1 week.
    - **If the value is still < 50,000 / mm^3 upon retest, permanently discontinue study intervention.**
    - For an improvement to 50,000 to 74,999 / mm^3 upon retest, continue to hold the study intervention and recheck the value within 1 week.
        - If a further downward trend is observed upon retest, **permanently discontinue study intervention**.
        - If no further downward trend is observed, the Investigator may reinitiate study intervention.
    - If the platelet count returns to ≥ 75,000 / mm^3 upon retest, the Investigator may reinitiate study intervention.

- For a decrease in platelet count to 50,000 to 74,999 / mm^3, temporarily discontinue study intervention and recheck the value within 1 week.
    - If a downward trend is observed upon retest, **permanently discontinue study intervention**.
    - If no downward trend is observed upon retest, the Investigator may reinitiate study intervention.
- For an increase in amylase to > 2.0 to 5.0 × ULN, temporarily discontinue study intervention and recheck the value within 72 hours of receipt (and no later than 1 week).
    - **If the value is still** > 2.0 × ULN **or higher upon retest, permanently discontinue study intervention.**
    - For an improvement to > 1.5 to 2 × ULN upon retest, continue to hold the study intervention and recheck the value within 1 week.
        - If the value does not decrease upon retest, **permanently discontinue study intervention**.
        - If a downward trend is observed, the Investigator may reinitiate study intervention.
    - If the amylase returns to ≤ 1.5 × ULN upon retest, the Investigator may reinitiate study intervention.
- For an increase in amylase to > 1.5 to 2 × ULN, temporarily discontinue study intervention and recheck the value within 72 hours of receipt (and no later than 1 week).
    - If the value does not decrease upon retest, **permanently discontinue study intervention**.
    - If a downward trend is observed, the Investigator may reinitiate study intervention.
- For an increase in lipase to > 2 to 5 × ULN, temporarily discontinue study intervention and recheck the value within 72 hours of receipt (and no later than 1 week).
    - If **the value is still > 2 to 5 × ULN or higher, permanently discontinue study intervention.**
    - For an improvement to > 1.5 to 2 × ULN upon retest, continue to hold the study intervention and recheck the value within 1 week.
        - If the value does not decrease upon retest, **permanently discontinue study intervention**.
        - If a downward trend is observed, the Investigator may reinitiate study intervention.
    - If the lipase returns to ≤ 1.5 × ULN upon retest, the Investigator may reinitiate study intervention.

- For an increase in lipase to > 1.5 to 2 × ULN, temporarily discontinue study intervention and recheck the value within 72 hours of receipt (and no later than 1 week).
    - If the value does not decrease upon retest, **permanently discontinue study intervention**.
    - If a downward trend is observed, the Investigator may reinitiate study intervention.
- For any increase in serum creatinine > 1.5 × ULN but < 3 × ULN, temporarily discontinue study intervention and recheck the value within 72 hours of receipt (and no later than 1 week).
    - If the value does not decrease upon retest, **permanently discontinue study intervention**
    - If a downward trend is observed, the Investigator may reinitiate study intervention.
- For an absolute lymphocyte count < 200 / mm^3, study intervention should be temporarily discontinued, and follow-up testing should be conducted as clinically indicated
    - If the absolute lymphocyte count returns to ≥ 500 / mm^3, the Investigator may reinitiate study intervention.
    - If there is persistent lymphopenia < 200 / mm^3, **permanently discontinue study intervention**.
- For any other laboratory increase/decrease (as relevant) from Baseline to a clinically significant higher severity grade, temporarily discontinue study intervention and recheck the value within 72 hours of receipt (and no later than 1 week). The Investigator may reinitiate study intervention if an improving trend is observed.

For other reasons:

- At the first sign or symptom suggestive of PML, or brain MRI suggestive of PML, temporarily discontinue the study intervention and check for CSF JCV PCR (performed by central laboratory). If positive, permanently withdraw the study intervention (see Appendix 7).

  The SoA specifies the data to collect at study intervention discontinuation and follow-up, and any additional evaluations that need to be completed.

7.1.1 Temporary Discontinuation

See Section 7.1 for specific criteria for temporary discontinuation related to abnormal laboratory values.

7.1.2 Rechallenge

See Section 7.1 for specific criteria for rechallenge following temporary discontinuation for abnormal laboratory values.

7.2 Participant Discontinuation/Withdrawal from the Study

- A participant may withdraw from the study at any time, at his/her own request or may be withdrawn at any time at the discretion of the Investigator for safety, behavioral, compliance, or administrative reasons.
- If a participant is permanently withdrawn from study intervention due to stopping rules (see Section 7.1), the participant should also be withdrawn from the study.
- At the time it is determined that a participant is withdrawing from the study, the participant should, if possible, return for an Early Discontinuation Visit, as listed in the SoA. The SoA specifies the data to collect at study discontinuation and follow-up, and any additional evaluations that need to be completed.
    - Subsequent to the Early Discontinuation Visit, the participant should, if possible, enter the 4-week Safety Follow-Up Period. At the end of that 4-week period, the participant should return for the 4-week Safety Follow-up Visit. After completing this visit, the participant would be discontinued from the study.
    - If the participant withdraws consent for future involvement in the study, any data collected up to that point may still be used, but no future data can be generated, and any biological samples collected will be destroyed.
    - A participant has the right at any time to request destruction of any biological samples taken. The Investigator must document this in the site study records.

Participants who have discontinued after randomization (e.g., due to AEs or lack of efficacy; see Section 7.1) will not be replaced and will not be eligible to participate in the OLE. Participants who discontinue from the study should return for the Early Discontinuation Visit and the Safety Follow-up Visit 4 weeks from the day of discontinuation of study interventions.

7.3 Lost to Follow-Up

A participant will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site.

The following actions must be taken if a participant fails to return to the clinic for a required study visit:

- The site must attempt to contact the participant and reschedule the missed visit as soon as possible, counsel the participant on the importance of maintaining the assigned visit schedule, and ascertain if the participant wants to or should continue in the study.
- Before a participant is deemed “lost to follow-up”, the Investigator or designee must make every effort to regain contact with the participant: 1) where possible, make 3 telephone calls, 2) if necessary, send a certified letter (or an equivalent local method) to the participant’s last known mailing address, and 3) if a participant has given the appropriate consent, contact the participant’s general practitioner for information. These contact attempts should be documented in the participant’s medical record.

- Should the participant continue to be unreachable, he/she will be considered to have withdrawn from the study.

8 Study Assessments and Procedures

- Study assessments and procedures and their timing are summarized in the SoA.
- **No** protocol waivers or exemptions are allowed.
- Immediate safety concerns should be discussed with the Sponsor immediately upon occurrence or awareness to determine if the participant should continue or discontinue study intervention.
- Adherence to the study design requirements, including those specified in the SoA, is essential and required for study conduct.
- All screening evaluations must be completed and reviewed to confirm that potential participants meet all eligibility criteria. The Investigator will maintain a screening log to record details of all participants screened, to confirm eligibility, and if applicable, record reasons for screening failure.
- Prior to performing any study assessments that are not part of the participant's routine medical care, the Investigator will obtain written informed consent as specified in Appendix 2 Study Governance.
- Procedures conducted as part of the participant's routine medical care (e.g., blood count) and obtained before signing of the ICF may be used for screening or Baseline purposes provided the procedures met the protocol-specified criteria and were performed within the time frame defined in the SoA, if reviewed and approved by the Sponsor.

The following data will be collected:

- Demography: date of birth, sex (gender), ethnicity, and race, as permitted by local regulations.
- Medical history (including diagnosis and duration of MS): previous illness and surgeries (e.g., all during the past year and only major ones prior to that), concomitant illness, allergies, prior therapies for the target indication and reason for switch (i.e., relevant previous medications), therapies stopped or changed at entry into the study (includes use of drugs, alcohol, tobacco, and caffeine), special diets and, for women, menstrual status and date of last menstrual period.
- Participants will be asked to record the following information daily in a paper participant diary: dosing date and time and food intake around time of dosing.

All blood and urine tests will be analyzed by a central laboratory, with the following exceptions:

- Urine testing for β-human chorionic gonadotropin will be conducted at the local laboratories.
- Urine dipstick results will be interpreted locally. Urine microscopic examination will be conducted centrally. Please see the SoA (see Section 1.3) and Appendix 5 for information regarding abnormal dipstick results.
- PK and biomarker samples will be analyzed by the analytical laboratories specified by the Sponsor.
- HIV testing should be conducted centrally. Where required by local regulations, can be conducted and analyzed locally.
- T-SPOT test will be conducted at the local laboratories after an indeterminate QuantiFERON test.
- Erythrocyte sedimentation rate will be conducted as part of the hepatic assessment at the local laboratories.
- In addition, ECG results will be interpreted locally by the Investigator (see Section 8.2.3).

Screening

See the SoA (see Section 1.3.1) for a list of assessments completed at Screening.

The Screening packet will be reviewed by the Medical Monitor. See Appendix 2 for further details.

The Screening Period may be extended after approval by the Medical Monitor (but cannot exceed 8 weeks) for participants who have used systemic corticosteroids for their MS before Screening. For a participant to be eligible, systemic corticosteroids should not have been administered between Screening and Baseline. See Appendix 2 for details of the steps to be performed.

Retesting before Baseline

In case the Screening laboratory samples are rejected by the central laboratory or the results are not assessable or abnormal, the tests need to be repeated once within 4 weeks. Any abnormal Screening laboratory value that is clinically relevant should be retested once in order to rule out any progressive or uncontrolled underlying condition. The last value before randomization must meet study criteria. In such circumstances, the Screening Period may need to be prolonged after approval by the Medical Monitor but should not exceed 8 weeks.

For screen failures, see Section 5.5.

Rescreening

Participants who are considered screen failures after a first Screening Period may undergo rescreening once, after approval by the Medical Monitor. For those participants, the Screening Period for rescreening may be extended after approval by the Medical Monitor (but cannot exceed 8 weeks) for the reasons described above. If a participant is rescreened, all Screening tests will need to be repeated except as follows:

a. Documented TB testing if occurred within 3 months prior to the rescreening visit.

b. Hepatitis and HIV testing if occurred within 1 month prior to the rescreening visit.

Unscheduled Visit

Participants should be instructed that if, at any point during the study, they suspect that they are experiencing new or worsening neurological symptoms, including possible relapse, they should contact the Investigator as soon as possible after the onset of symptoms. If necessary, the participant should be evaluated by the Treating Investigator and if appropriate, the Examining Investigator (assessor), within the clinic and every effort should be made to complete this evaluation within 1 week after the start of symptoms. Any assessments needed to confirm the relapse should be performed at the discretion of the Investigator. Details should be documented within the relevant section(s) of the eCRF. The definition of a qualifying and nonqualifying relapse is provided in Section 8.1.1.1. Once the assessments by the Treating Investigator and the Examining Investigator (assessor) are completed, the disability score will be adjudicated within approximately 20 days.

If an MRI scan is indicated at an Unscheduled Visit for Neurological Worsening and Relapse Assessment, it should be performed prior to initiating corticosteroid therapy, where possible. In addition, care should be taken to avoid the participant being exposed to gadolinium more than once in a 4-week period, i.e. it may be necessary to cancel the MRI scan at the next scheduled visit (all other assessments should be completed at the visit as normal). If the Unscheduled Visit is conducted for a safety concern the assessments described below should be completed and any additional assessments and further management will be at the discretion of the Investigator.

The following will be performed at an Unscheduled Visit:

- Review of concomitant medications and procedures, evaluation of AEs, disease activity assessment (relapse assessment, EDSS), complete physical examination, vital signs, and a neurologic exam.
- Blood sample collection for safety assessments (hematology, biochemistry, and coagulation).
- Urine collection for urinalysis, and, if necessary, microscopy.
- Additional assessments can be performed at the discretion of the Investigator.

Telephone Visit

If a clinic visit is not scheduled, Telephone Visits will be scheduled with participants as indicated in the SoA. During the Telephone Visits, the following should be discussed with the participant:

- Any AEs or SAEs that may have occurred since the last on-site visit/phone call
- Any changes in concomitant medications that may have occurred since the last on-site visit/phone call
- Occurrence of MS relapses since the last on-site visit/phone call
- Overall wellbeing
- Pregnancy, if applicable, including confirm completion of home pregnancy testing and discuss results
- Occurrence of other notable medical events (e.g., surgeries, injuries, laboratory tests performed outside the study)
- Details of any overdose with the study drug that may have occurred since the last on-site visit/phone call.

Open-label Extension Period

After completing the Treatment Period (Weeks 1 through 96), participants will be offered the opportunity to participate in an OLE Period where all participants will receive evobrutinib. Signed consent will also be obtained prior to participation in the optional OLE Period.

Scheduled assessments will be performed according to the SoA (see Section 1.3.2) before administration of the study intervention. All scheduled visits during the OLE Period may take place within the visit windows specified in the SoA. Participants who discontinue early must return for the OLE End of Treatment Visit (~Week 144 / E/D, Visit 26). For further details see Sections 1.3.2 and 4.

8.1 Efficacy Assessments and Procedures

8.1.1 Neurological Assessment

The Examining Investigator (assessor) will perform the neurological examination, document the functional system scores (FSS) and assess EDSS scores. The Examining Investigator (assessor) or Qualified Examining Designee will be also responsible for performing and documenting results from: Timed 25-foot Walk (T25-FW), 9-Hole Peg Test (9-HPT) and the Symbol Digit Modalities Test (SDMT). He or she will have access only to data from assessments listed above. The Examining Investigator (assessor) will not be involved with any aspect of medical management of the participant and will not have access to participant data. Every effort will be made to ensure that there is no change in the Examining Investigator (assessor) throughout the course of the study for any individual participant. The Examining Investigator (assessor) will be trained and instructed not to discuss what adverse effects (if

any) the participant is experiencing from their medication. The Examining Investigator (assessor) will receive training in performing EDSS assessments prior to the beginning of the study and must have successfully passed an examination on performance of the Neurostatus EDSS examination within 24 months of participation. All Examining Investigators (assessors) will maintain ongoing training on performance of the Neurostatus EDSS examination throughout the course of the study.

Prior to being examined by the Examining Investigator (assessor), Treating Investigator and/or study coordinators should remind participants not to discuss what (if any) adverse effects they may be experiencing; this should be documented in the source documents.

8.1.1.1 Qualified Relapse

A qualifying relapse is the occurrence of new or worsening neurological symptoms attributable to MS (for > 24 hours, no fever, infection, injury, AEs, and preceded by a stable or improving neurological state for ≥ 30 days). The relapse should be accompanied by an increase of ≥ 0.5 EDSS, or 2 points increase on 1 of the FSS, or 1-point increase on ≥ 2 of the FSS. The increase in FSS scores must be related to the neurological symptoms which were reported as new or worsening. The change must affect the selected FSS (i.e., pyramidal, ambulation, cerebellar, brainstem, sensory, or visual, excluding bladder or bowel).

Episodic spasms, sexual dysfunction, fatigue, and mood change will not suffice to establish a relapse.

Adjudication of qualified relapses (regardless of whether they are identified during a scheduled or unscheduled visit) will be performed by the Endpoint Adjudication Committee (EAC) based on prespecified criteria, applied to data collected by the Treating Investigator, in a blinded fashion. Any assessments needed to confirm the relapse should be performed, and details of the relapse should be documented within the relevant section(s) of the eCRF. The criteria for a qualified relapse should be clear and there should be documentation of how each potential relapse did or did not meet the criteria. Participants who have a documented relapse during treatment are not required to discontinue study intervention unless they meet any of the criteria for discontinuation from the study intervention (see Section 7). Relapse assessments may be conducted as phone assessments between study site visits.

The annualized relapse rates over 96 weeks will be calculated based on qualified relapses.

8.1.1.2 Disability progression and Expanded Disability Status Scale

Disability progression 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) when the Baseline score is 5.0 or less and an increase of ≥ 0.5 when the Baseline score is 5.5. Disability progression is considered sustained when the initial increase in the EDSS is confirmed at a regularly scheduled visit at least 12 weeks or 24 weeks, after the initial documentation of neurological worsening.

Confirmed disability progression, sustained for 12 and 24 weeks after the initial documentation of neurological worsening, will be analyzed as secondary endpoints.

8.1.1.3 Confirmed Disability Improvement

Disability improvement will be analyzed only for the subgroup of participants with a Baseline EDSS score ≥ 2.0. For participants with a Baseline EDSS score ≥ 2 and ≤ 5.5, disability improvement is defined as a reduction in EDSS score ≥ 1.0 compared to Baseline EDSS score sustained for at least 12 or 24-weeks. All participants without disability improvement will be counted as not improved, independent of follow-up time.

8.1.1.4 Timed Twenty-Five Foot Walk

The Timed Twenty-Five Foot Walk (T25-FW) is a quantitative mobility and leg function performance test based on a timed 25-foot walk. The participant is directed to one end of a clearly marked 25-foot course and is instructed to walk 25 feet as quickly as possible, but safely. The time is calculated from the initiation of the instruction to start and ends when the participant has reached the 25-foot mark. The task is immediately administered again by having the participant walk back the same distance. Participants may use assistive devices when doing this task. T25-FW will be administered by the Examining Investigator (assessor) or Qualified Examining Designee.

A worsening of ≥ 20% is considered to have occurred in this task, when the time it takes to complete the task is longer by equal to or more than 20% than the time it took at Baseline. The worsening is considered confirmed at 12 weeks if the following assessment (at a scheduled visit of 12 weeks or more after the initial observed worsening) confirms it.

8.1.1.5 Nine Hole Peg Test

The 9-HPT is a brief, standardized, quantitative test of upper extremity function. Both the dominant and nondominant hands are tested twice. The participant is seated at a table with a small, shallow container holding nine pegs and a wood or plastic block containing nine empty holes. On a start command when a stopwatch is started, the participant picks up the nine pegs one at a time as quickly as possible, puts them in the nine holes, and, once they are in the holes, removes them again as quickly as possible one at a time, replacing them into the shallow container. The total time to complete the task is recorded. Two consecutive studies with the dominant hand are immediately followed by two consecutive studies with the nondominant hand. 9-HPT will be administered by the Examining Investigator (assessor) or Qualified Examining Designee.

A worsening of > 20% (either hand) is considered to have occurred in this task, when the time it takes to complete the task is longer by equal to or more than 20% than the time it took at Baseline. The worsening is considered confirmed at 12 weeks if the following assessment (at a scheduled visit of 12 weeks or more after the initial observed worsening) confirms it.

8.1.1.6 Symbol Digit Modalities Test

The SDMT has demonstrated sensitivity in detecting not only the presence of cognitive impairment, but also changes in cognitive functioning over time and in response to treatment. The SDMT is brief, easy to administer, and involves a simple substitution task that normal children and adults can easily perform. Using a reference key, the examinee has 90 seconds to pair specific numbers with given geometric figures. Responses will be written and administration time is just 5 minutes. SDMT will be administered by the Examining Investigator (assessor) or Qualified Examining Designee.

8.1.2 Brain Magnetic Resonance Imaging Scans

If a participant discontinues the study more than 4 weeks after his or her most recent MRI, during the double blind, double dummy phase of the Treatment Period, an MRI may be obtained at the Discontinuation Visit. The Screening/Baseline MRI scan should be acquired before randomization and dosing to allow for the readouts to be read by the central MRI reader (at least 14 days prior to randomization).

Magnetic resonance imaging is a useful tool for monitoring CNS lesions in MS. Various MRI derived parameters have been related to disease activity, including T1 weighted gadolinium-enhancing (T1 Gd+) lesions and new or enlarging hyperintense T2 (active T2) lesions. Other MRI readouts such as chronic T1 hypo-intense lesions are reflective of long term brain damage (black holes). Combined unique active lesions are defined as the sum of new T1 Gd+ lesions and active T2 lesions (without double counting). It is hypothesized that changes in brain volume may reflect brain atrophy as a result of MS-related tissue loss and may thereby correlate with long term clinical outcome in these participants (De Stefano 2002).

Brain MRI scans will be performed according to a standardized imaging protocol before and after the administration of single-dose gadolinium. Further details, including the scans required and the optimal MRI workflow, will be provided in a separate Imaging Manual that will be provided to each study site.

Images will be assessed and reported by an independent, blinded, centralized MRI reading service, provided by NeuroRx Research. The assessment will be performed in the absence of clinical information. At the site level, the local radiologist assigned to this study will have access to the MRI scans. The Treating Investigator has access to MRI scans in case of safety-related issues. Independent safety-reads of all MRI scans can be obtained from the local radiology centers. In addition, if a scheduled MRI scan is delayed or an unscheduled MRI scan is indicated, care should be taken to avoid the participant being exposed to gadolinium more than once in a 4-week period, i.e., it may be necessary to cancel the MRI scan at the next scheduled visit (all other assessments should be completed at the visit as per schedule). If the next scheduled visit is the Week 96, the Week 96 MRI scan should be performed as soon as the 4-week period since previous exposure to gadolinium has elapsed.

Gadolinium will be used to enhance T1-weighted lesions and to optimize clarity and accuracy of reporting. As gadolinium is excreted renally, participants with acute renal insufficiency (eGFR < 60 mL/min/1.73 m^2) will be excluded from the study (see Section 5.2).

8.1.3 Patient Reported Outcomes

PRO data will be collected at the study visit with an electronic tablet device at specified study visits (see the SoA for details, Section 1.3.1). The tablet with the PRO instruments will be distributed by the Investigator staff and completed in their entirety by the participant.

PROs should be completed prior to administration of study intervention and prior to any other study assessment(s) to ensure the validity of the instruments is not compromised, and data quality meets requirements of the Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims (Food and Drug Administration [FDA] 2009).

PRO data will be elicited from participants in this study to better characterize the clinical profile of evobrutinib. These PRO measurements are described in Sections 8.1.3.1, 8.1.3.2, and 8.1.3.3. Please note that the methods for collecting and analyzing PRO data are different from those for the ascertainment of observed or volunteered AEs. Due to these differences, PRO data will not be reported as AEs and no attempt will be made to resolve any noticeable severe symptoms or functional status.

To ensure missing data are minimized, all participants will be given detailed information and training about the nature of the PRO assessments and importance of such information for the study, at or prior to study start (based on materials provided).

Further guidance/information for study sites on procedures for collecting electronic PROs (ePROs), as well as strategies to ensure complete and high-quality PRO data will be provided in the study manual of operations.

8.1.3.1 Patient Reported Outcomes Measurement Information System

The National Institute of Health Patient Reported Outcomes Measurement Information System (NIH PROMIS) comprises an extensive set of item banks and short-form measures created from the item banks that assess physical, mental, and social aspects of health in adults and children, including symptoms such as pain, fatigue, and sleep disturbance, and health domains such as PF (Cella 2007).

The PROMIS PF item bank was identified as having great potential for the evobrutinib program due to several factors, despite limited previous use in MS (Amtmann 2018), and lack of an MS-specific short form. First, the content includes all key aspects of PF domain, such as IADL, lower extremity (mobility), back and neck (central), and upper extremity functioning domains (Rose 2014). Second, the development process of PROMIS items included a rigorous development and calibration process, ensuring the technical quality of items. Further, items capture the full continuum of PF, from low to high levels, which is a

useful characteristic for capturing changes over time. A short form specific to MS has recently been derived with input from MS patients (n = 57) and is currently undergoing validation (Kamudoni 2018). Measures from the PF item bank are scored on a T-score metric (higher scores = higher PF). The minimal clinically important difference (MCID) score cut-off for the short form is yet to be established. A provisional MCID estimate of 5 points is proposed, based on previous research on the PF item bank (Yost 2011, Amtmann 2018).

The PROMIS Fatigue item bank includes 95 items assessing the experience (frequency, duration and intensity) as well as the impacts of fatigue on physical, mental and social activities (Lai 2011). Psychometric properties of this bank have been established across different clinic populations (Cella 2016). An 8-item short-form specific to MS, derived based on input from clinicians (n = 36) and participants with MS (n = 48), is available (Cook 2012). This short form is currently undergoing further validation. Measures from the fatigue item bank are scored on a T-score metric (higher scores = higher fatigue); a provisional MCID estimate of 4 points is proposed based on previous research on the fatigue item bank (Yost 2011, Amtmann 2018).

The PROMIS approach offers flexibility in the selection of items and how these are administered, including use of bespoke measures, fixed short forms, or computerized adaptive testing. PROMIS based short forms for physical functioning and fatigue are currently in preparation for FDA qualification as a drug development tool (DDT) in MS (MS Working Group, 2018; DDT COA, 2018).

Physical function deterioration is defined as a reduction in PROMIS PF T-score ≥ 5.0 compared to Baseline PROMIS PF T-score sustained for at least 12 weeks.

8.1.3.2 Medical Outcomes Study 36-Item Short Form Survey Instrument

The Medical Outcomes Study 36-Item Short Form Survey Instrument (SF-36v2) is a 36-item questionnaire that measures 8 areas of participant reported health rated from 0 to 100 (McHorney 1992, Ware 1992, McHorney 1993, McHorney 1994, and Freeman 2000). The areas are:

- Physical function
- Role limitations due to health problems
- Bodily pain
- Social functioning
- General mental health
- Role limitations due to emotional problems
- Energy/fatigue
- General health perceptions

The instrument will be used to calculate a normalized score for each of the 8 health domain scales, a physical component summary (PCS) score, and a mental component summary (MSC) score, with higher scores indicating better health.

8.1.3.3 EuroQoL 5 Dimension 5 Levels

The EuroQoL 5 Dimension (EQ-5D) is a standardized instrument developed as a measure of health-related quality of life that can be used in a wide range of health conditions and treatments. The 5-level EQ-5D version (EQ-5D-5L) was introduced in 2009 to improve the instrument's sensitivity and to reduce ceiling effects, as compared to the EuroQoL 5 Dimension 3 Levels (EQ-5D-3L). The EQ-5D-5L essentially consists of 2 pages: the EQ-5D descriptive system and the EQ visual analogue scale (EQ VAS).

The descriptive system comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems and extreme problems. The levels of the five dimensions can be combined into a 5-digit number that describes the patient's health state.

EQ-5D-5L health states, defined by the EQ-5D-5L descriptive system, may be converted into a single health utility index value using country specific value sets. Health utility values facilitate the calculation of quality-adjusted life years that are used to inform pharmacoeconomic evaluations of health care interventions.

The EQ VAS records the respondent's self-rated health on a 20 cm vertical, visual analogue scale with endpoints labelled 'the best health you can imagine' and 'the worst health you can imagine'. This information can be used as a quantitative measure of health as judged by the individual respondents.

Higher scores on both EQ-5D-5L health utility values and EQ VAS represent a better HRQoL.

8.2 Safety Assessments and Procedures

The safety profile of the study intervention will be assessed through the recording, reporting and analysis of Baseline medical conditions, AEs, physical examination findings, vital signs, ECGs, and laboratory tests including Ig and subclass concentration.

Comprehensive assessment of any potential toxicity experienced by each participant will be conducted starting when the participants give informed consent and throughout the study. The Investigator will report any AEs, whether observed by the Investigator or reported by the participant; the reporting period is specified in Section 8.3.1.

8.2.1 Physical Examinations

- A complete physical examination will include, at a minimum, assessments of the Cardiovascular, Respiratory, Gastrointestinal and Neurological systems.
- Investigators should pay special attention to clinical signs related to previous serious illnesses.

Any physical exam abnormality findings, which are identified as clinically significant before the ICF is signed, will be captured on the Medical History eCRF. After the ICF is signed, any new physical exam abnormality findings will be captured on the Adverse Event form.

8.2.2 Vital Signs

- Height at Screening and weight will also be measured and recorded. Weight will be measured and recorded at each visit where vital signs are recorded as noted in the SoA.
- Temperature, pulse rate, respiratory rate, and blood pressure will be assessed.
- Blood pressure and pulse measurements will be assessed semisupine with a completely automated device. Manual techniques will be used only if an automated device is not available.
- Blood pressure and pulse measurements should be preceded by at least 5 minutes of rest for the participant in a quiet setting without distractions (e.g., television, cell phones).
- Vital signs will be measured in a semisupine position after 5 minutes rest and will include temperature, systolic and diastolic blood pressure, and pulse and respiratory rate.

8.2.3 Electrocardiograms

- Single 12-lead ECG will be obtained as outlined in the SoA using an ECG machine that automatically calculates the heart rate (HR) and measures PR, QRS, QT, and QTcF intervals and RR duration.
- The 12-lead ECG recordings will be obtained after 10 minutes of rest in a semisupine position.
- ECG results will be interpreted locally by the Investigator. In the event of findings that could represent clinically relevant cardiac issues (including but not limited to new or worsening arrhythmia, significant changes in ECG parameters, signs and/or symptoms that could represent cardiac events such as syncope, chest pain, etc.), additional evaluations can be performed per the Investigator's clinical judgement (including but not limited to repeat ECGs, echocardiography, evaluation by a cardiologist). The printout of the ECG is to be signed, dated, and filed in the Investigator's Site File along with a signed and dated copy (if the printouts are not on archive-quality paper). In addition, ECGs will also be stored digitally by the Sponsor.

8.2.4 Clinical Safety Laboratory Assessments

- Blood and urine samples will be collected for the clinical laboratory tests listed in Appendix 5, at the time points listed in the SoA. All samples should be clearly identified.
- Additional tests may be performed at any time during the study, as determined necessary by the Investigator or required by local regulations.
- The tests will be performed by the central laboratory. See Appendix 5 for exceptions.

- Local laboratory results are only required when central laboratory results are not available in time for study intervention administration and/or response evaluation. If a local sample is required, it is important that the sample for central analysis is obtained at the same time. Additionally, if the local laboratory results are used to make a study intervention decision or response evaluation, the results must be entered in the CRF. Immunoglobulin levels should not be done locally unless clinically indicated, as these results may lead to unblinding (se Section 8.2.6).
- The Sponsor must receive a list of the local laboratory normal ranges before shipment of study intervention(s). Any changes to the ranges during the study must be forwarded to the Sponsor or designated organization.
- The Investigator must review each laboratory report, document their review, and record any clinically relevant changes occurring during the study in the AE section of the CRF. The laboratory reports must be filed with the source documents.
- Laboratory/analyte results that could unblind the study will not be reported to investigative sites or other blinded personnel until the study has been unblinded.
- Pregnancy testing (serum or highly sensitive urine, as required by local regulations) will be conducted at monthly intervals during study intervention administration.
- Pregnancy testing (serum or highly sensitive urine, as required by local regulations) will be conducted at the end of relevant systemic exposure of the study intervention and correspond with the time frame for female participant contraception in Section 5.1.
- Reflex analysis testing for HBV DNA by PCR will be conducted for participants who are anti-hepatitis B surface antibody positive and/or anti-hepatitis B core antibody positive at Screening (see Exclusion Criterion 34). Whole blood samples of approximately 2 mL will be collected at the times specified in SoA for main study and OLE.

8.2.5 Pregnancy

Pregnancy testing (urine or serum as required by local regulations) should be conducted as summarized in the SoA (see Section 1.3) during intervention.

Urine pregnancy testing will be performed at home or at the site. Urine pregnancy test kits will be provided to the participants at site visits. The Investigator and/or delegated site staff will train the relevant participants to self-administer the urine pregnancy test, and will contact the participant by telephone to confirm completion of urine pregnancy testing and discuss results.

8.2.6 Immunoglobulin levels

Blood samples for Ig levels (IgM, IgA, IgG, and IgE) will be collected as noted in the SoA (see Section 1.3).

Samples will be analyzed by the central laboratory selected by the Sponsor. Samples will be collected, labeled, processed, stored, and shipped according to the instructions in the Laboratory Manual.

Results will not be disclosed to the sites, Sponsor, or representative prior to database lock to avoid unblinding. However, the IDMC will have access to these data as applicable.

8.2.7 Columbia-Suicide Severity Rating Scale

The Columbia-Suicide Severity Rating Scale (C-SSRS) will be used for prospective suicidality assessment. C-SSRS is a tool used to assess the lifetime suicidality of a participant and to track suicidal events through the treatment. The structured interview prompts recollection of suicidal ideation, including the intensity of the ideation, behavior and attempts with actual/potential lethality.

The scale will be administered by the Treating Investigator at the timepoints indicated in the SoA (see Section 1.3). The C-SSRS "Screening/Baseline" will be collected at Screening and Baseline and the C-SSRS "since last visit" will be collected at subsequent visits.

Participants who answer "yes" to any suicidal behavior questions or to suicidal ideation questions 4 or 5 on the C-SSRS during the study should be referred for appropriate psychiatric care. The decision to discontinue the participant from the study should be taken by the Investigator in conjunction with the mental health provider. Any mental health issues considered to be life threatening should be recorded as appropriate.

Please note: assessing the risk of suicide is a difficult and complex task when applied to the individual participant. Certainly, no single clinical scale can replace a thorough medical examination and suicide risk assessment. Ultimately, the determination of the presence of suicidality depends on clinical judgment.

8.2.8 Independent Data Monitoring Committee, Endpoint Adjudication Committee, and Study Steering Committee

Independent Data Monitoring Committee

An IDMC will be formed for this study to monitor interim safety and disease activity data on a regular basis to ensure ongoing surveillance of participant safety and to monitor the conduct of the study to protect its integrity, including recommendations about additional monitoring measures or risk mitigation procedures that may be deemed necessary to protect the study participants. After consideration, the Sponsor will inform the IDMC of any decision that will be taken in response to the IDMC recommendations. The IDMC will consist of a minimum of at least 3 expert members who are independent of the Sponsor. The members will be

appointed by the Sponsor based on their expertise in biostatistics, MS and additional members with expertise in hepatology. All IDMC members will have experience in the conduct of clinical studies. Members will not be Investigators in the study, nor will they have any conflict of interest with the Sponsor. Sponsor representatives and study Investigators are not eligible for membership on the IDMC. Details regarding IDMC roles, responsibilities, activities, and possible recommendations will be provided in a separate IDMC charter.

The IDMC will review unblinded interim data. The recommendations of the IDMC will not contain unblinded data or other information that could lead to Investigators or Sponsor representatives becoming unblinded. An independent statistician, who is not involved with study conduct and not a member of IDMC, is responsible for producing the unblinded interim data for IDMC review.

Endpoint Adjudication Committee

An EAC will be formed for centralized, blinded review and determination of qualified relapses (see Section 8.1.1.1). The EAC will be comprised of subject matter experts convened for the purpose of ensuring consistency across study sites and to allow for an unbiased endpoint assessment. Sponsor representatives and study Investigators are not eligible for membership on the EAC. The EAC will be managed by a central vendor; the vendor will be under contract to the Sponsor. Specific membership, roles and responsibilities, required source documentation, data format, details about review process, and procedures and timing of meetings will be addressed in a separate EAC charter.

Study Steering Committee

A Study Steering Committee (SSC) will provide direction and oversight to the study from an Investigator's perspective, including protocol creation and amendments, study execution, and evaluation of study results at the end of study. Investigators in this clinical study and other experts who are not otherwise involved in the study may serve on this committee.

8.3 Adverse Events and Serious Adverse Events

The definitions of an AE and a SAE are in Appendix 4.

8.3.1 Time Period and Frequency for Collecting Adverse Event and Serious Adverse Event Information

The AE reporting period for safety surveillance begins when the participant is initially included in the study (date of first signature of informed consent/date of first signature of first informed consent) and continues until the Safety Follow-up Visit.

Any SAE assessed as related to study intervention must be recorded and reported, as indicated in Appendix 4, whenever it occurs, irrespective of the time elapsed since the last administration of study intervention.

The method of recording, evaluating, and assessing causality of AEs (including SAEs) and the procedures for completing and transmitting SAE reports are in Appendix 4.

8.3.2 Method of Detecting Adverse Events and Serious Adverse Events

At each study visit, the participant will be queried on changes in his or her condition. During the reporting period, any unfavorable changes in the participant's condition will be recorded as AEs, regardless if reported by the participant or observed by the Investigator.

Complete, accurate and consistent data on all AEs experienced for the duration of the reporting period (defined below) will be reported on an ongoing basis in the appropriate section of the CRF. All SAEs and nonserious AESIs must be additionally documented and reported using the appropriate Report Form as specified in Appendix 4.

8.3.3 Follow-up of Adverse Events and Serious Adverse Events

AEs are recorded and assessed continuously throughout the study, as specified in Section 8.3.1 and are assessed for their outcome at the Safety Follow-up Visit. All SAEs ongoing at the Safety Follow-up Visit must be monitored and followed up by the Investigator until stabilization or until the outcome is known, unless the participant is documented as "lost to follow-up". Reasonable attempts to obtain this information must be made and documented. It is also the responsibility of the Investigator to ensure that any necessary additional therapeutic measures and follow-up procedures are performed. Further information on follow-up procedures is given in Appendix 4.

8.3.4 Regulatory Reporting Requirements for Serious Adverse Events

The Sponsor will send appropriate safety notifications to Health Authorities in accordance with applicable laws and regulations.

The Investigator must comply with any applicable site-specific requirements related to the reporting of SAEs (particularly deaths) involving study participants to the IEC/IRB that approved the study.

In accordance with International Council for Harmonisation (ICH) Good Clinical Practice (GCP), the Sponsor/designee will inform the Investigator of findings that could adversely affect the safety of participants, impact the conduct of the study or alter the IEC's/IRB's approval/favorable opinion to continue the study. In line with respective regulations, the Sponsor/designee will inform the Investigator of AEs that are both serious and unexpected and considered to be related to the administered product ("suspected unexpected serious adverse reactions" or SUSARs). The Investigator should place copies of Safety Reports in the Investigator Site File. National regulations regarding Safety Report notifications to Investigators will be considered.

When specifically required by regulations and guidelines, the Sponsor/designee will provide appropriate Safety Reports directly to the concerned lead IEC/IRB and will maintain records of these notifications. When direct reporting is not clearly defined by national or site-specific regulations, the Investigator will be responsible for promptly notifying the concerned IEC/IRB of any Safety Reports provided by the Sponsor/designee and of filing copies of all related correspondence in the Investigator Site File.

For studies covered by the European Directive 2001/20/EC, the Sponsor's responsibilities regarding the reporting of SAEs/SUSARs/Safety Issues will be carried out in accordance with that Directive and with the related Detailed Guidance documents.

8.3.5 Pregnancy

Only pregnancies the Investigator considers to be related to the study intervention (e.g., resulting from a drug interaction with a contraceptive method) are AEs. However, all pregnancies with an estimated conception date during the period defined in Section 8.3.1 must be recorded in the AE page/section of the CRF for both pregnancies in female participants and pregnancies in female partners of male participants. The Investigator must notify the Sponsor/designee in an expedited manner of any pregnancy using the Pregnancy Report Form, which must be transmitted by the same process specified for SAE reporting in Appendix 4, section on Reporting Serious Adverse Events and Adverse Events of Special Interest.

Investigators must actively follow up, document and report on the outcome of all these pregnancies, even if the participants are withdrawn from the study.

The Investigator must notify the Sponsor/designee of these outcomes using the Pregnancy Report Form. If an abnormal outcome occurs, the SAE Report Form will be used if the participant sustains an event and the Parent-Child/Fetus Adverse Event Report Form if the child/fetus sustains an event. Any abnormal outcome (e.g., spontaneous abortion, fetal death, stillbirth, congenital anomalies, ectopic pregnancy) must be reported in an expedited manner, as specified in Section 8.3.1, while normal outcomes must be reported within 45 days after delivery.

In the event of a pregnancy in a participant occurring during the study, the participant must be discontinued from study intervention. The Sponsor/designee must be notified without delay and the participant must be followed as indicated above. As a pregnant participant may have received teriflunomide and due to the prolonged half-life of teriflunomide, an accelerated elimination procedure must be performed and a teriflunomide level < 0.02 mg/L has to be reached (for more details see Appendix 8).

8.4 Treatment of Overdose

For this study, any dose of study intervention greater than the highest total daily dose included in the protocol or planned for a participant in the study within a 24-hour time period ± 6 hours will be considered an overdose.

The Sponsor does not recommend specific treatment for an overdose if not associated with any clinical signs/symptoms, but assistive/supportive measures, as necessary, should be provided. However, if there are clinical signs/symptoms then accelerated elimination procedure should be performed as described in Appendix 8.

Even if it not associated with an AE or a SAE, any overdose is recorded in the CRF and reported to drug safety in an expedited manner. Overdoses are reported on a SAE Report Form, following the procedure in Appendix 4, section on Reporting Serious Adverse Events and Adverse Events of Special Interest.

8.5 Pharmacokinetics

- The following PK parameters will be calculated, when appropriate:

Symbol	Definition
AUC	Area under the plasma concentration-time curve
$CL_{/f}$	The apparent total body clearance of study intervention following extravascular administration, taking into account the fraction of dose absorbed. $CL_{/f}$ = Dose/AUC_{τ}. Either the observed or predicted AUC_{τ} should be used, depending on the study specific requirements.
C_{max}	Maximum observed concentration
T_{max}	The time to reach the maximum observed concentration collected during a dosing interval (unless otherwise defined, take the 1st occurrence in case of multiple/identical C_{max} values)
$Vz_{/f}$	The apparent volume of distribution during the terminal phase following extravascular administration, based on the fraction of dose absorbed. $Vz_{/f}$ = Dose/($AUC_{0-\infty} \times \lambda z$) following single dose. $Vz_{/f}$ = Dose/($AUC_{\tau} \times \lambda z$) following multiple dose.

- For the main study, whole blood samples of approximately 2 mL will be collected for measurement of plasma concentrations of evobrutinib, as specified in the SoA. The actual date and time (24-hour clock time) of each sample will be recorded to calculate actual time elapsed since the prior dose administration.
- The exact date/time of sample collection and drug administration must be recorded in the eCRF and will be used in the calculation of PK parameters. Time deviations from planned PK sampling times will not be considered a protocol deviation provided the exact date/time of sample collection and drug administration are recorded in the eCRF.
- The quantification of evobrutinib in plasma will be performed using a validated assay method. Concentrations will be used to evaluate the PK of evobrutinib.
- Remaining samples collected for analyses of evobrutinib concentration may also be used to evaluate safety or efficacy aspects related to concerns arising during or after the study.
- Details on processes for collection and shipment of these samples are in the Laboratory Manual. Retention time and possible analyses of samples after the end of study are specified in the respective ICF.

- Evobrutinib concentration information that may unblind the study will not be reported to investigative sites or blinded personnel until the study has been unblinded. As samples from participants randomized to teriflunomide will not be analyzed for plasma concentrations of evobrutinib, measurements of evobrutinib concentration will be performed by unblinded analysts as detailed in Section 6.3.2.
- All samples collected for PK, as noted in the SoA, still within the known stability of the evobrutinib at the time of receipt by the bioanalytical laboratory will be analyzed.

See Section 9.4.3 for further details of the PK analysis.

8.6 Pharmacodynamics

Not applicable.

8.7 Pharmacogenetics

- Where local regulations and IRB/IEC allow, approximately 4 mL blood (total) sample will be collected for DNA analysis from consenting participants. Participation in pharmacogenetic research is optional. Participants who do not wish to participate in the pharmacogenetic research may still participate in the study.
- If not collected on Day 1 or a redraw is needed, the pharmacogenetic sample may be obtained at any other point of time during the study.
- In the event of DNA extraction failure, a replacement sample for pharmacogenetic testing may be requested from the participant. Additional informed consent will not be required to obtain a replacement sample.
- Appendix 6 provides further information on pharmacogenetic research.

8.8 Biomarkers

- Collection of participant samples for biomarker research is also part of this study and is governed by the appropriate ICF.
- The following participant samples for biomarker research are required and will be collected from all participants in this study, as specified in the SoA:
  - Blood samples will be tested for biomarkers of disease to evaluate disease activity or treatment response.
  - Blood samples will be tested for biomarkers of disease to evaluate the exposure-response relationship between PD biomarkers.
  - Blood samples will be tested for novel liver function protein biomarkers and genomic biomarkers to evaluate their activity with respect to predictivity and sensitivity of the novel biomarkers of hepatic function compared to traditional clinical chemistry endpoints.
  - Blood samples will be tested for gene expression/gene products to evaluate disease activity or treatment response.

- In addition, participant samples will be collected for analysis of biomarkers thought to play a role in pharmacokinetics, safety endpoints, drug response, and treatment efficacy including, specific candidate genes/genome-wide analysis for genetic variations.

- In addition, participant samples may be used for additional research, as specified in the ICF.
- Approximately 400 mL total blood will be collected in the main study and approximately 360 mL total blood will be collected in the optional OLE. Further details are provided in the ICF and Laboratory Manual.
- Details on processes for collection and shipment of these samples are in the Laboratory Manual. The Sponsor will store the samples in a secure storage space with adequate measures to protect confidentiality. Retention time and possible analyses of samples after the end of study are specified in the respective ICF.

8.8.1 Biomarkers of Disease

Blood samples will be collected from all participants to assess the relationship between candidate disease biomarkers (e.g., neurofilament light, cytokines) and disease activity or treatment response. Samples will be obtained predose as noted in the SoA. Biomarkers of disease may be measured at the analytical laboratory selected by the Sponsor using an appropriately validated bioanalytical method.

8.8.2 Novel Liver Function Protein Biomarkers and Novel Liver Function Genomic Biomarkers

Blood samples will be collected from all study participants to evaluate the relationship of novel protein and genomic biomarkers of hepatic function levels compared to standard clinical chemistry endpoints. Levels of novel biomarkers of hepatic function may be measured using various methods in samples collected predose from participants at time points noted in the SoA.

Sampling for novel liver function biomarkers is performed in parallel with liver function test (LFT) clinical chemistry monitoring. To enable analysis of a variety of biomarkers, samples are needed both prior as well as during the occurrence of elevated ALT. Samples will be collected from all participants, stored, and a subset analyzed retrospectively, as appropriate.

8.8.3 Gene Expression

Blood samples for gene expression analysis will be collected, except from participants in countries where collection of samples is not allowed, predose at time points noted in the SoA. These samples should be obtained after the assessment of vital signs in all participants (where allowed by local regulations). The actual date and time of each sample will be recorded. Samples may be analyzed by an analytical laboratory selected by the Sponsor using an appropriately validated bioanalytical method. The purpose of this analysis is to test whether potential differences in expression of specific genes may be linked to PK, safety endpoints, drug response, and treatment efficacy.

8.9 Health Resource Utilization

Health resource utilization data, associated with medical encounters, will be collected by the Treating Investigator and study-site personnel for all participants throughout the study. Protocol-mandated procedures, tests, and encounters are excluded.

The data collected may be used to conduct exploratory economic analyses and will include: number of unplanned doctor/home/emergency visits, number of hospitalizations, paid assistance, and number of missed work days.

8.10 Immunogenicity Assessments

Not applicable.

9 Statistical Considerations

9.1 Statistical Hypotheses

9.1.1 Statistical Hypotheses Related to Primary Objective

The primary endpoint, ARR based on 96 weeks of follow-up and qualified relapse events from participants enrolled in the present study, will be assessed for superiority via a 1-sided test of the null hypothesis H_0: rRR ≥ 1, where rRR denotes qualified relapse rate ratio comparing evobrutinib to teriflunomide. The alternative hypothesis is H_1: rRR < 1. The rRR effect measure will be estimated from a negative binomial (NB) model for qualified relapse count that adjusts for covariates based on stratification factors.

9.1.2 Statistical Hypotheses Related to Secondary Objectives

There are four efficacy secondary endpoints and two HRQoL secondary endpoints: time to first occurrence of 12-week CDP over 96 weeks (pooled), time to first occurrence of 24-week CDP over 96 weeks (pooled), Change from Baseline (CFB) in PROMIS PF score at 96 weeks (pooled), CFB in PROMIS Fatigue score at 96 weeks (pooled), total number of T1 Gd+ lesions based on assessments at Week 24, 48, and 96, and total number of new or enlarging T2 lesions based on assessments at Week 24, 48, and 96.

The secondary endpoint, time to 12-week CDP (pooled), will be assessed for superiority via a 1-sided stratified logrank test of the null hypothesis H_0: $S_e(t) \leq S_c(t)$, where $S_e(t)$ denotes the survival function for time to 12-week CDP in the experimental (evobrutinib) group, $S_c(t)$ denotes the survival function for time to 12-week CDP in the control (teriflunomide) group, and the variable t denotes time since randomization. The alternative hypothesis is H_1: $S_e(t) > S_c(t)$. Strata in the logrank test will be based on randomization strata and study ID. The secondary endpoint time to 24-week CDP (pooled) will be analyzed similarly, with a similar null hypothesis tested.

The secondary endpoint, CFB in PROMIS PF score at 96 weeks (pooled), will be assessed for superiority via a 1-sided test of the null hypothesis H_0: $\Delta_{PF} \leq 0$, where Δ_{PF} denotes difference in PROMIS PF score CFB at 96 weeks least-squares mean, comparing evobrutinib to teriflunomide, based on pooled data (higher score corresponds to improved physical function). The alternative hypothesis is H_1: $\Delta_{PF} > 0$. Covariates in the model used to model CFB will be based on randomization strata and study ID. The secondary endpoint CFB in PROMIS Fatigue score at 96 weeks (pooled) will be analyzed similarly, with a similar null hypothesis tested.

The secondary endpoint, total number of T1 Gd+ lesions based on scans at Weeks 24, 48, and 96, from participants enrolled in the present study, will be assessed for superiority via a 1-sided test of the null hypothesis H_0: lRR ≥ 1, where lRR denotes lesion rate ratio comparing evobrutinib to teriflunomide. The alternative hypothesis is H_1: lRR < 1. The lRR effect measure will be estimated from a NB model for total number of T1 Gd+ lesions that adjusts for covariates based on stratification factors. The secondary endpoint total number of active T2 lesions based on scans at Weeks 24, 48, and 96, will be analyzed similarly, with a similar null hypothesis tested.

9.2 Sample Size Determination

The sample size is estimated based on the primary endpoint, ARR over 96 weeks. The ARR over 96 weeks among participants receiving evobrutinib is estimated to be 0.13 (NB dispersion parameter ≈ 2.5), as compared with 0.23 (NB dispersion parameter ≈ 2.5) among participants receiving the control treatment, teriflunomide. This represents a reduction in qualified relapse rate of 43% due to evobrutinib relative to teriflunomide. A 1-sided test of superiority based on the qualified relapse rate ratio from a NB model for qualified relapse count, H_0: rRR ≥ 1.0, H_1: rRR < 1.0, requires an evaluable sample size of 372 participants per group, followed for 96 weeks, to provide 90% power, assuming a 1-sided significance level of 0.025 for the evobrutinib versus teriflunomide comparison. To account for a drop-out rate of 20% over 2 years, the enrollment per group is targeted to be 465.

The assumption for evobrutinib ARR is based on 48 week data from the Phase II study of evobrutinib in RMS, in which the highest dose, 75 mg twice daily administered under fasted conditions, resulted in an ARR of 0.11 (95% CI: 0.04, 0.25). Due to the wide CI, an estimate of 0.13 is assumed for evobrutinib in Phase III study powering. The assumption of 0.23 for teriflunomide ARR is based on recent Phase III studies (ASCLEPIOS I and II), wherein ARR ranged from 0.22 to 0.25 (Hauser 2019). The assumption of 2.5 for the dispersion parameter is derived from point and CI estimates for ARR from active treatment groups of recent Phase III studies and modeling of relapse data from the Phase II study of evobrutinib. If the distribution of relapse count is more disperse than assumed (i.e., dispersion ≈ 3.0 in both groups), the power provided by an enrolled sample size of 465 participants per group will be reduced from 90% to 87.5%.

It is assumed that an IA will be conducted on the basis of the secondary endpoint, time to 12-week CDP, based on unblinded pooled data from the present study and the second Phase III study, that a group sequential approach will be used to conduct this IA, and that the IA will be an opportunity for sample size re-estimation (SSR). In particular, the group sequential design has an IA triggered when approximately 35% to 45% of the expected 12-week CDP events in the pooled analysis are observed (as specified in the Interim Analysis Plan), an efficacy boundary derived from the γ-spending function with γ = -24, and a futility boundary derived from the γ-spending function with γ = -24 (i.e., essentially zero probability of boundary crossing). A total of 201 events (defined as 12-week CDP) from the pooled studies is required to provide 71.5% power to detect hazard ratio = 0.70; it is expected that these events will be achieved with 465 enrolled participants per group per study followed for 96 weeks, assuming a log rank test, 1-sided significance level of 0.025 (see Section 9.4.4.4), 2 year CDP rate in the teriflunomide group of 15% (i.e., teriflunomide hazard rate of 0.0813), and 2 year dropout rate of 20%.

The trial-wise and family-wise type-1 error will be preserved at the 0.025 1-sided level in the presence of multiple testing due to both multiple looks at the data and multiple endpoints (primary and secondary), as described in Section 9.4.4.4.

Sample size and power calculations for the ARR and 12-week CDP endpoints were performed using the East v6.5 package.

9.3 Populations for Analyses

The analysis populations are specified below. The final decision to exclude participants from any analysis population will be made during a blinded data review meeting prior to database lock and unblinding.

Analysis Set	Description
Screening (SCR)	All participants, who provided informed consent, regardless of the participant's randomization and study intervention status in the study.
Full Analysis Set (FAS)	All participants who were randomized to study intervention. Participants will be analyzed per the intervention group to which they were randomized (i.e., intention-to-treat principle).
Safety (SAF)	All participants, who were administered any dose of any study intervention. Participants will be analyzed per the actual study intervention they received.
PK	All participants who receive at least 1 dose of evobrutinib and have at least 1 quantifiable evobrutinib plasma concentration at a scheduled PK time point postdose without any important deviations or events that may impact the quality of the data or alter the evaluation of PK. Participants who receive active control will not be included.
Open Label Extension	All participants who receive at least 1 dose of evobrutinib during the OLE.

The following subgroups of the FAS will be considered for efficacy analyses:

- Sex, age, region, severity of disease, ethnic origin, prior treatment history (including but not limited to type, number and duration of prior treatments as well as reason for switch), EDSS.

9.4 Statistical Analyses

This section provides a description of the statistical methods to be used to analyze efficacy, safety, and other endpoints. Prior to locking the database, a detailed Integrated Analysis Plan (IAP) will be finalized.

Unless otherwise specified, the FAS will be the primary analysis set for all efficacy analyses, PRO analyses, and reporting of demographic and Baseline characteristics. The Safety analysis set will be used for all safety data reporting.

The secondary objectives based on disability progression and patient reported symptoms and functional status will be evaluated based on pooled data from the present study and the second Phase III study (i.e., MS200527_0080).

9.4.1 Efficacy Analyses

Endpoint	Statistical Analysis Methods
Primary	
ARR over 96 weeks	Primary analysis based on NB model of qualified relapse count over 96 weeks, with terms for intervention group and randomization strata; test based on adjusted relapse rate ratio (evobrutinib versus comparator) from the model. Estimation of treatment effect (and 95% 2-sided CI) based on adjusted relapse rate ratio from NB model. In the primary analysis of ARR, missing data assumed to be missing at random, with missing data status noninformative for qualified relapse.
Secondary efficacy and HRQoL	
Time to 12-week CDP	Analyzed for participants pooled from both Phase III studies. Primary analysis based on stratified logrank test of distribution of time to 12-week CDP with strata defined by randomization strata and study ID (i.e., Study MS200527_0080 or _0082), and with Cui, Hung, Wang (CHW) statistic for combining incremental Z statistics (pre- and post-IA). Estimation of treatment effect (and 95% 2-sided CI) based on hazard ratio from stratified Cox model of 12-week CDP hazard rate, with terms for intervention group and study ID, and strata defined by randomization strata. Cumulative distribution function for time to 12-week CDP will be estimated via Kaplan-Meier method by intervention group. In the primary analysis of time to 12-week CDP, censoring assumed to be noninformative for 12-week CDP.

Endpoint	Statistical Analysis Methods
Time to 24-week CDP	Analyzed for participants pooled from both Phase III studies. Primary analysis based on stratified logrank test of distribution of time to 24-week CDP with strata defined by randomization strata and study ID. Estimation of treatment effect (and 95% 2-sided CI) based on hazard ratio from stratified Cox model of 24-week CDP hazard rate, with terms for intervention group and study ID, and strata defined by randomization strata. Cumulative distribution function for time to 24-week CDP will be estimated via Kaplan-Meier method by intervention group. In the primary analysis of time to 24-week CDP, censoring assumed to be noninformative for 24-week CDP.
PROMIS physical function (PF) score CFB at Week 96, PROMIS Fatigue score CFB at Week 96	Analyzed for participants pooled from both Phase III studies. Primary analysis based on Mixed-Effect Model for Repeated Measures (MMRM) where score CFB is modeled, with terms for intervention group, visit, intervention group by visit interaction, Baseline score, randomization strata, and study ID. Test and treatment effect estimator (with 95% 2-sided CI) based on difference (evobrutinib versus comparator) in least squares means of Week 96 CFB from the model. Cumulative distribution function for Week 96 CFB will be estimated by intervention group. In the primary analysis of CFB, missing data assumed to be missing at random (MAR), with missing data status noninformative for CFB.
Total T1 Gd+ lesions over scans at Weeks 24, 48, and 96; Total new or enlarging T2 lesions over scans at Weeks 24, 48, and 96	Primary analysis based on NB model of total lesion count, with terms for intervention group, Baseline lesion activity, and randomization strata. Test and treatment effect estimator (with 95% 2-sided CI) based on adjusted lesion rate ratio (evobrutinib versus comparator) from model. In the primary analysis of lesion rate, missing data assumed to be MAR, with missing data status noninformative for lesion count.
Tertiary/Exploratory	
ARR at 48 weeks	Analysis follows that of ARR over 96 weeks.
Time to qualified relapse;	Stratified logrank test of distribution of time to event with strata defined by randomization strata. Estimation of treatment effect (and 95% 2-sided CI) based on hazard ratio from stratified Cox model of event hazard rate. Cumulative distribution function will be estimated via Kaplan-Meier method by intervention group. Censoring assumed to be noninformative for qualified relapse.
Time to 12-week confirmed disability based on composite score; Time to ≥ 20% increase (confirmed at 12 weeks) in T25-FW; Time to ≥ 20% increase (confirmed at 12 weeks) in 9-HPT. Time to first occurrence of 12 week confirmed PF deterioration	Analyzed for participants pooled from both Phase III studies. Primary analysis based on stratified logrank test of distribution of time to event with strata defined by randomization strata and study ID. Estimation of treatment effect (and 95% 2-sided CI) based on hazard ratio from stratified Cox model, with terms for intervention group and study ID, and strata defined by randomization strata. Cumulative distribution function will be estimated via Kaplan-Meier method by intervention group. Censoring assumed to be noninformative for event of interest.
Relapse-free status at Week 96; Disease activity free status at Week 48 or Week 96; NEPAD at Week 48 or Week 96; new Gd+ T1 lesion free status at Week 96; new or enlarging T2 lesion free status at Week 96; CUA lesion-free status at Week 96	Proportions within intervention groups compared using Cochran-Mantel-Haenszel (CMH) $\chi 2$ test stratified by randomization strata.
12-week (or 24-week) confirmed EDSS progression free status at Week 96;	Analyzed for participants pooled from both Phase III studies. Proportions within intervention groups compared using Cochran-Mantel-Haenszel (CMH) $\chi 2$ test stratified by randomization strata and study ID.

Endpoint	Statistical Analysis Methods
Total new T1 hypo-intense lesions over scans at Weeks 24, 48, and 96; Total CUA lesions over scans at Weeks 24, 48, and 96.	NB model of total lesion count, with terms for intervention group, Baseline lesion activity, and randomization strata. Estimation of treatment effect (and 95% 2-sided CI) based on adjusted lesion rate ratio from NB model. Missing data assumed to be MAR, with missing data status noninformative for lesion count.
Mean number of T1 Gd+ lesions per scan based on assessments up to Week 96; Mean number of new or enlarging T2 lesions per scan; Mean number of CUA lesions per scan.	Hodges-Lehman estimate (and 95% 2-sided CI) of shift in lesion count distribution between intervention groups and Wilcoxon rank sum test, stratified according to randomization strata and Baseline lesion activity.
Change in volume of T1 Gd+ lesions from Baseline to Week 96; Change in volume of T2 lesions from Baseline to Week 96; Change in normalized T1 intensity within pre-existing nonenhancing T2 weighted lesion volume from Baseline to Week 96.	MMRM analysis where CFB is modeled, with terms for intervention group, visit, intervention group by visit interaction, randomization strata, and Baseline value of endpoint. Estimation of treatment effect (and 95% 2-sided CI) based on adjusted difference in least-squares means from the model. Missing data assumed to be MAR, with missing data status noninformative for CFB.
Volume of slowly evolving lesions (SELs) at Weeks 24, 48, and 96	For each timepoint, Hodges-Lehman estimate (and 95% 2-sided CI) of shift in distribution between intervention groups and Wilcoxon rank sum test, stratified according to randomization strata and Baseline T2 lesion volume tertiles.
Percentage change in BV from Week 24 to Week 96; Percentage change in Thalamic volume from Week 24 to Week 96; Percentage change in cortical grey matter volume from Week 24 to Week 96.	MMRM analysis where percentage change from Week 24 is modeled, with terms for intervention group, visit, intervention group by visit interaction, Week 24 value of endpoint, and randomization strata. Estimation of treatment effect (and 95% 2-sided CI) based on adjusted difference in least-squares means from the model. Missing data assumed to be MAR, with missing data status noninformative for percentage change from Week 24.
CFB in PROMIS physical function score at Week 48; CFB in PROMIS Fatigue score at Week 48	Analyzed for participants pooled from both Phase III studies. MMRM analysis of CFB at Week 48 follows that of CFB at Week 96.
CFB in SDMT score at Week 48 and Week 96; CFB in SF-36v2 score at Week 48 and Week 96; CFB in EQ-5D-5L score at Week 48 and Week 96	Analyzed for participants pooled from both Phase III studies. MMRM analysis where change from Baseline is modeled, with terms for intervention group, visit, intervention group by visit interaction, Baseline value of endpoint, randomization strata, and study ID. Estimation of treatment effect (and 95% 2-sided CI) based on adjusted difference in least-squares means from the model. Missing data assumed to be MAR, with missing data status noninformative for CFB.
12-week and 24-week confirmed disability improvement status	Analyzed for the subgroup of participants from the pooled Phase III studies with Baseline EDSS score ≥ 2.0. Participants without confirmed disability improvement counted as not improved, independent of follow-up time. Cochran-Mantel-Haenszel (CMH) test stratified according to randomization strata and study ID.
Absolute values of HRU endpoints at study visit over 96 weeks, including but not limited to doctor/home/ emergency visits, hospitalizations, paid assistance, and missed work days	Analyzed for participants pooled from both Phase III studies. Descriptive statistics (mean, SD, median, minimum/maximum, 25^{th} and 75^{th} percentile; 95% CI) at the various time points.

9-HPT = 9-Hole Peg Test, ARR = Annualized Relapse Rate, BV = brain volume, CDP = Confirmed disability progression, CFB = change from Baseline, CHW = Cui, Hung, Wang, CMH = Cochran-Mantel-Haenszel, CI = confidence interval, CUA = combined unique active, EDSS = Expanded Disability Status Scale, EQ-5D-5L = EuroQoL 5 Dimension 5 Levels, Gd = gadolinium, HRQoL = health-related quality of life, HRU = Health Resource Utilization, MAR = Missing at Random, MMRM = Mixed Effect model for repeated measures, NB = negative binomial, NEPAD = No evidence of progression or active disease, PROMIS = Patient Reported Outcomes Measurement Information System, SD = standard deviation, SDMT = Symbol Digit Modalities Test, SEL = slowly evolving lesions, SF-36v2 = 36-Item Short Form Survey Instrument Version 2, T25-FW = Timed 25-Foot Walk.

9.4.1.1 Efficacy Analyses Related to Primary Objective

Per ICH E9(R1), Addendum on Estimands and Sensitivity Analysis in Clinical Trials (November 2019), the primary estimand targeting the primary objective in this study is defined by the following attributes:

- **Variable (endpoint):** The primary endpoint is ARR over 96 weeks. Measurements on the participant-level are number of qualified relapses experienced by the participant and follow-up time over which those relapses occurred (potentially less than 96 weeks).
- **Treatment:** The intervention of interest is evobrutinib 45 mg twice daily for 96 weeks and the alternative intervention is teriflunomide 14 mg once daily for 96 weeks.
- **Population:** The population of participants targeted by the clinical question is defined by the inclusion/exclusion criteria.
- **Strategies for handling intercurrent events:** The main intercurrent event envisaged is discontinuation of assigned treatment. The strategy for dealing with this is Treatment Policy, which requires that the qualified relapse count over 96 weeks post randomization be used, if available, regardless of treatment discontinuation. The Compositive Variable strategy is used to handle the intercurrent event of death attributable to MS or treatment. The While Alive strategy is used to handle the intercurrent event of death unattributable to MS or treatment.
- **Population-level summary:** The population-level summary comparing the intervention groups is relapse rate ratio, based on a NB model for qualified relapse count, with terms for intervention group and randomization strata, with offset equal to log follow-up time (in years) over which the qualified relapses experienced by a participant are observed.

In accordance with Treatment Policy strategy, for participants discontinuing treatment, any qualified relapses occurring through Safety Follow-up will be included in the analysis up to 96 weeks post randomization. For participants who have not been followed to 96 weeks post randomization, the missing data will be assumed to be missing at random (MAR); treatment discontinuers will be assumed to experience relapse through 96 weeks post randomization at the same rate as treatment completers within the same intervention group and stratum. This is the assumption underlying the rate ratio estimator implemented by the NB model that uses observed relapse events and observed follow-up time for each participant, regardless of treatment completer/discontinuer status.

In accordance with the Composite Variable strategy, for participants experiencing death attributable to MS or treatment, the death will be counted as a qualified relapse. In accordance

with the While Alive strategy, for participants experiencing death unattributable to MS or treatment, the death will not be counted as a qualified relapse. The adjusted relapse rate ratio (RR) comparing evobrutinib to teriflunomide will be estimated from the NB model. The adjusted RR, 95% 2-sided CI, and 1-sided p-value will be reported, together with adjusted ARR (and 95% CI) for each intervention group from the model, and unadjusted ARR (and 95% CI) for each intervention group estimated nonparametrically. The analysis of ARR will be based on data from this study alone.

Unadjusted ARR over 96 weeks is calculated by dividing the total number of qualified relapse events experienced by all participants in a given group through Week 96 by the person-time (in years) observed for those participants through Week 96.

The covariates defined by randomization strata are region (4 levels: North America, Western Europe, Eastern Europe, Rest of World) and Baseline (Day 1) EDSS (2 levels: < 4.0, ≥ 4.0). The NB model assumes a common dispersion parameter for all participants, independent of intervention group or Baseline covariates. The adjusted relapse rate ratio estimate is given by exponentiation of the estimate for the treatment coefficient from the NB model. In the PA, only observed events and observation time will be included in the analysis; there will be no imputation of events for participants discontinuing study early.

Sensitivity analyses of the primary estimand for the primary endpoint are specified in Section 9.4.1.3.

The IAP will further specify rates, timings, and potential roles of intercurrent events, strategies for handling them, and potential consequences on the analyses and results.

9.4.1.2 Efficacy Analyses Related to Secondary Objectives

The hypotheses tested in support of secondary efficacy objectives are described, and the inheritance of alpha graphically depicted, in Section 9.4.4.4.

The primary estimands targeting the secondary objectives in this study share some of the attributes of the primary estimand in Section 9.4.1.1, such as the treatment comparison attribute, the population attribute, and the strategy for dealing with the intercurrent events of treatment discontinuation. Distinctive attributes of each estimand are described below for each secondary endpoint.

For the secondary efficacy endpoint, time to 12-week CDP (clinical disability progression based on EDSS), the Treatment Policy strategy requires that all EDSS data, up to 96 weeks post randomization, will be used if available, regardless of treatment discontinuation. In accordance with Treatment Policy strategy, for participants discontinuing treatment, all EDSS data through Safety Follow-up will be included in the analysis up to 96 weeks post randomization. The population-level summary comparing the intervention groups is 12-week CDP hazard ratio, based on a stratified Cox model for 12-week CDP hazard rate, with terms for intervention group and study ID, strata defined by randomization strata, and where the data are pooled from the FAS of the present study and the FAS of the second Phase III study.

Participants who did not experience 12-week CDP by 96 weeks post randomization, by time of early study discontinuation, or before being lost to follow up, will be censored at the date of the last EDSS assessment during the 96 weeks post randomization. Censoring will be assumed to be noninformative for 12-week CDP, conditional on intervention group, stratum, and study ID. This is the assumption underlying the hazard ratio estimator implemented by the stratified Cox model that uses observed follow-up time for each participant, regardless of treatment completer/discontinuer status, where participants are censored for 12-week CDP at the latest EDSS assessment. For participants assigned to an additional 12 weeks of treatment, for a total of 108 weeks of treatment, confirmation (or lack thereof) at Week 108 will be included in the analysis.

In accordance with the Composite Variable strategy, for participants experiencing death attributable to MS or treatment, the death will be counted as 12-week CDP. In accordance with the While Alive strategy, for participants experiencing death unattributable to MS or treatment, the death will not be counted as 12-week CDP.

An unblinded IA will be performed for the purpose of SSR based on the secondary efficacy endpoint, time to 12-week CDP, pooled across the present study and the second Phase III study (see Section 9.4.4.1). At the time of the IA, the increase in sample size, up to a maximum of 35%, will be based on the CP of the stratified logrank test at the time of the IA, according to a rule specifying required sample size as a decreasing step function of CP in the promising zone, $0.3 \leq CP < 0.8$, as prespecified in the IDMC Charter or related documentation. At the time of the PA of 12-week CDP, the Cui, Hung, Wang (CHW) approach will be used to combine the incremental Z statistics from the IA and the PA (Cui 1999). Each incremental Z statistic will compare evobrutinib and teriflunomide on the basis of pooled 12-week CDP data via a stratified logrank test, with strata defined by randomization strata and study ID (i.e., Study MS200527_0080 or _0082).

The PA of time to 12-week CDP will report the hazard ratio comparing evobrutinib to teriflunomide estimated via a stratified Cox model, 95% 2-sided CI, and 1-sided p-value for the CHW test, together with a Kaplan-Meier estimate of cumulative probability of experiencing 12-week CDP over time for each intervention group. Prior to pooling, the validity of pooling data across the present study and the second Phase III study will be assessed by reviewing consistency of demographics, Baseline characteristics, ARR and 12-week CDP results.

The secondary endpoint, time to 24-week CDP, will be analyzed only at the time of the PA, so a conventional stratified logrank test will be used, not the CHW test. In all other respects, the analysis of the 24-week CDP endpoint will be the same as that of the 12-week CDP endpoint. The primary estimand attributes for the time to 24-week CDP endpoint are the same as the primary estimand attributes for the time to 12-week CDP endpoint.

For the secondary efficacy endpoint, CFB in PROMIS PF score at Week 96, the Treatment Policy strategy requires that the CFB in PROMIS PF score at 96 weeks post randomization be used if available, regardless of treatment discontinuation. In accordance with Treatment Policy strategy, for participants discontinuing treatment, all PROMIS PF data through Safety Follow-up will be included in the analysis. The population-level summary comparing the

intervention groups is difference in least-squares means, based on a MMRM for CFB, where the model includes terms for intervention group, visit, intervention group by visit interaction, Baseline score, randomization strata, and study ID, and where the data are pooled from the FAS of the present study and the FAS of the second Phase III study.

For participants with missing PROMIS PF data, the missing data will be assumed to be MAR; such participants will be assumed to have the same mean PROMIS PF score CFB trajectory through 96 weeks as participants with available data within the same intervention group and stratum, and having the same Baseline score and study ID. This is the assumption underlying the estimator of difference of LS means at Week 96 implemented by the MMRM that uses observed score CFB data from each participant, regardless of treatment completer/discontinuer status.

In accordance with the While Alive strategy, for participants experiencing death, all PROMIS PF data up to the time of death will be used in the analysis.

The difference (comparing evobrutinib and teriflunomide) in least-squares mean CFB, 95% 2-sided CI, and 1-sided p-value will be reported. For each intervention group, the adjusted least-squares mean Week 96 score CFB and associated 95% 2-sided CI will be reported, as will the estimated cumulative distribution function for Week 96 score CFB.

The secondary endpoint, CFB in PROMIS Fatigue score at Week 96, will be analyzed in the same manner as CFB in PROMIS PF score at Week 96. The primary estimand attributes for these two endpoints are the same.

For the secondary efficacy endpoint, total number of T1 Gd+ lesions over scans at Weeks 24, 48, 96, the Treatment Policy strategy requires that the lesion count over scans at 24, 48, and 96 weeks post randomization be used, if available, regardless of treatment discontinuation. In accordance with Treatment Policy strategy, for participants discontinuing treatment, all lesion data through Safety Follow-up will be included in the analysis. The population-level summary comparing the intervention groups is lesion rate ratio, based on a NB model for lesion count, with terms for intervention group, randomization strata, and Baseline lesion activity, with offset equal to log number of scans over which the lesions experienced by a participant are observed.

For participants missing one or more scans at Weeks 24, 48, 96, the missing data will be assumed to be MAR; such participants will be assumed to experience lesions through 96 weeks at the same rate as participants with available data within the same intervention group and stratum, and having the same Baseline lesion activity. This is the assumption underlying the rate ratio estimator implemented by the NB model that uses observed lesion count and number of scans giving rise to that count for each participant, regardless of treatment completer/discontinuer status.

In accordance with the While Alive strategy, for participants experiencing death, all T1 Gd+ lesion data up to the time of death will be used in the analysis.

The adjusted lesion RR comparing evobrutinib to teriflunomide estimated from the NB model, 95% 2-sided CI, and 1-sided p-value will be reported, together with adjusted lesion rate for each intervention group and associated 95% 2-sided CI. The analysis of total number of T1 Gd+ lesions will be based on data from the present study only.

The secondary endpoint, total number of new or enlarging T2 lesions over scans at Weeks 24, 48, 96, will be analyzed in the same manner as total number of T1 Gd+ lesions over scans at Weeks 24, 48, and 96. The primary estimand attributes for these two endpoints are the same.

Sensitivity analyses of the primary estimands of the secondary endpoints are specified in Section 9.4.1.3.

9.4.1.3 Sensitivity Analyses

For the primary estimand of the primary endpoint, ARR over 96 weeks, it is important to assess the sensitivity of results to the MAR assumption that treatment discontinuers experience relapse through 96 weeks at the same rate as treatment completers within the same intervention group and stratum. The following sensitivity analyses for the primary endpoint ARR over 96 weeks will be detailed in the IAP:

1. Analysis to evaluate the potential influence of informative treatment discontinuation according to discontinuation reason. If a participant's discontinuation is unrelated to treatment (i.e., unrelated to efficacy or safety issues), the participant's relapse count will be multiply imputed using ARR in comparator group, for the interval between discontinuation and 96 weeks post randomization. If a participant's discontinuation is related to treatment, relapse count will be multiply imputed at a rate higher than ARR in comparator group, for the interval between discontinuation and 96 weeks post randomization.

2. Analysis in which participants who discontinued treatment early during the 96-week Treatment Period without qualified relapse in the 30 days prior to discontinuation, are assumed to have a qualified relapse event at the date of discontinuation.

3. Analysis in which additional covariates are included in the NB model, such as number of relapses occurring within 2 years prior to study entry, Baseline presence/absence of T1 Gd+ lesions, prior MS treatment, and age, to evaluate the potential impact of model misspecification.

For the primary estimand of the secondary endpoint, time to 12-week CDP, it is important to assess the sensitivity of results to the assumption that censoring is noninformative for 12-week CDP, conditional on intervention group, study ID, and stratum. The following sensitivity analyses for the secondary efficacy endpoint time to 12-week CDP will be detailed in the IAP:

1. Analysis to evaluate the potential influence of informative treatment discontinuation, according to discontinuation reason or occurrence of initial progression event at last assessment. If a participant's discontinuation is unrelated to treatment (i.e., unrelated to efficacy or safety issues) and no initial progression event occurred, the participant's time

to CDP will be multiply imputed using hazard rate in comparator group, for the interval between discontinuation and 96 weeks post randomization. If a participant's discontinuation is related to treatment or an initial unconfirmed progression event occurred, the participant's time to CDP will be multiply imputed using hazard rate higher than that in comparator group, for the interval between discontinuation and 96 weeks post randomization.

2. Analysis in which participants who discontinued treatment early during the 96-week Treatment Period after having an initial progression event, but prior to 12-week confirmation, are assumed to have a 12-week CDP event at the date of initial progression.

3. Analysis with additional covariates in the stratified Cox model, such as number of relapses occurring within 2 years prior to study entry, Baseline presence/absence of T1 Gd+ lesions, prior MS treatment, and age.

Sensitivity analyses for time to 24-week CDP will be similar.

For the primary estimand of the secondary endpoint, PROMIS PF CFB, it is important to assess the sensitivity of results to the MAR assumption that participants with missing data have the same mean CFB trajectory through 96 weeks as participants with complete data, conditional on intervention group, Baseline score, study ID, and stratum. The following sensitivity analysis for the secondary endpoint PROMIS PF score CFB at Week 96 will be detailed in the IAP:

1. Analysis to evaluate the potential influence of informative treatment discontinuation according to discontinuation reason. If a participant's discontinuation is unrelated to treatment (i.e., unrelated to efficacy or safety issues), the participant's score CFB at each timepoint will be multiply imputed using CFB distribution at that timepoint in comparator group, for the interval between discontinuation and 96 weeks post randomization. If a participant's discontinuation is related to treatment, the participant's score CFB at each timepoint will be multiply imputed using a worse CFB distribution at that timepoint than that in comparator group, for the interval between discontinuation and 96 weeks post randomization.

Sensitivity analyses for PROMIS Fatigue score CFB at Week 96 will be similar.

For the primary estimand of the secondary endpoint, total number of T1 Gd+ lesions over scans at Weeks 24, 48, 96, it is important to assess the sensitivity of results to the MAR assumption that participants with missing data experience lesions through 96 weeks at the same rate as participants with complete data, conditional on intervention group, Baseline lesion activity, and stratum. The following sensitivity analysis for the secondary efficacy endpoint total T1 Gd+ lesions will be detailed in the IAP:

1. Analysis to evaluate the potential influence of informative treatment discontinuation according to discontinuation reason. If a participant's discontinuation is unrelated to treatment (i.e., unrelated to efficacy or safety issues), the participant's lesion count will be multiply imputed using rate in comparator group, for the interval between

discontinuation and 96 weeks post randomization. If a participant's discontinuation is related to treatment, lesion count will be multiply imputed at a rate higher than that in comparator group, for the interval between discontinuation and 96 weeks post randomization.

The following supplemental analyses for the secondary efficacy endpoint total number of T1 Gd+ lesions over scans at Weeks 24, 48, 96 will be detailed in the IAP:

1. Analysis wherein lesion count at a single scan is modeled as Poisson-distributed, with a random effect for participant, and terms for intervention group, randomization strata, and Baseline lesion activity. This model will compare treatment on the basis of conditional adjusted lesion rate ratio, where conditioning is on the random effect.

2. Analysis wherein total lesion count is compared between groups on the basis of the shift in location of the lesion count distribution (Hodges-Lehmann estimate) and Wilcoxon rank sum test, stratified according to randomization strata and Baseline lesion activity.

Sensitivity and supplemental analyses for total number of new/enlarging T2 lesions over scans at Weeks 24, 48, 96 will be similar.

9.4.2 Safety Analyses

All safety analyses will be performed on the Safety Analysis population.

Endpoint	Statistical Analysis Methods
Primary	Not applicable.
Secondary	
Adverse Events	Descriptive statistics, AESI summaries, 3-Tier AE summaries
Clinical Laboratory Test Values	Descriptive statistics, shift tables, boxplots, individual participant line plots, Kaplan-Meier analyses of time to event, Evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) figures
ECG Parameters	Descriptive statistics
Vital Signs	Descriptive statistics
Immunoglobulin Levels	Descriptive statistics
Tertiary/Exploratory	Not applicable

9.4.2.1 Adverse Events

Adverse events will be coded using the MedDRA. All TEAEs will be summarized by intervention group. Treatment-emergent AEs are defined as AEs that occurred or worsened on or after the first dose of study intervention. The number and percentage of participants who experienced at least 1 TEAE will be summarized by SOC and preferred term. The percentage will be based on the number of participants in each intervention group. Treatment-emergent AEs will also be summarized by relationship to intervention and by severity within each

intervention group. Deaths, SAEs, AESIs, and AEs leading to study discontinuation will be tabulated and presented in data listings. Participant level data listings of all AEs will be presented.

Summary and analysis of AEs will be performed based on the 3-tier approach (Crowe 2009) as further detailed in the IAP. Tier 1 AEs and AESIs will be predefined in the IAP.

9.4.2.2 Clinical Laboratory Test Values

Clinical laboratory results (chemistry, hematology, and urinalysis) will be summarized using descriptive statistics for each visit by intervention group. Observed values at each visit and changes from Baseline to each postbaseline visit will be presented. For clinical laboratory parameters with associated normal ranges, number and percentage of participants having high/low/normal findings for worst on-treatment laboratory value will be summarized by intervention group; shift tables will be used to summarize changes from Baseline finding to worst on-treatment finding. For clinical laboratory parameters with National Cancer Institute (NCI)-CTCAE grades, shift tables will be used to summarize changes from Baseline grade to worst on-treatment grade. The distribution of selected laboratory parameters by time point and intervention group will be displayed via boxplots. All laboratory data will be provided in participant data listings.

Analyses of liver enzyme tests will include Kaplan-Meier estimates of time to ALT or AST events, plots supporting evaluation of Drug-Induced Serious Hepatotoxicity (eDISH), and individual participant profiles.

9.4.2.3 Vital Signs

Observed values at each visit and changes from Baseline to each postbaseline visit in vital signs (blood pressure, pulse rate, respiratory rate, and temperature) will be summarized by time point and intervention group using descriptive statistics. Similar summaries of descriptive statistics will be provided for the vital signs collected before and after the first dose of study intervention. Out-of-range values of vital signs will be tabulated as appropriate. All vital signs will be provided in participant data listings.

9.4.2.4 Electrocardiogram Parameters

Observed values at each visit and change from Baseline to Week 96 in ECG parameters (e.g., PR, HR, QRS, RR, QT, and QTcF) will be summarized by intervention group using descriptive statistics. QTc will be reported based on Fridericia's method. Percentage and counts of participants with normal and abnormal ECG findings will be summarized by intervention group. Out-of-range values of ECG parameters will be tabulated as appropriate. All ECG data will be provided in participant data listings.

For all ECG parameters, the results for categorical analysis will be summarized by intervention group and visit (or time point) in frequency tables with counts and percentages of participants.

Categories will cover absolute values and changes from Baseline (predose Day 1):

HR:

- Absolute < 50 bpm, < 40 bpm, < 30 bpm
- Change from Baseline > 20 bpm, > 30 bpm, > 40 bpm

PR:

- Absolute > 200 msec and > 220 msec
- Change from Baseline > 30 msec

QRS:

- Absolute > 110 msec

QTcF:

- Absolute > 450 msec, > 480 msec, and > 500 msec
- Change from Baseline > 30 msec and > 60 msec

Electrocardiogram parameters will be summarized using descriptive statistics for continuous variables such as QTcF intervals, and frequency counts and percentages for categorical variables.

9.4.2.5 Immunoglobulin Levels

Data on Ig levels (observed values, change, and percent change from Baseline, with Baseline defined as the Day 1 sample) will be descriptively summarized in tabular and/or graphic format, as appropriate.

Correlative analyses may be explored and reported separately.

9.4.2.6 Concomitant Medication and Procedures

Prior and concomitant medications will each be categorized by therapeutic class and preferred term using WHO Drug coding dictionary. The number and percent of participants using each prior and concomitant medication will be summarized by therapeutic class and preferred drug name for each intervention group. Participants who reported more than 1 medication for a particular preferred term will be counted once for each preferred term and therapeutic class.

Concomitant procedures will be categorized by SOC and preferred term using MedDRA. The number of and percent of participants experiencing each prior and concomitant procedure will be summarized by type of procedure for each intervention group.

9.4.2.7 Columbia-Suicide Severity Rating Scale

Results of the C-SSRS will be listed for each visit by participant.

9.4.3 Other Analyses

PK and biomarker exploratory analyses will be specified in the IAP finalized before database lock. Integrated analyses across studies, such as the population PK analysis will be presented separately from the main clinical study report (CSR).

9.4.3.1 Pharmacokinetic Parameters and Biomarkers

In general, biomarkers assessed at planned visits and premature early discontinuation from treatment will be summarized in a manner similar to safety laboratory parameters.

The PK data of the current study will be used for the development of population PK models (possibly in combination with the corresponding data of earlier studies, e.g. MS200527_0019, MS200527_0017, MS200527_0086, and/or others), to describe the concentration-time profiles for the participants in the study. Population PK model-based exposure metrics such as C_{max} and AUC (from 0 to 24 hour) at steady state will be derived.

The derived PK profiles and exposure metrics will be used to develop exposure response models (longitudinal or cross-sectional) for MRI, efficacy and safety endpoints. In particular, a cross-sectional model of ARR as a function of steady state AUC, a longitudinal repeated time to event (qualified relapse) model, a longitudinal model for EDSS, cross-sectional or longitudinal exposure response models for MRI lesions (T1 Gd+ lesion and new or enlarging T2 lesion count), and a longitudinal (if possible) exposure response model for ALT elevations, will be developed.

Other endpoints of interest depending on the signals that will be detected based on the exploratory/inferential statistical analysis of the study data may also be included in the exposure response analyses.

Full details of the planned population PK modeling and other exposure related modeling will be described in the study IAP. The results of the corresponding analyses will be reported separately from the study CSR.

9.4.3.2 Demographics, Baseline Characteristics, Disposition, and Compliance

Participant demographics, such as age, sex, race, will be summarized by intervention group using descriptive statistics. Baseline disease characteristics (including MS history and MRI characteristics), such as Baseline EDSS, number of relapses in the 1 year and 2 years prior to randomization, time (in years) since onset of MS symptoms, time (in years) since MS diagnosis, Baseline number of T1 Gd+ lesions, Baseline number of T2 lesions, and Baseline T2 lesion volume, prior MS study treatment will also be summarized.

Disposition of participants (i.e., discontinuation from treatment by reason, discontinuation from study by reason) and compliance of participants to intervention will be summarized by intervention group using descriptive statistics.

9.4.3.3 Patient Reported Outcome Analyses

The endpoint CFB in PRO score at a given time point (Week 48 or Week 96), will be compared between evobrutinib and teriflunomide via a MMRM based on pooled data, adjusted for Baseline PRO score and covariates defined by randomization strata and study ID. The adjusted difference in least-squares mean change from Baseline at a given time point (Week 48 or Week 96), comparing evobrutinib to teriflunomide, 2-sided 95% CI, and 1-sided p-value will be reported. A similar analysis will be performed for all PRO score change from Baseline endpoints. The exploratory endpoint, time to deterioration in PROMIS PF score confirmed at 12 weeks will be analyzed in a manner similar to that of time to 24-week CDP.

9.4.3.4 Health Resource Utilization

For the exploratory endpoint HRU observed values at each visit and CFB to each postbaseline visit in HRU (doctor/home/emergency visits, hospitalizations, paid assistance, and missed work) will be summarized by time point and intervention group using descriptive statistics (mean, standard deviation, median, minimum/maximum, 25th and 75th percentile; 95% CI).

9.4.3.5 Analysis of Open Label Extension Period Endpoints

Efficacy and HRQoL data collected during the 144-week OLE Period will be summarized. Details will be provided in the IAP.

Safety data collected during the 144-week OLE Period will be analyzed as described in Section 9.4.2.

9.4.4 Sequence of Analyses

Three analyses are planned for the study: (1) an unblinded IA for SSR based on the secondary endpoint 12-week CDP, with data pooled across the present study and the second Phase III study, performed by the IDMC and the independent Data Analysis Center (DAC), triggered when approximately 35% to 45% of the expected number of 12-week CDP events required to power this secondary endpoint have been observed, as specified in the Interim Analysis Plan, (2) a PA, performed by the Sponsor, with timing and endpoint evaluation as described in Section 9.4.4.2, and (3) a final analysis, performed by the Sponsor, triggered when 100% of participants enrolled in the OLE complete the OLE, or discontinue prematurely from the OLE. The IA will not be performed for futility.

If the sample size is not increased at the time of the IA, and the PA of the present study occurs prior to the PA of the second Phase III study, there will be an analysis of endpoints based on data pooled from the two studies at the time of the PA of the second Phase III study.

In addition, analyses will be performed at regular intervals for the purpose of safety monitoring by the IDMC, as described in the IDMC Charter.

Prior to the IA, a detailed Interim Analysis Plan will be provided, specifying the timing of the IA, the endpoint that will be unblinded and analyzed, boundary of the promising zone, decision rule, alpha spending, and sample size adjustment. Details of the primary and final analyses will be described in the IAP. Details of the IDMC analyses, including unblinded sample size re-estimation, will be described in the IDMC Statistical Analysis Plan and the IDMC Charter and related documentation.

9.4.4.1 Unblinded Interim Analysis for Sample Size Re-estimation

When approximately 35% to 45% of the 12-week CDP (as measured by EDSS) events (i.e., 71 to 91 events out of a total of 201) required to power the 12-week CDP endpoint, based on pooled data from the present study and the second Phase III study have been observed, as specified in the Interim Analysis Plan, an IA for unblinded sample size re-estimation based on the secondary endpoint 12-week CDP will be performed. This pooled IA will be conducted by the IDMC with the assistance of an independent DAC. Only data relevant to 12-week CDP must be cleaned for the IA. The SSR will be based on determining whether CP falls in the promising zone, defined by $0.3 \leq CP < 0.8$, as prespecified in the IDMC Charter or related documentation. Based on an efficacy boundary derived from the γ-spending function, $\gamma = -24$, it is expected that 1-sided alpha of less than 1.0E-6 (i.e., essentially zero) will be spent on this IA. Based on a futility boundary derived from the γ-spending function, $\gamma = -24$, there is essentially zero spend of beta (a test of futility does not result in a spend of alpha).

The independent DAC will share with the IDMC the CP estimate and the result of the prespecified map from CP to sample size increase recommendation. If the recommendation is to increase the sample size, the independent DAC will also provide the IDMC with the recalculated number of 12-week CDP (as measured by EDSS) events from the pooled studies required to achieve the target power. The IDMC will share their recommendation to continue with the planned enrollment, or increase enrollment up to a maximum of 35% additional participants (i.e., up to 628 per group per study) as described in the IDMC Charter or related documentation, with a Firewall team from the Sponsor, with whom the final decision rests. The Firewall team will follow a charter to be finalized prior to the IA. All personnel involved in the conduct of the study will remain blinded to the result of the IA.

9.4.4.2 Primary Analysis

If the sample size is not increased when the IA is performed, the PA of the present study will be triggered when the last participant reaches 96 weeks of treatment or discontinues from treatment prematurely during the blinded Treatment Period, and completes Safety Follow-up, or discontinues from the study prematurely during the blinded Treatment Period. Given that the primary analyses of this study and the second Phase III study are unlikely to coincide, endpoints based on data pooled from both studies will be evaluated at the time of the last study-specific PA.

If the sample size is increased when the IA is performed, the PA will be triggered when the recalculated number of events from the pooled studies is reached based on events occurring during the blinded Treatment Period, or when the last participant from both studies from the

original enrollment reaches 96 weeks of treatment or discontinues from treatment prematurely during the blinded Treatment Period, and completes Safety Follow-up or discontinues from study prematurely during the blinded Treatment Period, whichever occurs last.

After protocol deviations are determined, and the database is locked for the PA, the drug codes will be broken and made available for the primary data analysis. All endpoints based on Baseline to Week 96 data will be evaluated.

9.4.4.3 Final Analysis

The final analysis will occur only when the last participant enrolled in the OLE completes the OLE or discontinues prematurely, the protocol deviations are determined, and the database is locked for the final analysis. All endpoints based on OLE data will be evaluated.

9.4.4.4 Multiplicity

To control family-wise type I error at the 1-sided 0.025 level in the presence of SSR based on pooled 12-week CDP (as measured by EDSS), the CHW method will be used to combine the incremental statistic corresponding to the first approximately 35% to 45% of pooled 12-week CDP information (as specified in the Interim Analysis Plan) with the incremental statistic corresponding to the remaining pooled 12-week CDP information, to form the CHW test statistic for pooled 12-week CDP to be used in the PA (Cui 1999).

To control trial-wise and family-wise type I error at the 1-sided 0.025 level in the presence of multiple endpoint testing, a graphical approach to sequentially rejective multiple testing will be employed (Bretz 2009, Hung 2013). The 1-sided null hypotheses in the graph are as follows, where "Study 1" denotes the present study, and "Study 2" denotes the second Phase III study:

Primary Endpoint Null Hypotheses:

H_{011}: $rRR_1 \geq 1$, where rRR_1 denotes qualified relapse rate ratio comparing evobrutinib to teriflunomide in Study 1

H_{012}: $rRR_2 \geq 1$, where rRR_2 denotes qualified relapse rate ratio comparing evobrutinib to teriflunomide in Study 2

Secondary Endpoint Null Hypotheses:

H_{02}: $S_{12e}(t) \leq S_{12c}(t)$, where $S_{12e}(t)$ denotes the survival function for time to 12-week CDP in the experimental (evobrutinib) group based on pooled data, $S_{12c}(t)$ denotes the survival function for time to 12-week CDP in the comparator (teriflunomide) group based on pooled data, and the variable t denotes time since randomization.

H_{03}: $S_{24e}(t) \leq S_{24c}(t)$, where $S_{24e}(t)$ and $S_{24c}(t)$ denotes the survival functions for time to 24-week CDP based on pooled data

H_{04}: $\Delta_{PF} \leq 0$, where Δ_{PF} denotes difference in PROMIS Physical Function score CFB at 96 weeks least-squares mean, comparing evobrutinib to teriflunomide, based on pooled data (higher score corresponds to improved physical function).

H_{05}: $\Delta_{Fatigue} \leq 0$, where $\Delta_{Fatigue}$ denotes difference in PROMIS Fatigue score CFB at 96 weeks least-squares mean, comparing evobrutinib to teriflunomide, based on pooled data (higher score corresponds to reduced fatigue).

H_{061}: $lRR_{61} \geq 1$, where lRR_{61} denotes T1 Gd+ lesion rate ratio comparing evobrutinib to teriflunomide in Study 1

H_{062}: $lRR_{62} \geq 1$, where lRR_{62} denotes T1 Gd+ lesion rate ratio comparing evobrutinib to teriflunomide in Study 2

H_{071}: $lRR_{71} \geq 1$, where lRR_{71} denotes new or enlarging T2 lesion rate ratio comparing evobrutinib to teriflunomide in Study 1

H_{072}: $lRR_{72} \geq 1$, where lRR_{72} denotes new or enlarging T2 lesion rate ratio comparing evobrutinib to teriflunomide in Study 2

At the PA, the primary efficacy endpoint, ARR, will be tested at the $0.025 – \alpha_1/2$ (1-sided) level in each study, assuming α_1 (1-sided) is spent in the conduct of the IA of 12-week CDP based on pooled data. Figure 2 shows the multiple testing procedure involving study-specific endpoints and pooled endpoints, with the simplifying assumption that $\alpha_1 = 0$.

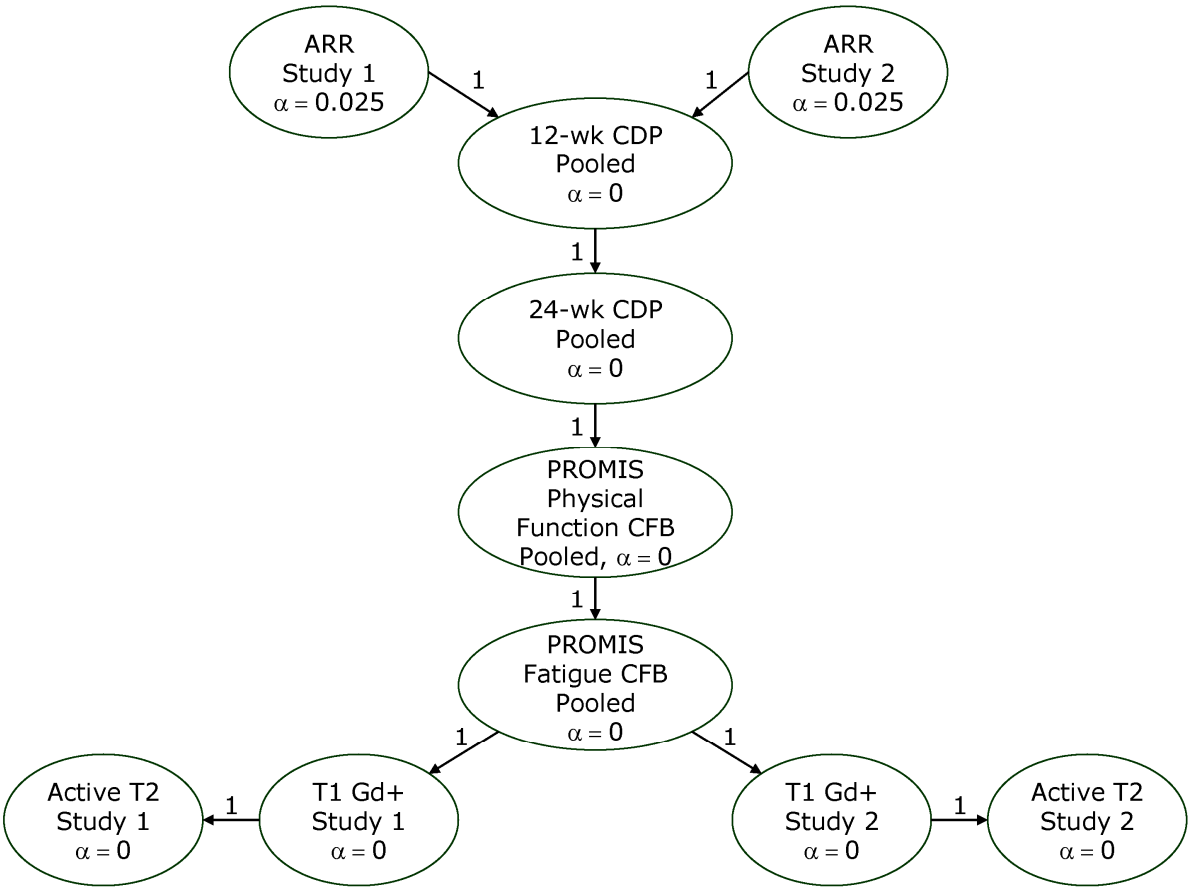
At the PA, the 12-week CDP pooled endpoint will be tested at the $0.025 – \alpha_1$ level, 1-sided, only if ARR is significant in both studies at the $0.025 – \alpha_1/2$ level, 1-sided.

If the 12-week CDP pooled endpoint is significant at the $0.025 – \alpha_1$ level, 1-sided, the subsequent pooled endpoints (24-week CDP, PROMIS PF CFB at Week 96, PROMIS Fatigue CFB at Week 96) will be tested in a hierarchical order at $0.025 – \alpha_1$ level, 1-sided.

If all the pooled endpoints (12-week CDP, 24-week CDP, PROMIS PF CFB at Week 96, PROMIS Fatigue CFB at Week 96) are significant at the $0.025–\alpha_1$ level, 1-sided, then the T1 Gd+ endpoint will be tested within each study at the $0.025– \alpha_1/2$ level, 1-sided.

If the T1 Gd+ endpoint within a study is significant at the $0.025– \alpha_1/2$ level, 1-sided, then the active T2 endpoint for that study will be tested at the same level.

Figure 2 Multiplicity Graph



ARR = Annualized Relapse Rate, CDP = Confirmed Disability Progression, CFB = change from Baseline, Gd^+ = gadolinium positive, PROMIS = Patient Reported Outcomes Measurement Information System, T1 and T2 = type of Magnetic Resonance Image, Wk = Week.

10 References

Alankus YB, Grenningloh R, Haselmayer P, et al. BTK Inhibition Prevents Inflammatory Macrophage Differentiation: A Potential Role in MS. EMD Serono, Global Clinical Development Center, Billerica, Massachusetts, USA. 2018 P557.

Amtmann D, Bamer AM, Kim J, et al. People with multiple sclerosis report significantly worse symptoms and health related quality of life than the US general population as measured by PROMIS and NeuroQol outcome measures. Disabil Health J. 2018;11(1):99-107.

Bar-Or A, Fawaz L, Fan B, et al. Abnormal B-cell cytokine responses a trigger of T-cell-mediated disease in MS? Ann Neurol. 2010;67(4):452-61.

Block H, Zarbock A. The Role of the Tec Kinase Bruton's Tyrosine Kinase (Btk) in Leukocyte Recruitment. Int Rev Immunol. 2012;31:104-18.

Boschert U, Crandall T, Pereira A, et al. T Cell Mediated Experimental CNS Autoimmunity Induced by PLP in SJL Mice is modulated by Evobrutinib (M2951) a Novel Bruton's Tyrosine Kinase Inhibitor. EMD Serono Research and Development Institute, Billerica, Massachusetts, USA. 2017 P678.

Brañas P, Jordan R, Fry-Smith A, et al. Treatments for fatigue in multiple sclerosis: a rapid and systematic review. Health Technol Assess. 2000;4(27):1-61.

Bretz F, Maurer W, Brannath W, et al. A graphical approach to sequentially rejective multiple test procedures. Stat Med. 2009;28:586-604.

Cella D, Gershon R, Lai JS, et al. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. Qual Life Res. 2007;16:133-41.

Cella D, Lai JS, Jensen SE, et al. PROMIS Fatigue Item Bank had Clinical Validity across Diverse Chronic Conditions. J Clin Epidemiol. 2016;73:128-34.

Cook KF, Bamer AM, Amtmann D, et al. Six Patient-Reported Outcome Measurement Information System Short Form Measures Have Negligible Age- or Diagnosis-Related Differential Item Functioning in Individuals With Disabilities. Arch Phys Med Rehabil. 2012;93:1289-91.

Crowe BJ, Xia HA, Berlin JA, et al. Recommendations for safety planning, data collection, evaluation and reporting during drug, biologic and vaccine development: a report of the safety planning, evaluation, and reporting team. Clin Trials. 2009;6(5):430-40.

Cui L, Hung HMJ, Wang SJ. Modification of Sample Size in Group Sequential Clinical Trials. Biometrics. 1999;55:853-7.

De Stefano N, Iannucci G, Sormani MP, et al. MR correlates of cerebral atrophy in patients with multiple sclerosis. J Neurol. 2002;249(8):1072-7.

Drug Development Tools Clinical Outcome Assessments #000069: PROMIS® Physical Functioning Item Bank. Available at: https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm626855.htm. Accessed 11 April 2019.

Fraussen J, de Bock L, Somers V et al. B cells and antibodies in progressive multiple sclerosis: Contribution to neurodegeneration and progression. Autoimmun Rev. 2016;15(9):896-9.

Freeman JA, Hobart JC, Langdon DW, et al. Clinical appropriateness: a key factor in outcome measure selection: the 36-item short-form health survey in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2000;68(2):150-6.

Haselmayer P, Camps M, Liu-Bujalski L, et al. Pharmacodynamic Modeling of BTK Occupancy Versus Efficacy in Ra and SLE Models Using the Novel Specific BTK Inhibitor Evobrutinib. ACR/ARHP Annual Meeting; 2017: Abstract number:2565.

Hauser SL, Bar-Or A, Cohen J, et al. Efficacy and safety of ofatumumab versus teriflunomide in relapsing multiple sclerosis: results of the phase 3 ASCLEPIOS I and II trials. Mult Scler. 2019;25(S2):336. Presentation available at https://onlinelibrary.ectrims-congress.eu/ectrims/2019/stockholm/279581/stephen.hauser.efficacy.and.safety.of.ofatumumab.versus.teriflunomide.in.html?f=listing%3D0%2Abrowseby%3D8%2Asortby%3D1%2Asearch%3DASCLEPIOS.

Hauser SL, Waubant E, Arnold DL, et al. B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. N Engl J Med. 2008;358(7):676-88.

Hawker K, O'Connor P, Freedman MS, et al. Rituximab in Patients with Primary Progressive Multiple Sclerosis, Results of a Randomized Double-Blind Placebo-Controlled Multicenter Trial. Ann Neurol. 2009;66:460–71.

Hung HMJ and Wang S. Multiple comparisons in complex clinical trial designs. Biometrical Journal. 2013;55(3):420-9.

Jelcic I, Nimer FA, Wang J, et al. Memory B cells Activate Brain-Homing, Autoreactive CD4+ T Cells in Multiple Sclerosis. Cell. 2018;175(1):85-100 Epub.

Jongen PJ. Health-Related Quality of Life in Patients with Multiple Sclerosis: Impact of Disease-Modifying Drugs. CNS Drugs. 2017;31(7):585-602.

Kamudoni P, Henke C, Barrett A, et al. A new PROMIS physical function short form for use in relapse and progressive multiple sclerosis types. Qual Life Res. 2018;27(1):S1–S190.

Kappos L, Hartung H-P, Freedman MS, et al. Atacicept in multiple sclerosis (ATAMS): a randomised, placebo-controlled, double-blind, phase 2 trial. Lancet Neurol. 2014;13:353–63.

Lai JS, Cella D, Choi S, et al. How item banks and their application can influence measurement practice in rehabilitation medicine: a PROMIS fatigue item bank example. Arch Phys Med Rehabil. 2011;92:S20-7.

LaRocca NG. Impact of Walking Impairment in Multiple Sclerosis Perspectives of Patients and Care Partners. Patient. 2011;4(3):189-201.

Li R, Rezk A, Miyazaki Y, et al. Proinflammatory GM-CSF–producing B cells in multiple sclerosis and B cell depletion therapy. Sci Transl Med. 2015;7(310):1-10.

López-Herrera G, Vargas-Hernández A, González-Serrano ME et al. Bruton's tyrosine kinase-an integral protein of B cell development that also has an essential role in the innate immune system. J. Leukoc Biol. 2014;95:243-50.

Martin SA, Olayinka-Amao OC, Henke C, et al. Conceptual Model and Instrument Review in Multiple Sclerosis. Value Health 2017;20(9):A726-7.

McHorney CA, Ware JE Jr, Rachel Lu JF, et al. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care. 1994;32(1):40-66.

McHorney CA, Ware JE Jr, Rogers W, et al. The validity and relative precision of MOS short and long form health status scales and Dartmouth COOP charts. Med Care. 1992;30(5):MS253-65.

McHorney CA, Ware JE, Raczek AE. The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care. 1993;31(3):247-63.

Montalban X, Hemmer B, Rammohan K, et al. A Randomized, Double-Blind, Parallel-Group, Placebo-Controlled Phase III Trial to Evaluate Efficacy and Safety of Ocrelizumab in Primary Progressive Multiple Sclerosis. Poster session presented at: Americas Committee for Treatment and Research in Multiple Sclerosis (ACTRIMS) Forum; 2016 Feb 18-20; New Orleans, LA.

Montalban X, Arnold DL, Weber MS, et al. Placebo-Controlled Trial of an Oral BTK Inhibitor in Multiple Sclerosis. N Engl J Med. 2019; 380:2406-17.

Multiple Sclerosis Working Group, presented at the Ninth Annual PRO Consortium Workshop. Available at: https://c-path.org/wp-content/uploads/2018/05/2018_MS_WG_Poster.pdf. Accessed 11 April 2019.

Patti F, Amato MP, Trojano M, et al. Quality of life, depression and fatigue in mildly disabled patients with relapsing-remitting multiple sclerosis receiving subcutaneous interferon beta-1a: 3-year results from the COGIMUS (COGnitive Impairment in MUltiple Sclerosis) study. Mult Scler. 2011;17(8):991-1001.

Patti F, Vila C. Symptoms, prevalence and impact of multiple sclerosis in younger patients: a multinational survey. Neuroepidemiology. 2014;42:211-8.

Polman CH, Reingold SC, Barkhof F, et al. Ethics of placebo-controlled clinical trials in multiple sclerosis: a reassessment. Neurology. 2008;70(13 Pt 2):1134-40.

Poser CM and Brinar VV. Epilepsy and multiple sclerosis. Epilepsy Behav. 2003;4(1):6-12.

Rawji KS, Mishra MK, Michaels NJ, et al. Immunosenescence of microglia and macrophages: impact on the ageing central nervous system. Brain. 2016;139:653-61.

Rose M, Bjorner JB, Gandek B, et al. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. J Clin Epidemiol. 2014;67:516-26.

Teriflunomide (Aubagio®) SmPC. https://www.ema.europa.eu/en/medicines/human/EPAR/aubagio. Last accessed on 25 November, 2019.

Teriflunomide (Aubagio®) USPI. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=202992. Last accessed on 25 November, 2019.

Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018;17(2):162-73.

Torke S, Grenningloh R, Boschert U, et al. Inhibition of Bruton's tyrosine kinase selectively prevents antigen-activation of B cells and ameliorates B cell-mediated experimental autoimmune encephalomyelitis. University Medical Center, Gottingen, Germany. 2018 P575.

Vogel DYS, Vereyken EJF, Glim JE et al. Macrophages in inflammatory multiple sclerosis lesions have an intermediate activation status. J Neuroinflammation. 2013;10(35):epub.

Ware JE, Sherbourne CD. The MOS 36-ltem Short-Form Health Survey (SF-36) I. Conceptual Framework and Item Selection. Med Care. 1992;30(6):473-83.

Whang JA, Chang BY. Brutons tyrosine kinase inhibitors for the treatment of rheumatoid arthritis. Drug Discovery Today. 2014, doi:10.1016/j.drudis.

WHO Guidelines on Tuberculosis Infection Prevention and Control. 2019 Update. Geneva. World Health Organization. 2019.

WHO TB Burden Estimates. https://extranet.who.int/tme/generateCSV.asp?ds=estimates. Last accessed on 09 December 2019.

Wolinsky J, Arnold DL, Bar-Or A, et al. Efficacy of Ocrelizumab in Patients With PPMS With and Without T1 Gadolinium-Enhancing Lesions at Baseline in a Phase III, Placebo-Controlled Trial. Poster session presented at: Americas Committee for Treatment and Research in Multiple Sclerosis (ACTRIMS) Forum; 2016 Feb 18-20; New Orleans, LA.

Yost KJ, Eton DT, Garcia SF, et al. Minimally important differences were estimated for six PROMIS-Cancer scales in advanced-stage cancer patients. J Clin Epidemiol. 2011;64(5):507-16.

11 Appendices

Appendix 1 Abbreviations

9-HPT	9 Hole Peg Test
AE	Adverse Event
AESI	Adverse Event of Special Interest
ALT	Alanine aminotransferase
ARR	Annualized Relapse Rate
AST	Aspartate aminotransferase
AUC	Area under the curve
BEA	Blinded Extension Analysis
BTK	Bruton's Tyrosine Kinase
BV	Brain volume
CDP	Confirmed disability progression
CFB	Change from Baseline
CHW	Cui, Hung, Wang
CI	Confidence interval
CNS	Central nervous system
CP	Conditional power
CRF	Case Report Form
CRO	Contract Research Organization
CSF	Cerebrospinal fluid
CSF JCV PCR	Cerebrospinal fluid JC virus polymerase chain reaction
CSR	Clinical study report
C-SSRS	Columbia-Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
CUA	Combined unique active
CYP3A	Cytochrome P450 3A
DAC	Data Analysis Center
DDI	Drug-drug interactions

DMD	Disease-modifying drugs
EAC	Endpoint Adjudication Committee
ECG	Electrocardiogram
EDC	Electronic Data Capture
EDSS	Expanded Disability Status Scale
eGFR	Estimated glomerular filtration rate
EQ-5D-5L	EuroQoL 5 Dimension 5 Levels
FAS	Full Analysis Set
FDA	Food and Drug Administration
FLAIR	Fluid-attenuated inversion recovery
FSH	Follicle-stimulating hormone
FSS	Functional system scores
GCP	Good Clinical Practice
Gd	Gadolinium
HCV	Hepatitis C virus
HFE	High Iron Fe (human hemochromatosis protein)
HIV	Human immunodeficiency virus
HRQoL	Health-related quality of life
HRU	Health Resource Utilization
HRT	Hormonal replacement therapy
IA	Interim analysis
IAP	Integrated Analysis Plan
ICF	Informed consent form
ICH	International Council for Harmonisation
IDMC	Independent Data Monitoring Committee
IEC	Independent Ethics Committee
Ig	Immunoglobulin
IRB	Institutional Review Board
IWRS	Interactive Web Response System
LFT	Liver function test
LTBI	Latent TB infection

MCID	Minimal clinically important difference
MedDRA	Medical Dictionary of Regulatory Activities
MRI	Magnetic Resonance Imaging
MMRM	Mixed effect model for repeated measures
MS	Multiple sclerosis
NB	Negative binomial
NCI	National Cancer Institute
NYHA	New York Heart Association
OLE	Open Label Extension
PA	Primary Analysis
PCR	Polymerase chain reaction
PF	Physical function
PK	Pharmacokinetics
PML	Progressive multifocal leukoencephalopathy
PRO	Patient Reported Outcomes
PROMIS	Patient Reported Outcomes Measurement Information System
QoL	Quality of Life
RA	Rheumatoid arthritis
RMS	Relapsing multiple sclerosis
RR	Rate ratio
SAE	Serious Adverse Event
SCR	Screening
SDMT	Symbol Digit Modalities Test
SEL	Slowly evolving lesions
SF-36v2	36-Item Short Form Survey Instrument Version 2
SLE	Systemic lupus erythematosus
SoA	Schedule of activities
SOC	System Organ Class
SPMS	Secondary progressive multiple sclerosis
SSC	Study Steering Committee
SSR	Sample size re-estimation

SUSAR	Suspected Unexpected Serious Adverse Reaction
T25-FW	Timed 25-Foot Walk
TB	Tuberculosis
TEAE	Treatment emergent adverse event
ULN	Upper limit of normal
WHO	World Health Organization
WOCBP	Woman of Childbearing Potential

Appendix 2 Study Governance

Financial Disclosure

Investigators and Sub-Investigators will provide the Sponsor with sufficient, accurate financial information, as requested, for the Sponsor to submit complete and accurate financial certification or disclosure statements to the appropriate regulatory authorities. This information is required during the study and for 1 year after completion of the study.

Informed Consent Process

- The Investigator or his/her representative will explain the nature of the study to the participant or his/her legally authorized representative and answer all questions on the study.
- Participants must be informed that their participation is voluntary.
- Participants or their legally authorized representative will be required to sign a statement of informed consent that meets the requirements of 21 CFR 50; local regulations; ICH guidelines; Health Insurance Portability and Accountability Act (HIPAA) requirements, where applicable; and the IRB/IEC or study center.
- The medical record must include a statement that written informed consent was obtained before the participant was enrolled in the study and the date the written consent was obtained. The authorized person obtaining the informed consent must also sign the ICF.
- If the ICF is updated during their participation in the study, participants must be reconsented to the most current, approved version.
- A copy of the ICF(s) must be provided to the participant or the participant's legally authorized representative.
- The original signed and dated consent will remain at the Investigator's site and must be safely archived so that it can be retrieved at any time for monitoring, auditing and inspection purposes.
- As this study includes optional pharmacogenetic examinations, including collection and storage of biological samples, participants may consent to a separate pharmacogenetic analysis, the process of which will need to be documented in the participant's medical records.
- Participants who are rescreened are required to sign a new ICF.
- Consenting participant will enter the 4-week screening period to be evaluated for eligibility. Please see the SoA (see Section 1.3.1) for details. Participants must fulfill all entry criteria for participation in the study.

- The screening period may be extended to a total period of 8 weeks after approval by the Medical Monitor. The following should be performed:
  - An Eligibility Review Form [ERF] documenting the Investigator's assessment of each screened participant with regard to the protocol's inclusion and exclusion criteria is to be completed by the Investigator.
  - Each participant screened must be registered in the IWRS by the Investigator or the Investigator's research staff at Screening. A screen failure record must be maintained by the Investigator, and reasons must be captured in the IWRS.
  - It should be stated in the medical record that the participant is participating in this clinical study.
  - Eligibility will be evaluated and confirmed by the study eligibility team. Sites will be required to submit an eligibility packet to the Medical Monitor (consisting of an eligibility checklist and appropriate documentation) for potential eligible participants.

Data Protection

- The Sponsor will assign a unique identifier to participants after obtaining their informed consent. All participant records or datasets transferred to the Sponsor will contain the identifier only; participant names or any identifiable information will not be transferred.
- The Sponsor must inform participants that their personal study-related data will be used per local data protection and privacy laws. The level of disclosure must also be explained to the participant.
- The participant must be informed that his/her medical records may be examined by Clinical Quality Assurance auditors or other Sponsor-appointed, authorized personnel, by appropriate IRB/IEC members, and by regulatory authority inspectors. All such persons will strictly maintain participants' confidentiality.

Study Administrative

This clinical study will be sponsored by Merck Healthcare KGaA Darmstadt, Germany for sites outside of the US and Canada and EMD Serono Research & Development Institute, Inc., Billerica, MA, US for sites in the US and Canada.

The study will be conducted at approximately 200 global sites anticipated from approximately 30 countries (approximately 30 sites in the US). Sites will be a mixture of academic centers and outpatient clinics.

The Coordinating Investigator listed on the title page represents all Investigators for decisions and discussions on this study, per ICH GCP. The Coordinating Investigator will provide expert medical input and advice on the study design and execution and is responsible for the review and signoff of the clinical study report.

An IDMC, EAC, and SSC will perform specific study-related activities as detailed in each committee's charter (see Section 8.2.8).

The study will appear in the following clinical studies registries: ClinicalTrials.gov and EudraCT.

Details of structures and associated procedures will be defined in a separate Pharmacy Manual.

Regulatory and Ethical Considerations

- This study will be conducted in accordance with the protocol and the following:
    - Consensus ethical principles derived from international guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines
    - Applicable ICH GCP Guidelines
    - Applicable laws and regulations
- The Investigator must submit the protocol, protocol amendments (if applicable), ICF, Investigator Brochure, and other relevant documents (e.g., advertisements) to an IRB/IEC and the IRB/IEC must review and approve them before the study is initiated.
- Any protocol amendments (i.e., changes to the protocol) will be documented in writing and require IRB/IEC approval before implementation of changes, except for changes necessary to eliminate an immediate hazard to study participants. When applicable, amendments will be submitted to the appropriate Health Authorities.
- The Investigator will be responsible for the following:
    - Providing written summaries of the status of the study to the IRB/IEC annually or more frequently per the IRB's/IEC's requirements, policies, and procedures.
    - Notifying the IRB/IEC of SAEs or other significant safety findings, as required by IRB/IEC procedures
    - Providing oversight of the study conduct at the site and adherence to requirements of 21 CFR, ICH guidelines, the IRB/IEC, European regulation 536/2014 for clinical studies (if applicable), and all other applicable local regulations
- The protocol and any applicable documentation will be submitted or notified to the Health Authorities in accordance with all local and national regulations for each site.

Emergency Medical Support

- The Sponsor or designee will provide Emergency Medical Support cards to participants for use during the study. These provide the means for participants to identify themselves as participating in a clinical study. Also, these give health care providers access to any information about this participation that may be needed to determine the course of medical

treatment for the participant. The information on the Emergency Medical Support card may include the process for emergency unblinding (if applicable).

- The first point of contact for all emergencies will be the clinical study Investigator caring for the participant. Consequently, the Investigator agrees to provide his or her emergency contact information on the card. If the Investigator is available when an event occurs, they will answer any questions. Any subsequent action (e.g., unblinding) will follow the standard process established for Investigators.

When the Investigator is not available, the Sponsor provides the appropriate means to contact a Sponsor (or designee) physician. This includes provision of a 24-hour contact number at a call center, whereby the health care providers will be given access to the appropriate Sponsor (or designee) physician to assist with the medical emergency and to provide support for the potential unblinding of the participant concerned.

Clinical Study Insurance and Compensation to Participants

Insurance coverage will be provided for each country participating in the study. Insurance conditions shall meet good local standards, as applicable.

Clinical Study Report

After study completion, the Sponsor will write a clinical study report in consultation with the Coordinating Investigator.

Publication

- The results of this study may be published or presented at scientific meetings. If this is foreseen, the Investigator agrees to submit all manuscripts or abstracts to the Sponsor before submission. This allows Merck to protect proprietary information and to provide comments.
- The Sponsor will comply with the requirements for publication of study results. Per standard editorial and ethical practice, the Sponsor will generally support publication of multicenter studies only in their entirety and not as individual site data. In this case, a coordinating Investigator will be designated by agreement.
- Authorship will be determined by agreement and in line with International Committee of Medical Journal Editors authorship requirements.

Dissemination of Clinical Study Data

After completion of the study, a clinical study report will be written by the Sponsor in consultation with the Coordinating Investigator following the guidance in ICH Topic E3, and will be submitted in accordance with local regulations.

Any and all scientific, commercial, and technical information disclosed by the Sponsor in this protocol or elsewhere should be considered the confidential and proprietary property of the Sponsor. The Investigator shall hold such information in confidence and shall not disclose the

information to any third party except to such of the Investigator's employees and staff as have been made aware that the information is confidential and who are bound to treat it as such and to whom disclosure is necessary to evaluate that information. The Investigator shall not use such information for any purpose other than for determining mutual interest in performing the study and, if the parties decide to proceed with the study, for the purpose of conducting the study.

The Investigator understands that the information developed from this clinical study will be used by the Sponsor in connection with the development of the study drug and therefore may be disclosed as required to other clinical Investigators, to the US Food and Drug Administration, and to other government agencies. The Investigator also understands that, to allow for the use of the information derived from the clinical study, the Investigator has the obligation to provide the Sponsor with complete test results and all data developed in the study.

No publication or disclosure of study results will be permitted except under the terms and conditions of a separate written agreement.

Data Quality Assurance

- All participant study data will be recorded on printed or electronic CRFs or transmitted to the Sponsor or designee electronically (e.g., laboratory data). The Investigator is responsible for verifying that data entries are complete, accurate, legible, and timely by physically or electronically signing the CRF. Details for managing CRFs are in the Pharmacy Manual.
- For PRO data (e.g., QoL and pain assessments), ePRO will be used.
- The Investigator must maintain accurate documentation (source data) that supports the information in the CRF.
- The Investigator must permit study-related monitoring, quality assurance audits, IRB/IEC review, and regulatory agency inspections and provide direct access to the study file and source data.
- Monitoring details describing strategy (e.g., risk-based initiatives in operations and quality such as Risk Management and Mitigation Strategies and Analytical Risk-Based Monitoring), methods, responsibilities and requirements, including handling of noncompliance issues and monitoring techniques (central, remote, or on-site monitoring) are in the Monitoring Plan.
- The Sponsor or designee is responsible for data management of this study, including quality checking of the data and maintaining a validated database. Database lock will occur once quality control and quality assurance procedures have been completed. PDF files of the CRFs will be provided to the Investigators at study completion.
- Study monitors will perform ongoing source data verification to confirm that data in the CRF are accurate, complete, and verifiable; that the safety and rights of participants are being protected; and that the study is being conducted per the currently approved protocol and any other study agreements, ICH GCP, and all applicable regulatory requirements.

- Records and documents, including signed ICFs, pertaining to the conduct of this study must be retained by the Investigator for 15 years after study completion, unless local regulations, institutional policies, or the Sponsor requires a longer retention. No records may be destroyed during the retention period without the Sponsor's written approval. No records may be transferred to another location or party without the Sponsor's written notification.

Source Documents

- Source documents provide evidence for the existence of the participant and substantiate the integrity of the data collected.
- The Investigator must keep a paper or electronic file (medical file and original medical records) at the site for each study participant. The file must identify each participant, contain the following demographic and medical information for the participant, and should be as complete as possible:
    - Participant's full name, date of birth, sex, height, and weight
    - Medical history and concomitant diseases
    - Prior and concomitant therapies (including changes during the study)
    - Study identifier (i.e., the Sponsor's study number) and participant's study number.
    - Dates of entry into the study (i.e., signature date on the informed consent) and each visit to the site
    - Any medical examinations and clinical findings predefined in the protocol
    - All AEs
    - Date that the participant left the study, including any reason for early withdrawal from the study or study intervention, if applicable.
- All source data must be filed (e.g., CT or MRI scan images, ECG recordings, and laboratory results). Each document must have the participant number and the procedure date; ideally, printed by the instrument used for the procedure. As necessary, medical evaluation of these records should be performed, documented, signed and dated by the Investigator.
- Data recorded on printed or electronic CRFs that are transcribed from source documents must be consistent with the source documents or the discrepancies must be explained. The Investigator may need to request previous medical records or transfer records, depending on the study. Also, current medical records must be available.
- The study monitors will use printouts of electronic files for source data verification. These printouts must be signed and dated by the Investigator and kept in the study file.
- Source documents are stored at the site for the longest possible time permitted by the applicable regulations, and/or as per ICH GCP guidelines, whichever is longer. The Investigator ensures that no destruction of medical records is performed without the Sponsor's written approval.

- Definition of what constitutes source data is found in the eCRF guidelines.

Study and Site Closure

- The Sponsor reserves the right to close the study site or terminate the study at any time and for any reason. Study sites will be closed upon study completion. A study site is considered closed when all required documents and study supplies have been collected and a site closure visit has been completed.
- The Investigator may initiate site closure at any time, provided there is reasonable cause and sufficient notice is given in advance of the intended termination.
- Reasons for the early closure of a study site by the Sponsor or Investigator may include:
  - Failure of the Investigator to comply with the protocol, the requirements of the IRB/IEC or local health authorities, the Sponsor's procedures, or GCP guidelines
  - Inadequate recruitment of participants by the Investigator
  - Discontinuation of further development of the Sponsor's compound

Appendix 3 Contraception

Definitions:

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming postmenopausal unless permanently sterile, as specified below.

If fertility is unclear (e.g., amenorrhea in adolescents or athletes) and a menstrual cycle cannot be confirmed before the first dose of study intervention, consider additional evaluation.

A WOCBP is **not**:

1. Premenarchal

2. A premenopausal female with 1 of the following:
   - Documented hysterectomy
   - Documented bilateral salpingectomy
   - Documented bilateral oophorectomy

   Documentation can come from the site personnel's review of the female's medical records, medical examination, or medical history interview.

   For a female with permanent infertility due to an alternate medical cause other than the above, (e.g., Müllerian agenesis, androgen insensitivity), Investigator discretion applies to determine study entry.

3. A postmenopausal female
   - A postmenopausal state is defined as no menses for 12 months without an alternative medical cause.
     - A high follicle-stimulating hormone (FSH) level in the postmenopausal range may be used to confirm a postmenopausal state in a female not using hormonal contraception or hormonal replacement therapy (HRT). However, in the absence of 12 months of amenorrhea, more than 1 FSH measurement is required in the postmenopausal range.
   - A female on HRT and whose menopausal status is in doubt will be required to use one of the nonestrogen hormonal highly effective contraception methods if she wishes to continue her HRT during the study. Otherwise, she must discontinue HRT to allow confirmation of postmenopausal status before study enrollment.

Contraception Guidance:

CONTRACEPTIVES ALLOWED DURING THE STUDY INCLUDE:

Highly Effective Methods That Have Low User Dependency

- Implantable progestogen-only hormone contraception associated with inhibition of ovulation[a]
- Intrauterine device (IUD)
- Intrauterine hormone-releasing system (IUS)[a]
- Bilateral tubal occlusion
- Vasectomized partner: a highly effective contraceptive method provided that the partner is the sole sexual partner of a WOCBP and the absence of sperm has been confirmed. Otherwise, use an additional highly effective method of contraception. The spermatogenesis cycle is approximately 90 days.

Highly Effective Methods That Are User Dependent

- Combined (estrogen- and progestogen-containing) hormonal contraception associated with inhibition of ovulation[a]
  - Oral
  - Intravaginal
  - Transdermal
  - Injectable
- Progestogen-only hormone contraception associated with inhibition of ovulation[a]
  - Oral
  - Injectable
- Sexual abstinence: a highly effective method only if defined as refraining from intercourse during the entire period of risk associated with the study intervention. The reliability of sexual abstinence needs to be evaluated in relation to the duration of the study. Abstinence is only acceptable as a contraceptive method for study purposes if it is in line with the preferred and usual lifestyle of the participant as evaluated by the Investigator.

Barrier Methods

- Male or female condom with or without spermicide
- Cap, diaphragm, or sponge with spermicide

Notes:

Contraceptive use by men or women is consistent with local regulations on the use of contraceptive methods for clinical study participants.

Highly effective methods are those with a failure rate of < 1% per year when used consistently and correctly. Typical use failure rates differ from those when used consistently and correctly.

If locally required, in accordance with Clinical Trial Facilitation Group (CTFG) guidelines, acceptable contraceptive methods are limited to those which inhibit ovulation as the primary mode of action.

Evobrutinib has been characterized as both an inducer and a time-dependent inhibitor of CYP3A4/5 in vitro, an enzyme involved in the metabolism of estrogen and progestin. Although it is a low risk, it is possible that coadministration with evobrutinib results in increased metabolism of estrogen, progestin, and other hormones used for contraception, increasing the likelihood that hormonal contraception methods, marked above with [a], might fail. However, all participants that are WOCBP are required to use a barrier method as backup (see Inclusion Criterion 6) in part to mitigate this potential risk.

Teriflunomide may increase the systemic exposures of ethinylestradiol and levonorgestrel. Consideration should be given to the type or dose of contraceptives used in combination with teriflunomide (see Appendix 9 and refer to the locally approved product information [e.g., relevant SmPC or USPI]). There was an increase in mean ethinylestradiol C_{max} and AUC_{0-24} (1.58- and 1.54-fold, respectively) and levonorgestrel C_{max} and AUC_{0-24} (1.33- and 1.41-fold, respectively) following repeated doses of teriflunomide.

Periodic abstinence (calendar, symptothermal, postovulation methods), withdrawal (coitus interruptus), spermicides only, and lactational amenorrhoea method (LAM) are **not** acceptable methods of contraception for this study. Male condom and female condom cannot be used together (due to risk of failure with friction).

Appendix 4 Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting

Definitions

Adverse Event

An AE is any untoward medical occurrence in a participant administered a pharmaceutical product, regardless of causal relationship with this treatment. Therefore, an AE can be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal product, regardless if it is considered related to the medicinal product.

For surgical or diagnostic procedures, the condition/illness leading to such a procedure is considered as the AE rather than the procedure itself.

The Investigator is required to grade the severity or toxicity of each AE.

Investigators will reference the NCI-CTCAE, version 5.0 (publication date: 27 November 2017), a descriptive terminology that can be used for AE reporting.

A general grading (severity/intensity; hereafter referred to as severity) scale is provided at the beginning of the above referenced document, and specific event grades are also provided.

If the severity for an AE is not specifically graded by NCI-CTCAE, the Investigator is to use the general NCI-CTCAE definitions of Grade 1 through Grade 5, using his or her best medical judgment.

The 5 general grades are:

Grade 1 or Mild

Grade 2 or Moderate

Grade 3 or Severe

Grade 4 or Life-threatening

Grade 5 or Death

Any clinical AE with severity of Grade 4 or 5 must also be reported as an SAE. However, a laboratory abnormality of Grade 4, such as anemia or neutropenia, is considered serious only if the condition meets one of the serious criteria specified below.

If death occurs, the primary cause of death or event leading to death should be recorded and reported as an SAE. “Fatal” will be recorded as the outcome of this specific event and death will not be recorded as separate event. Only, if no cause of death can be reported (e.g., sudden death, unexplained death), the death per se might then be reported as an SAE.

Investigators must also systematically assess the causal relationship of AEs to study intervention (including any other nonstudy interventions, radiation therapy, etc.) using the following definitions. Decisive factors for the assessment of causal relationship of an AE to the study interventions (evobrutinib and teriflunomide) include, but may not be limited to, temporal relationship between the AE and the study interventions, known side effects of study interventions, medical history, concomitant medication, course of the underlying disease, and study procedures.

Unrelated: Not reasonably related to the study intervention. AE could not medically (pharmacologically/clinically) be attributed to the study intervention under study in this clinical study protocol. A reasonable alternative explanation must be available.

Related: Reasonably related to the study intervention. AE could medically (pharmacologically/clinically) be attributed to the study intervention under study in this clinical study protocol.

Abnormal Laboratory Findings and Other Abnormal Investigational Findings

Abnormal laboratory findings and other abnormal investigational findings (e.g., on an ECG trace) should not be reported as AEs unless they are associated with clinical signs and symptoms, lead to study intervention discontinuation or are considered otherwise medically important by the Investigator. If a laboratory abnormality fulfills these criteria, the identified medical condition (e.g., anemia or increased ALT) must be reported as the AE rather than the abnormal value itself.

Serious Adverse Events

An SAE is any untoward medical occurrence that at any dose:

- Results in death
- Is life-threatening. Life-threatening refers to an event in which the participant is at risk of death at the time of the event, not an event that hypothetically might have caused death if it was more severe.
- Requires inpatient hospitalization or prolongs an existing hospitalization
- Results in persistent or significant disability or incapacity
- Is a congenital anomaly or birth defect
- Is otherwise considered to be medically important. Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered as SAEs when, based upon appropriate medical judgment, they may jeopardize the participant or may require medical or surgical intervention to prevent one of the outcomes listed above. Examples of such events include allergic bronchospasm requiring intensive treatment in an emergency room or at home, blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse.

For the purposes of reporting, any suspected transmission of an infectious agent via a study intervention is also considered an SAE, as specified below for reporting SAEs and AESIs.

Events that Do Not Meet the Definition of an SAE

Elective hospitalizations to administer, or to simplify study intervention or procedures (e.g., an overnight stay to facilitate intravenous therapy) are not considered SAEs. However, all events leading to unplanned hospitalizations or unplanned prolongation of an elective hospitalization (i.e., undesirable effects of any administered treatment) must be documented and reported as SAEs.

Events Not to Be Considered as AEs/SAEs

Medical conditions present at the initial study visit that do not worsen in severity or frequency during the study are defined as Baseline Medical Conditions, and are not to be considered AEs.

Adverse Events of Special Interest

Adverse events of special interest are liver AEs (possible drug-induced, non-infectious, non-alcoholic, and immune-mediated), infections (serious and opportunistic infections), lipase and amylase elevation, and seizure, as described in Section 6.9.

Other Adverse Events to be Reported Following a Specialized Procedure

The following procedures should be followed for reporting overdoses (refer to eCRF guidelines for further details):

- Overdoses without an AE should be reported using the paper SAE form only, stating if the overdose was accidental or intentional.
- Overdoses associated with a nonserious AE should be recorded on the AE eCRF and also the paper SAE form.
- Overdoses associated with an SAE should be recorded on the AE eCRF and the SAE reporting procedure outlined below should be followed.

Recording and Follow-Up of AE and/or SAE

It is important that each AE report include a description of the event, its duration (onset and resolution dates and also onset and resolution times, when it is important to assess the time of AE onset relative to the recorded study intervention administration time), its severity, its causal relationship with the study intervention, any other potential causal factors, any treatment given or other action taken, including dose modification or discontinuation of the study intervention, and its outcome. In addition, serious cases should be identified and the appropriate seriousness criteria documented.

Specific guidance is in the CRF Completion and Monitoring Conventions provided by the Sponsor.

Reporting Serious Adverse Events and Adverse Events of Special Interest

Serious Adverse Events

In the event of any new SAE occurring during the reporting period, the Investigator must immediately (within a maximum of 24 HOURS after becoming aware of the event) inform the Sponsor or its designee using the electronic SAE report form in the Electronic Data Capture (EDC) system.

Reporting of SAEs using a paper report form is required as a back-up method only for an EDC system failure. Names, addresses, and telephone and fax numbers will be included on the paper form. All information from the paper form must be transcribed into the electronic form as soon as the system becomes available.

In exceptional circumstances, an SAE (or follow-up information) may be reported by telephone; in these cases, an electronic SAE report form must be completed immediately thereafter.

Relevant pages from the CRF may be provided in parallel (e.g., medical history, concomitant drugs). Additional documents may be provided by the Investigator, if available (e.g., laboratory results, hospital report, autopsy report).

The Investigator must respond to any request for follow-up information (e.g., additional information, outcome, final evaluation, other records where needed) or to any question the Sponsor/designee may have on the AE within the same timelines as those noted above for initial reports. This is necessary to ensure prompt assessment of the event by the Sponsor or designee and (as applicable) to allow the Sponsor to meet strict regulatory timelines associated with expedited safety reporting obligations.

Requests for follow-up will usually be made via the study monitor, although in exceptional circumstances the drug safety department may contact the Investigator directly to obtain further information or to discuss the event.

Adverse Events of Special Interest

In the event of a nonserious AESI, the Investigator will complete the AESI Report Form and send it to the Sponsor/designee within 7 days (nonserious AESIs), or 24 hours (AESIs classified as serious). Names, addresses, and telephone and fax numbers for AESI reporting will be included on the Report Form. Serious AESIs must be reported in an expedited manner as SAEs as outlined above.

Appendix 5 Clinical Laboratory Tests

Table 4 Protocol-Required Clinical Laboratory Assessments

Laboratory Assessments	Parameters			
Hematology	Platelet count		Mean Corpuscular Volume (MCV)	White Blood Cell (WBC) Count with Differential: • Neutrophils • Lymphocytes • Monocytes • Eosinophils • Basophils
	Reticulocytes			
	Hemoglobin		Mean corpuscular hemoglobin (MCH) Mean corpuscular hemoglobin concentration (MCHC)	
	Hematocrit			
Coagulation	International normalized ratio	Partial thromboplastin time		
Supplementary LFT visits	Aspartate aminotransferase		γ-Glutamyl-transferase	Total Bilirubin
	Alanine aminotransferase		Alkaline phosphatase	
Biochemistry	Blood Urea Nitrogen	Potassium	Aspartate Aminotransferase	Bilirubin
	Creatinine and eGFR calculation	Sodium	Alanine Aminotransferase	Protein
	Glucose	Calcium	Alkaline phosphatase	Albumin
	Chloride	Lactate dehydrogenase	Lipase	Amylase
		Magnesium	Total carbon dioxide	Phosphate
		γ-Glutamyl-transferase	Immunoglobulin and subclass concentrations (as specified in the SoA, see Section 1.3)[a]	
Details of liver chemistry stopping criteria and required actions and follow-up assessments after liver stopping or monitoring event are given in Section 7.1.				
Hepatic/Autoimmune Panel (to be performed in the event of elevated LFTs, see Section 7.1)	Antinuclear antibody, antismooth muscle antibody, antibody to liver kidney microsomes	Alkaline phosphatase, Albumin	Anti-HAV IgM, HBsAg, anti-HBc, anti- HBsAg, anti-HCV, anti-HEV IgG and IgM, anti-VCA IgG and IgM, anti-EA IgG, anti-EBNA IgG, anti-CMV IgG and IgM, EBV PCR, and CMV PCR	
	Ferritin/ Transferrin saturation[c]	Fibrinogen, ESR[b], hsCRP	Focused Genetic Testing[c]	
Routine Urinalysis	• Specific gravity • pH, glucose, protein, blood, ketones, bilirubin, urobilinogen, nitrite, leukocyte esterase by dipstick • Microscopic examination (if blood or protein is abnormal). • βhCG (women only)			

Laboratory Assessments	Parameters
Reflex Testing for HBV DNA	• HBV DNA PCR
Other Screening Tests	• Serology (HCV antibodies, HBV antibodies, HIV testing[d], HBsAg, HCV RNA PCR, QuantiFERON TB test) • FSH and estradiol (as needed if **not** a WOCBP **only**) • Serum or highly sensitive urine human chorionic gonadotropin (hCG) pregnancy test (as needed for a WOCBP). Note: Local urine testing will be standard for the protocol unless serum testing is required by local regulation or the IRB/IEC. • Ferritin and transferrin saturation • Teriflunomide levels[e] • All study-required laboratory assessments will be performed by a central laboratory, except for urine dipstick (microscopic examination done centrally), urine pregnancy, and T-SPOT.

βhCG = β-Human Chorionic Gonadotropin, CMV = Cytomegalovirus, DNA = deoxyribonucleic acid, EA = Early Antigen, EBNA = Epstein-Barr Nuclear Antigen, eGFR = Estimated Glomerular Filtration Rate, ESR = Erythrocyte Sedimentation Rate, FSH = Follicle Stimulating Hormone, HAV = Hepatitis A Virus, HBc = Hepatitis B Core Antigen, HBsAg = Hepatitis B Surface Antigen, HBV = Hepatitis B Virus, HCV = Hepatitis C Virus, HEV = Hepatitis E Virus, HFE = High Iron Fe (human hemochromatosis protein), HIV = Human Immunodeficiency Virus, hsCRP = High Sensitivity C Reactive Protein, IDMC = Independent Data Monitoring Committee, IEC = Independent Ethics Committee, IRB = Institutional Review Board, Ig = Immunoglobulin, MCH = mean corpuscular hemoglobin, MCHC = mean corpuscular hemoglobin concentration, OLE = Open Label Extension, PCR = polymerase chain reaction, TB = tuberculosis, VCA = Viral Capsid Antigen, WBC = White Blood Cell, WOCBP = Women of Childbearing Potential.

a Results will not be disclosed to the sites, Sponsor, or representative, to avoid unblinding. However, the IDMC will have access to these data as applicable.

b ESR will be analyzed locally.

c Focused genetic testing for variants that confer risk for liver diseases and/or drug-related liver injury, including but not limited to testing for variants in the HFE gene (C282Y, H63D) in the setting of abnormal ferritin/transferrin saturation values as defined in Exclusion Criterion 10.

d HIV testing will be done at Screening centrally, unless indicated otherwise by local regulations.

e Teriflunomide levels will be determined in male participants planning to enter the OLE Period.

Appendix 6 Pharmacogenetics

Use/Analysis of DNA

Genetic variation may impact a participant's response to therapy, susceptibility to, and severity and progression of disease. Variable response to therapy may be due to genetic determinants that impact study intervention absorption, distribution, metabolism, and excretion; mechanism of action of the study intervention; disease etiology; and/or molecular subtype of the disease being treated.

DNA samples may be used for research related to pharmacokinetics, safety endpoints, drug response, and treatment efficacy of evobrutinib or MS and related diseases. They may also be used to develop tests or assays, including diagnostic tests related to pharmacokinetics, safety endpoints, drug response, and treatment efficacy of evobrutinib and MS. Pharmacogenetic research may consist of the analysis of one or more candidate genes or the analysis of genetic markers throughout the genome or analysis of the entire genome (as appropriate).

The results of pharmacogenetic analyses may be reported in the CSR or in a separate study summary.

Details on processes for collection and shipment of these samples can be found in the Laboratory Manual. The Sponsor will store the DNA samples in a secure storage space with adequate measures to protect confidentiality.

Retention time and possible analysis of DNA sample after the study ends are specified in the respective ICF.

Appendix 7 Guidance for Diagnosis of PML

The safety monitoring algorithm presented in Figure 3 will be implemented in this study.

Comprehensive neurological assessments will be performed every 12 weeks at the regular study visits. Additionally, telephone interviews will be conducted to assess for new or worsening neurological symptoms and a neurological evaluation will be conducted if clinically indicated. This neurological exam will include calculation of an EDSS score at the scheduled 12-week visit or in the event of new/worsening symptoms. This exam requires that FSS also be determined. The examination to calculate the FSS includes cognitive, visual and motor assessments, as well as assessments of other neurological systems. These neurological systems are often affected by PML, and by MS as well.

Should a non-MS etiology, such as PML, be considered as a differential etiology for any change in the clinical picture (neurological symptoms and/or exam), further assessments should be done. The evaluation of PML may include a brain MRI scan and CSF analysis per the proposed treatment algorithm (Figure 3).

Action Steps if PML is suspected:

If the clinical presentation is suggestive of PML, further investigations should include brain MRI evaluation as soon as possible. If MRI evaluation reveals lesions suspicious for PML a lumbar puncture with evaluation of the CSF for the detection of JCV DNA should be undertaken. A diagnosis of PML can potentially be made by evaluating clinical and MRI findings plus the identification of JCV in the CSF.

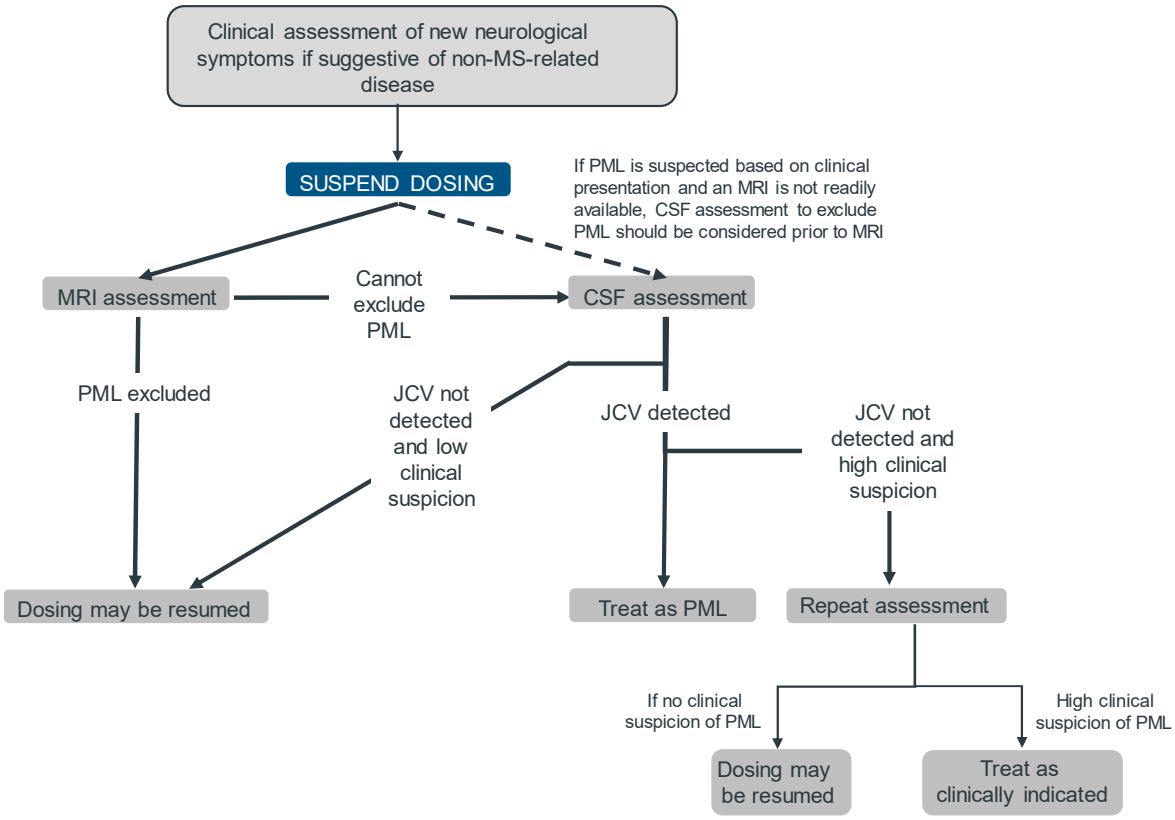
MRI Assessments

Although there are no pathognomonic findings that differentiate PML from MS, a brain MRI scan that includes fluid-attenuated inversion recovery (FLAIR) and T2-weighted and T1-weighted sequences, with and without Gd, should be performed to assess patients with neurological changes suggestive of PML.

CSF Assessment

- The detection of JCV DNA in the CSF of a participant with clinical and MRI features suggestive of PML establishes the diagnosis of PML.
- If JCV DNA is not detected in CSF and if clinical suspicion of PML remains high, a repeat lumbar puncture should be performed.
- If diagnosis remains uncertain and suspicion of PML remains high, a brain biopsy may be considered to establish a definitive diagnosis.
- CSF will be analyzed centrally.

Figure 3 Diagnostic Algorithm for PML – Suggested Diagnostic Algorithm



CSF = cerebrospinal fluid, JCV = JC virus, MRI = Magnetic Resonance Imaging, MS = multiple sclerosis, PML = progressive multifocal leukoencephalopathy.

Appendix 8 Procedure for Accelerated Elimination of Teriflunomide

Teriflunomide is eliminated slowly from the plasma based on the teriflunomide SmPC and teriflunomide USPI. Without an accelerated elimination procedure, it takes on average 8 months to reach plasma concentrations < 0.02 mg/L, although because of individual variations in drug clearance it may take as long as 2 years. An accelerated elimination procedure could be used at any time after discontinuation of teriflunomide. Elimination can be accelerated by either of the following procedures:

- Administration of cholestyramine 8 g every 8 hours for 11 days. If cholestyramine 8 g three times a day is not well tolerated, cholestyramine 4 g three times a day can be used.
- Administration of 50 g oral activated charcoal powder every 12 hours for 11 days.

If either elimination procedure is poorly tolerated, treatment days do not need to be consecutive unless there is a need to lower teriflunomide plasma concentration rapidly. At the end of 11 days, both regimens successfully accelerated teriflunomide elimination, leading to more than 98% decrease in teriflunomide plasma concentrations. Use of the accelerated elimination procedure may potentially result in return of disease activity if the participant had been responding to treatment with teriflunomide.

Appendix 9 Teriflunomide Drug-Drug Interactions

Teriflunomide has the potential for drug interactions and due to the blinded treatment assignment, these should be considered when treating participants. The following are the potential drug interactions as defined in the teriflunomide USPI (teriflunomide USPI).

DRUG INTERACTIONS

Effect of AUBAGIO on CYP2C8 Substrates

- Teriflunomide is an inhibitor of CYP2C8 in vivo. In patients taking AUBAGIO, exposure of drugs metabolized by CYP2C8 (e.g., paclitaxel, pioglitazone, repaglinide, rosiglitazone) may be increased. Monitor these patients and adjust the dose of the concomitant drug(s) metabolized by CYP2C8 as required.

Effect of AUBAGIO on Warfarin

- Coadministration of AUBAGIO with warfarin requires close monitoring of the INR because AUBAGIO may decrease peak INR by approximately 25%.

Effect of AUBAGIO on Oral Contraceptives

- AUBAGIO may increase the systemic exposures of ethinylestradiol and levonorgestrel. Consideration should be given to the type or dose of contraceptives used in combination with AUBAGIO.

Effect of AUBAGIO on CYP1A2 Substrates

- Teriflunomide may be a weak inducer of CYP1A2 in vivo. In patients taking AUBAGIO, exposure of drugs metabolized by CYP1A2 (e.g., alosetron, duloxetine, theophylline, tizanidine) may be reduced. Monitor these patients and adjust the dose of the concomitant drug(s) metabolized by CYP1A2 as required.

Effect of AUBAGIO on Organic Anion Transporter 3 (OAT3) Substrates

- Teriflunomide inhibits the activity of OAT3 in vivo. In patients taking AUBAGIO, exposure of drugs which are OAT3 substrates (e.g., cefaclor, cimetidine, ciprofloxacin, penicillin G, ketoprofen, furosemide, methotrexate, zidovudine) may be increased. Monitor these patients and adjust the dose of the concomitant drug(s) which are OAT3 substrates as required.

Effect of AUBAGIO on BCRP and Organic Anion Transporting Polypeptide B1 and B3 (OATP1B1/1B3) Substrates

- Teriflunomide inhibits the activity of BCRP and OATP1B1/1B3 in vivo. For a patient taking AUBAGIO, the dose of rosuvastatin should not exceed 10 mg once daily. For other substrates of BCRP (e.g., mitoxantrone) and drugs in the OATP family (e.g., methotrexate, rifampin), especially HMG-Co reductase inhibitors (e.g., atorvastatin, nateglinide, pravastatin, repaglinide, and simvastatin), consider reducing the dose of these drugs and monitor patients closely for signs and symptoms of increased exposures to the drugs while patients are taking AUBAGIO.

DRUG INTERACTION STUDIES

- Teriflunomide is not metabolized by Cytochrome P450 or flavin monoamine oxidase enzymes.

The potential effect of AUBAGIO on other drugs:

- CYP2C8 substrates
    - There was an increase in mean repaglinide C_{max} and AUC (1.7- and 2.4-fold, respectively) following repeated doses of teriflunomide and a single dose of 0.25 mg repaglinide, suggesting that teriflunomide is an inhibitor of CYP2C8 in vivo. The magnitude of interaction could be higher at the recommended repaglinide dose.
- CYP1A2 substrates
    - Repeated doses of teriflunomide decreased mean C_{max} and AUC of caffeine by 18% and 55%, respectively, suggesting that teriflunomide may be a weak inducer of CYP1A2 in vivo.
- OAT3 substrates
    - There was an increase in mean cefaclor C_{max} and AUC (1.43- and 1.54-fold, respectively), following repeated doses of teriflunomide, suggesting that teriflunomide is an inhibitor of organic anion transporter 3 (OAT3) in vivo.
- BCRP and OATP1B1/1B3 substrates
    - There was an increase in mean rosuvastatin C_{max} and AUC (2.65- and 2.51-fold, respectively) following repeated doses of teriflunomide, suggesting that teriflunomide is an inhibitor of BCRP transporter and organic anion transporting polypeptide 1B1 and 1B3 (OATP1B1/1B3).
- Oral contraceptives
    - There was an increase in mean ethinylestradiol C_{max} and AUC0-24 (1.58- and 1.54-fold, respectively) and levonorgestrel C_{max} and AUC0-24 (1.33- and 1.41-fold, respectively) following repeated doses of teriflunomide.
- Teriflunomide did not affect the pharmacokinetics of bupropion (a CYP2B6 substrate), midazolam (a CYP3A4 substrate), S-warfarin (a CYP2C9 substrate), omeprazole (a CYP2C19 substrate), and metoprolol (a CYP2D6 substrate).

The potential effect of other drugs on AUBAGIO

- Potent CYP and transporter inducers: Rifampin did not affect the pharmacokinetics of teriflunomide.

Appendix 10 Sponsor Signature Page

Study Title:	A Phase III, Multicenter, Randomized, Parallel Group, Double Blind, Double Dummy, Active Controlled Study of Evobrutinib Compared with Teriflunomide, in Participants with Relapsing Multiple Sclerosis to Evaluate Efficacy and Safety.
Regulatory Agency Identifying Numbers:	EudraCT: 2019-004980-36 US IND: 129428
Clinical Study Protocol Version:	Version 1.0 / 13 February 2020

I approve the design of the clinical study:

Signature

02/14/2020
Date of Signature

Name, academic degree:	Matthew Mandel, MD
Function/Title:	Global Clinical Development/Medical Director
Institution:	EMD Serono Research & Development Institute, Inc.
Address:	Global Clinical Development Center - Neurology EMD Serono Research & Development Institute, Inc. 45A Middlesex Turnpike, Billerica, MA 01821, USA
Telephone number:	+1 781 427 4250
Fax number:	Not Applicable
E-mail address:	matthew.mandel@emdserono.com

Appendix 11 Coordinating Investigator Signature Page

Study Title:	A Phase III, Multicenter, Randomized, Parallel Group, Double Blind, Double Dummy, Active Controlled Study of Evobrutinib Compared with Teriflunomide, in Participants with Relapsing Multiple Sclerosis to Evaluate Efficacy and Safety.
Regulatory Agency Identifying Numbers:	EudraCT: 2019-004980-36 US IND: 129428
Clinical Study Protocol Version:	Version 1.0 / 13 February 2020
Site Number:	

I approve the design of the clinical study, am responsible for the conduct of the study at this site, and understand and will conduct it per the clinical study protocol, any approved protocol amendments, International Council on Harmonisation Good Clinical Practice (Topic E6) and all applicable Health Authority requirements and national laws.

Signature

Feb 18, 2020

Date of Signature

Name, academic degree:	Professor Xavier Montalban, MD, PhD
Function/Title:	Professor
Institution:	Vall d'Hebron University Hospital and Research Institute
Address:	Pg. Vall d'Hebron, 119-129 08035 Barcelona, Spain
Telephone number:	(+34) 932-746-202/ ext 6202
Fax number:	(+34) 932-746-084
E-mail address:	MontalbanX@smh.ca

Appendix 12 Principal Investigator Signature Page

Study Title:	A Phase III, Multicenter, Randomized, Parallel Group, Double Blind, Double Dummy, Active Controlled Study of Evobrutinib Compared with Teriflunomide, in Participants with Relapsing Multiple Sclerosis to Evaluate Efficacy and Safety.
Regulatory Agency Identifying Numbers:	EudraCT: 2019-004980-36 US IND: 129428
Clinical Study Protocol Version:	Version 1.0 / 13 February 2020
Site Number:	

I am responsible for the conduct of the study at this site, and understand and will conduct it per the clinical study protocol, any approved protocol amendments, International Council on Harmonisation Good Clinical Practice (Topic E6) and all applicable Health Authority requirements and national laws.

I also understand that Health Authorities may require the Sponsors of clinical studies to obtain and supply details about ownership interests in the Sponsor or Investigational Medicinal Product and any other financial ties with the Sponsor. The Sponsor will use any such information solely for complying with the regulatory requirements. Therefore, I agree to supply the Sponsor with any necessary information regarding ownership interest and financial ties including those of my spouse and dependent children, and to provide updates as necessary to meet Health Authority requirements.

_____ _____

Signature Date of Signature

Name, academic degree:	
Function/Title:	
Institution:	
Address:	
Telephone number:	
Fax number:	
E-mail address:	