STATISTICAL ANALYSIS PLAN

Protocol title:	**A Phase 2, double-blind, randomized, placebo-controlled study assessing efficacy and safety of SAR441344, a CD40L-antagonist monoclonal antibody, in participants with relapsing multiple sclerosis.**
Protocol number:	**ACT16877**
Compound number (INN/Trademark):	**SAR441344** **Not applicable**
Study phase:	**Phase 2**
Short Title:	**Proof-of-concept study for SAR441344 in relapsing multiple sclerosis**
Statistician:	[redacted]
Statistical project leader:	[redacted]
Date of issue:	**18-Oct-2022**
Regulatory agency identifier number(s):	
IND:	138742
EudraCT:	2020-004785-19
NCT:	NCT04879628
WHO:	U1111-1260-3962
EUDAMED:	Not applicable
Other:	Not applicable

Total number of pages: 46

TABLE OF CONTENTS

LIST OF TABLES

VERSION HISTORY

This statistical analysis plan (SAP) for Study ACT16877 is based on the protocol dated 20-May-2021. This section summarizes major changes to the statistical analysis features in the SAP.

The first participant was randomized on 29-Jun-2021. This SAP is approved before the analysis of the 12-week double-blind period is conducted (early analysis).

Major changes in statistical analysis plan

SAP Version	Approval Date	Changes	Rationale
1	28-Jun-2022	Change: In the primary analysis, baseline score of activation of CD40/CD40L pathway and baseline score of activation of CD40/CD40L pathway by treatment interaction will not be included as covariates. Baseline GdE T1 lesion count by treatment interaction will be deleted as well. They will be included as covariates in a supplementary analysis.	Due to missing data and lack of evidence for the prognostic value of baseline score of activation of CD40/CD40L pathway. Interactions not included in the scope of the primary analysis.
		Clarification: In the objectives and endpoints Section of the protocol, the change in PGIC-Fatigue from baseline over time should be read as PGIC-Fatigue value at each planned visit.	PGIC already represents a change in itself.
		Correction: PQATv3 items will be described at W12.	No descriptive summaries of PQATv3 scores by treatment arm at Week 24.
		Addition: SAR441344 will be assessed additionally in a subgroup of participants with baseline CD40/CD40L pathway activation score above a pre-defined threshold (Section 3.7.2).	A threshold will be determined based on the baseline activation score distribution (percentile).
		Change: Exploratory subgroup analyses will be done using the same model and population as that utilized for the primary analysis within each subgroup and not in a model including CD40/CD40L pathway activation score entered as an effect (score above threshold Y/N).	To assess the treatment effect within each subgroup separately.
		AESI category “Tuberculosis or initiation of medications for suspected tuberculosis” was renamed into “Tuberculosis or suspected tuberculosis leading to initiation of medications”	Clarification for AE coding and analysis purposes

SAP Version	Approval Date	Changes	Rationale
2	11-Oct-2022	Addition: All Part A descriptive summaries (including disposition, demography and baseline characteristics, and safety analyses) will include a pooled placebo arm (ie, pooling SC and IV placebo arms)	Have a thorough description of placebo participants
		Change: Inclusion of MRIs done until 3 days after IMP administration at Week 12 in the MRI endpoints analyses (Section 5.4).	Include all available MRI results judged clinically clinically evaluable
		Change: For the count of new lesions at Week 12 and Week 24, inclusion of all Week 12 (and Week 24) MRI results performed within 20 to 40 days after the previous MRI will be considered (Section 3.2.1 & Section 3.4.1)	Include all available MRI results judged clinically evaluable
		An additional cut-off of 0.25 will be used for LRV (ie, 75% relative reduction to placebo) (Section 3.2.2)	To check the outcome of the decision using a smaller LRV
		Addition: A statistical model has been added to include additionally baseline score of activation of CD40/CD40L pathway as covariate (Section 3.2.3)	To assess the prognostic effect of baseline score of activation of CD40/CD40L pathway
		Addition: Sensitivity analysis excluding MRIs done between 1 to 3 days after IMP administration at Week 12 (Section 3.2.4)	To assess the impact of the inclusion of the MRIs on the treatment effect
		Addition: Plots have been added for MRI parameters (Section 3.4.2) and biomarkers (Section 3.7.1.4.3)	To provide a complete description of some parameters
		Efficacy and biomarker parameters collected over both double-blind and open-label periods will be described from double-blind baseline to last open-label visit, in OLE population (in addition to a description on Part A only).	To describe the evolution of parameters on a long-time period
		Addition: CXCL13 endpoint will be analyzed at Week 12 using an ANCOVA (possibly after log-transformation or rank analysis). Boxplots of raw data and absolute change from baseline will also be provided for all biomarkers. Percent changes from baseline will be presented as well.	To assess SAR441344 groups differences with placebo
		Change: Most common AEs will be selected AEs using a cut-off of >=4% in at least one treatment group	Use a more stringent cut-off to restrict this TEAEs table to the most frequent AEs
		Addition: data handling rules for the analysis of biomarkers (Section 3.7.1.4)	Exclude unreliable results
		Clarification: No individual listing will be provided for Double-blind data analysis (Section 3.8)	To maintain the blind outside of Biostatistics&Programming
3	18-Oct-2022	Change: In the primary analysis and other similar analysis, baseline GdE T1 lesion count will be included as a categorical covariate instead of continuous. A	Account for the outliers in the distribution of baseline Gde T1 count

SAP Version	Approval Date	Changes	Rationale
		sensitivity analysis based on baseline GdE T1 lesion count as a continuous covariate will be done.	

1 INTRODUCTION

1.1 STUDY DESIGN

This is a multicenter, multinational, randomized, placebo-controlled, double-blind, 4 parallel-group, Phase 2 study, including an open-label extension in which placebo arms participants switch to the SAR441344 arm corresponding to the route of administration they have been randomized to.

After a screening phase of up to 4 weeks, participants are centrally randomized (using permuted block randomization schedule) via Interactive Response Technology (IRT) in a 1:1:4:4 ratio to 1 of the 4 intervention groups: Placebo SC q2w, Placebo IV q4w, SAR441344 300 mg SC q2w, or SAR441344 1200 mg IV q4w, and treated double-blind for 12 weeks (Part A). The route of administration (IV or SC) and dose is open-label, while the treatment assignment (SAR441344 or Placebo) is double-blind. After 12 weeks of double-blind treatment, all participants switch to an open-label SAR441344 treatment period of 8 to 76 weeks expected duration for individual participants (Part B). 129 participants were randomized from 37 sites.

Study primary analysis will be conducted after completion of the 12-week double-blind period (Part A).

1.2 OBJECTIVES AND ENDPOINTS

The study objectives are detailed in Table 1 below.

Table 1 - Objectives and endpoints

Objectives	Endpoints
Primary	
• To determine the efficacy of SAR441344 as measured by reduction of the number of new active brain lesions	• Number of new gadolinium (Gd)-enhancing T1-hyperintense (GdE T1) lesions at Week 12 as measured by brain magnetic resonance imaging (MRI)
Secondary	
• To evaluate efficacy of SAR441344 on disease activity as assessed by other MRI measures	• Number of new or enlarging T2 lesions at Week 12 • Total number of GdE T1 lesions at Week 12
• To evaluate the safety and tolerability of SAR441344	• Adverse events (AEs), serious adverse events (SAEs), potentially clinically significant abnormalities (PCSAs) in laboratory tests, electrocardiogram (ECG), and vital signs during the study period • Anti-drug antibodies (ADAs)
• To evaluate pharmacokinetics of SAR441344	• SAR441344 plasma concentrations over time. Pharmacokinetic (PK) parameters (maximum concentration [C_{max}], time to C_{max} [t_{max}], area under the curve over the dosing interval [$AUC_{0\text{-}tau}$], and elimination half-life [$t_{1/2}$])

Objectives	Endpoints
Exploratory	
• To evaluate efficacy of SAR441344 on disease activity, assessed by clinical, imaging measures, and patient-reported outcomes	• Change from baseline over time in volume, number, and intensity (T1) of slowly evolving lesions (SEL) • Change in number of phase rim lesions in susceptibility weighted imaging (SWI) MRI from baseline over time (subset of centers with capacity of 3 Tesla MRI) • Change in magnetization transfer imaging (MTR) of GdE T1 lesions from Week 8 over time (selected centers) • Number of new GdE T1 lesions over time • Change in volume of T2 lesions from baseline over time • Change in brain volume, including regional changes, from baseline over time • Change in total number of T1-hypointense lesions from baseline over time • Proportion of participants with no new MRI disease activity at the end of 12 weeks of treatment and over time • Number of new or enlarging T2 lesions over time • Number of relapses (annualized relapse rate) over 12 weeks of treatment and over time • Proportion of relapse-free participants at the end of 12 weeks of treatment and over time • Change in EDSS from baseline over time • Change in MSIS-29 physical and psychological domains scoring from baseline over time • Change in PROMIS-Fatigue-MS-8 scoring from baseline over time • Descriptive summaries of PQATv3 scores by treatment arm at Week 12 • PGIC-Fatigue from baseline over time • Change in PGIS-Fatigue from baseline over time
• To explore genetic and plasma-based biochemical biomarkers that correlate with disease pathophysiology	• Analysis of messenger ribonucleic acid (mRNA) signature of cluster of differentiation 40/cluster of differentiation-40 ligand (CD40/CD40L) pathway activation in blood • Change in lymphocyte phenotype subset at the EOS compared to baseline • Change in plasma neurofilament light chain (NfL) through EOS compared to baseline • Change in serum chitinase-3-like 1 (CHI3L1) through EOS compared to baseline • Change in sTREM2 through EOS compared to baseline • Change in CXCL13 through EOS compared to baseline • Change in IgG, IgM through EOS compared to baseline Soluble CD40L (sCD40L) in plasma (at baseline, Week 12, and during 24 weeks of follow-up after EOT) • Presence of SNPs in genes related to CD40/CD40L signaling pathway or MS disease

2 ANALYSIS POPULATIONS

The following populations for analyses are defined

Table 2 - Populations for analyses

Population	Description
Screened	All participants who signed the ICF.
Randomized	All participants from screened population who have been allocated to a randomized intervention by IRT regardless of whether the intervention was received.
Intent-to-treat (ITT)	All randomized participants. Participants will be analyzed according to the intervention allocated by randomization.
Efficacy	All participants from the randomized population who take all Part A doses[a] of study intervention (one SC dose skipped is allowed) and with an evaluable primary endpoint. The primary endpoint is evaluable when the following conditions are met: • Availability of the blinded MRI central assessment of the primary endpoint[b]. • No systemic corticosteroids[c] administration 30 days before baseline MRI and during Part A period. Participants will be analyzed according to the intervention they actually received.
Safety	All randomized participants who take at least 1 dose[d] (regardless of the amount) of study intervention. Participants will be analyzed according to the intervention they actually received.
Pharmacokinetic (PK)	All randomized and treated participants (safety population) with at least one post-baseline PK sample with adequate documentation of dosing and sampling dates and times. Participants will be analyzed according to the intervention they actually received.
Open-Label Extension (OLE)	All participants who received at least one dose of intervention during the open-label extension phase.
ADA	All participants from the Safety population treated with SAR441344 in Part A with at least one post-baseline ADA result (positive, negative or inconclusive). Participants will be analyzed according to the intervention they actually received.
ADA Open-Label Extension (ADA-OLE)	All participants from the Open-Label Extension population initially randomized to placebo with at least one post-baseline ADA result (positive, negative or inconclusive) in Part B. Participants will be analyzed according to the intervention they actually received.

Note:
a From baseline to Week 10
b Assessment of new GdE T1 lesions available at Week 12 (see Table 8 for time windows definition)
c ATC level 2 =" CORTICOSTEROIDS FOR SYSTEMIC USE"
d In Part A

Participants exposed to study intervention before or without being randomized will not be considered randomized and will not be included in any analysis population. The safety experience of these participants will be reported separately.

Randomized participants for whom it is unclear if they took the study intervention will be considered as exposed and will be included in the safety population as randomized.

For any participant randomized more than once, only the data associated with the first randomization will be used in any analysis population. The safety experience associated with any later randomization will be reported separately.

For participants receiving more than one study intervention during the study, the intervention group for as-treated analyses will be the intervention group to which the participant was the most frequently exposed between Day 1 and Week 12 IMP administrations (included).

The intervention group to which the participant is considered to be exposed at a given administration is detailed below:

Table 3 - As-treated intervention assignment by administration

IMP administration visit	Actual administration	Intervention group[a] to which the participant is considered to be exposed for that administration
Day 1	6 SAR441344 vials +/- placebo vials	IV SAR441344
	only placebo vials	IV placebo if randomized in IV
		SC placebo if randomized in SC
	2 SAR441344 vials +/- placebo vials	SC SAR441344
	None of above cases	As randomized arm with route entered in the CRF
Other Part A visits	4 SAR441344 vials +/- placebo vials	IV SAR441344
	only placebo vials	IV placebo if randomized in IV
		SC placebo if randomized in SC
	1 SAR441344 vial +/- placebo vials	SC SAR441344
	None of above cases	As randomized arm with the route entered in the CRF
Week 12	6 SAR441344 vials +/- placebo vials	IV placebo
	4 SAR441344 vials +/- placebo vials	IV SAR441344
	2 SAR441344 vials +/- placebo vials	SC placebo
	1 SAR441344 vial +/- placebo vials	SC SAR441344
	None of above cases	As randomized arm with route entered in the CRF

a Note: route of administration is based on the route entered in the CRF

3 STATISTICAL ANALYSES

3.1 GENERAL CONSIDERATIONS

Study primary analysis will be conducted after completion of the 12-week double-blind period, using all complete and cleaned data from the 12-week double-blind period, but prior to the formal completion of the study (early analysis after database lock on the 12-week double-blind). In addition, the main efficacy and safety data (see Section 3.8) available in open-label study period will be presented descriptively in the early analysis after database lock on the 12-week double-blind. Besides, this analysis will be completed by a purely descriptive analysis of the open-label study period that will be conducted after formal completion of the study and final database lock.

In general, continuous data will be summarized using the number of observations available, mean, standard deviation (SD), median, [Q1, Q3,] minimum, and maximum. Categorical and ordinal data will be summarized using the count and percentage of participants.

The baseline value is defined as the last available value before the day of the first dose of double-blind investigational medicinal product (IMP). For participants randomized but not treated, the baseline value is defined as the last available value before randomization.

Unless otherwise specified, analyses will be performed by randomized study intervention group and in pooled placebo arms ie, SC Placebo, IV Placebo, Placebo (pooling SC and IV placebo arms), SC SAR441344 300 mg, IV SAR441344 1200 mg (and overall, for baseline and demographics characteristics).

Observation period

All data collected until the participant's first IMP administration in the open-label period IMP (excluded) if any, or otherwise up to last double-blind IMP administration + 180 days will be considered in the statistical analyses of Part A (12-week double-blind period).

Data collected after the participant's first IMP administration in the open-label period IMP (included) will be considered in the statistical analyses of Part B (open-label period).

The observation period will be divided into 4 segments:

- The **pre-treatment period** is defined as the period up to first IMP administration.
- The **treatment-emergent (TE) period** is defined as the period from the first IMP administration to the last IMP administration + 180 days. The treatment-emergent period includes the following 2 periods:
  - The **on-treatment period** is defined as the period from the first IMP administration to the last administration of the IMP + 24 days for the SC group or + 38 days for the IV group (maximal administration interval allowed by the protocol)
  - The **residual treatment period** is defined as the period from the end of the on-treatment period to the end of the treatment-emergent period.

- The **post-treatment period** is defined as the period from the end of the treatment-emergent period.

3.2 PRIMARY ENDPOINT(S) ANALYSIS

3.2.1 Definition of endpoint(s)

The primary endpoint is the number of new gadolinium (Gd)-enhancing T1-hyperintense (GdE T1) lesions at Week 12 relative to Week 8 as measured by brain magnetic resonance imaging.

MRI readings are performed by NeuroRX who will provide the number of new GdE T1 lesions at Week 12 relative to previous visit (protocol Week 8). Analysis time window for Week 12 is defined in Section 5.4. Only Week 12 MRIs performed within 20 to 40 days after the previous MRI will be considered in the primary endpoint analyses.

Intervention groups considered in the analysis of the primary endpoint will be: Placebo (pooling SC and IV placebo arms), SC SAR441344 300 mg, and IV SAR441344 1200 mg.

3.2.2 Main analytical approach

The primary analysis of the number of new GdE T1 lesions at Week 12 (count data endpoint) at Week 12 relative to Week 8 will be performed in the Efficacy Population through a negative binomial regression model.

This model will include the baseline GdE T1 lesion activity (presence/absence) and treatment as factors. Placebo will be used as the reference in the model. In order to account for the varying follow-up time of participants between Week 8 and Week 12, the duration (in months) between the Week 12 MRI and the previous MRI will be also taken into account as an offset variable after natural logarithm transformation.

The goodness of fit of the negative binomial regression model to the primary endpoint data will be assessed through the comparison of the model deviance with the degrees of freedom. Convergence status will be checked as well. In case of model fitting issue, other options such as eg, an exact Poisson regression model will be used.

The mean number of new GdE T1 lesions per month at Week 12 in each intervention group, as well as the relative reduction under treatment with SAR441344 when compared to placebo will be estimated based on the regression coefficients. Exponential of LSmeans estimates, rate ratio estimates (ie, relative treatment effect) and associated 95% CIs will be provided. The LS-means coefficients will be proportional to those found in the analysis data set (OBSMARGINS SAS option).

The estimated mean count of new GdE T1 lesions per month at Week 12 by intervention group will be displayed graphically using a bar chart.

For Quantitative Decision-Making methodology, the asymmetric CI of the estimate of the rate ratio in each SAR441344 group compared to placebo will also be obtained from the negative binomial regression model (or eg, exact Poisson regression model in case of model fitting issues). Using a significance criterion of 86% (P[Non-negative conclusion|LRV]≈14%), upper limits of these asymmetric CIs will correspond to upper limits of one-sided 86% CIs. Using a relevance criterion of 15% (P [Negative conclusion|TV]≈15%), the lower limits of the asymmetric CIs will correspond to lower limits of one-sided 85% CIs. The evaluation will be based on cut-offs described in Table 4.

Table 4 - Conclusion rules

Lower limit of the asymmetric confidence interval	Upper limit of the asymmetric confidence interval	Outcome
< TV	< LRV	Positive
< TV	≥ LRV	Intermediate
≥ TV	< LRV	Intermediate
≥ TV	≥ LRV	Negative

TV=0.1 (ie, 90% relative reduction to placebo) – For LRV, two cut-offs will be used: LRV=0.3 (ie, 70% relative reduction to placebo) as planned in the protocol and additionally LRV=0.25 (ie, 75% relative reduction to placebo)

Note: Subpopulation analyses described in Section 3.7.2 will be performed whatever the outcome of the decision made in the whole Efficacy population.

3.2.3 Sensitivity analysis

The treatment effect will be estimated in a model including baseline GdE T1 lesion count as a continuous covariate, which is likely to be associated with the primary endpoint. The primary analysis model will be rerun on Efficacy Population but adjusting for the continuous covariate baseline GdE T1 lesion count instead of the categorical covariate baseline GdE T1 lesion activity. This new model will include the baseline GdE T1 lesion count as covariate and treatment as factor. For model fitting check purposes, a graphical display of the predicted and observed monthly counts (Y-axis) depending on the baseline GdE T1 lesion count (X-axis) will be performed, in the Efficacy Population, with one line by intervention group.

The treatment effect will be estimated in a model including the baseline score of activation of CD40/CD40L pathway, which is also likely to be associated with the primary endpoint. A negative binomial regression including the baseline GdE T1 activity (presence/absence) and the baseline score of activation of CD40/CD40L pathway as covariates, and treatment as factor will be done on the Efficacy Population.

To assess the impact of a high proportion of participants with no new GdE T1 lesions at Week 12 (0 counts), a zero-inflated negative binomial model will be tested on the Efficacy Population. Our study population may be perceived as a mixture of patients with high probability to have only zero lesions under treatment and patients who may have zero or more new lesions, which could be related to the presence of T1 lesions at baseline as this is known as prognostic for the occurrence of new lesions. The zero-inflated negative binomial model will consist of two parts:

- A logistic regression in order to model the probability of zero new lesions associated with the presence of GdE T1 lesions at baseline (Y/N)
- A negative binomial regression to model the count of new GdE T1 lesions at Week 12. The negative binomial regression model will include the same covariates as the primary analysis model (See Section 3.2.2). Rate ratio estimates (ie, relative treatment effect) and associated 95% CIs will be provided, based on the negative binomial regression part of the model.

Impact of the activation of CD40/CD40L pathway

The impact of the activation of CD40/CD40L pathway on SAR441344 efficacy will be assessed through the baseline score of activation of CD40/CD40L pathway by treatment interaction.

Note: For more details on how the score of activation of CD40/CD40L pathway is obtained, please refer to Section 3.7.1.4.1.

More specifically, the number of new GdE T1 lesions at Week 12 (count data endpoint) will be analyzed in the Efficacy Population through a negative binomial regression including the baseline GdE T1 activity (presence/absence) and treatment as factors, and term for baseline score of activation of CD40/CD40L pathway by treatment interaction.

A graphical display of the predicted monthly rate (and 95% CI) (Y-axis) depending on CD40/CD40L pathway activation score (X-axis) will be performed, in the Efficacy Population, by intervention group. The SAS OBSMARGINS option will be used in the calculation of the predicted monthly rate.

Other analyses assessing the relationship between activation of CD40/CD40L pathway and SAR441344 efficacy and related graphical displays will be provided in Section 3.7.1.4.2, in the subsection “Determination of a threshold in the activation score”.

3.2.4 Supplementary analyses

Analyses in ITT population

For the following analysis in the ITT Population, MRI assessments for participants who received systemic corticosteroids (ATC level 2=” CORTICOSTEROIDS FOR SYSTEMIC USE”) within 30 days of the MRI assessment date will be excluded from the analyses.

The comparison between intervention groups in terms of participants without new GdE T1 lesions at Week 12 relative to Week 8 will be done in the ITT population through a logistic regression model considering participants with no available assessment for the primary endpoint (missing or measured after corticosteroids intake) or with ≥1 new lesions at Week 12 as failures (composite intercurrent event handling strategy). This model will include the same factors as that included in the primary analysis model. Odds ratios of success (success = “no new GdE T1 lesions at Week 12”) versus placebo will be provided for each SAR441344 group (estimates and 95% CIs).

Missing data assessment

As detailed in Section 3.4.2, descriptive statistics and a plot for number of new GdE T1 lesions will be provided. Besides, the reason for non-available primary endpoint will be provided in the table (including Premature study discontinuation, Systemic corticosteroids intake within 30 days prior to the assessment, SARS-CoV-2 infection, Emergency situation related to Ukranian war, Other).

Analyses excluding all Week 12 MRIs done later than the first open-label administration (the day after)

To assess the impact of MRIs done after first open-label administration at Week 12, the primary analysis will be rerun in the Efficacy population after exclusion of all Week 12 MRIs done after the study day corresponding to the first open-label administration for participants who entered in the open-label period (ie, MRIs done 1 to 3 days after the first open-label administration will be excluded).

3.3 SECONDARY ENDPOINT(S) ANALYSIS

The secondary endpoints detailed in this section are the main efficacy endpoints assessed at Week 12 by MRI measures: number of new or enlarging T2 lesions at Week 12 and total number of GdE T1 lesions at Week 12.

Intervention groups considered in these secondary endpoint analyses will be: Placebo (pooling SC and IV placebo arms), SC SAR441344 300 mg, and IV SAR441344 1200 mg.

Other secondary endpoints analyses are defined in Section 3.6.2 (AE, SAE), Section 3.6.3.1 (laboratory abnormalities), Section 3.7.1.1 (PK) and Section 3.7.1.2 (immunogenicity).

3.3.1 Key/Confirmatory secondary endpoint(s)

Not applicable.

3.3.2 Supportive secondary endpoint(s)

- Number of new or enlarging T2 lesions at Week 12:

Only Week 12 MRIs performed within the analysis time window (Section 5.4) and between 20 to 40 days after the previous MRI will be considered in the analysis of the number of new or enlarging T2 lesions at Week 12.

The number of new or enlarging T2 lesions at Week 12 relative to Week 8 as measured by brain MRI (count data endpoint) will be analyzed in the Efficacy Population through a negative binomial regression model.

The similar methodology as for the primary endpoint will be applied: this model will include the baseline T2 lesion count, and treatment as factor. In order to account for the varying follow-up time of participants between Week 8 and Week 12, the duration (in months) between the Week 12 MRI and the previous MRI will be also taken into account as an offset variable after natural logarithm transformation. In case of model fitting issue, other options such as eg, an exact Poisson regression model will be used.

The estimated number of lesions per month and associated 2-sided 95% CI will be provided for each intervention group. The rate ratio and associated 2-sided 95% CI will be provided for comparing SAR441344 intervention groups to placebo.

Please refer to Section 3.2.2 for further details on the statistical analysis model.

- Total number of GdE T1 lesions at Week 12:

The total number of GdE T1 lesions at Week 12 as measured by brain MRI (count data endpoint) will be analyzed in the Efficacy Population through a negative binomial regression model.

The same methodology as for the primary endpoint will be applied: this model will include the baseline GdE T1 activity (presence/absence) and treatment as factors. Log transformed number of MRI scans between baseline and Week 12 will be the offset variable. In case of model fitting issue, other options such as eg, an exact Poisson regression model will be used.

The estimated number of lesions per scan and associated 2-sided 95% CI will be provided for each intervention group. The rate ratio and associated 2-sided 95% CI will be provided for comparing SAR441344 intervention groups to placebo.

Please refer to Section 3.2.2 for further details on the statistical analysis model.

3.4 TERTIARY/EXPLORATORY ENDPOINT(S) ANALYSIS

The tertiary endpoints detailed in this section are related to disease activity, assessed by clinical imaging measures, and patient-reported outcomes.

Results will be presented on Part A and on Part A and B. In this Section, intervention groups considered in the descriptive analyses over the double-blind period (Part A) will be: Placebo (pooling SC and IV placebo arms), SC SAR441344 300 mg, and IV SAR441344 1200 mg.

On the overall study period (over Part A and B), statistical outputs will be displayed by randomized intervention group (Placebo SC, Placebo IV, SAR441344 300 mg, and IV SAR441344 1200 mg).

Magnetic Resonance Imaging assessments for participants who received systemic corticosteroids (ATC level 2=" CORTICOSTEROIDS FOR SYSTEMIC USE") within 30 days of the MRI assessment date will be excluded from analyses (over Parts A and B).

3.4.1 Definition of endpoint(s)

MRI endpoints

- Number and volume of slowly evolving lesions (SELs) from baseline unenhancing T2 lesions at Week 12 and EOS
- Normalized T1 (nT1) intensity evolution in SELs from baseline unenhancing T2 lesions at Week 12 and EOS
- Number of phase rim lesions in susceptibility weighted imaging (SWI) MRI (subset of centers with capacity of 3 Tesla MRI) at each timepoint
- Number of new phase rim lesions in SWI MRI from baseline (subset of centers with capacity of 3 Tesla MRI) at each timepoint
- Number of new phase rim lesions from active lesions in SWI MRI from baseline (subset of centers with capacity of 3 Tesla MRI) at each timepoint
- MTR recovery at Week 24 in new MTR lesions detected at Week 8 and 12 (selected centers)
- MTR recovery at EOS in new MTR lesions detected at Week 12 (selected centers)
- Number of new GdE T1 lesions at each timepoint
- Total number of GdE T1 lesions at each timepoint
- Relative change from baseline in volume of T2 lesions at each timepoint
- Normalized brain volume at baseline and relative change from baseline in brain volume, including regional changes (thalamus and cerebral cortex), at EOS
- Number of new T1-hypointense lesions at each timepoint
- Relative change from baseline in volume of T1-hypointense lesions at each timepoint
- Proportion of participants with no new MRI disease activity over 12, 24, 48 and 72 weeks. No new MRI disease activity means that the participant has no new Gd-enhancing T1-hyperintense lesions and no new and enlarging T2 lesions at the MRI assessment during the corresponding period.
- Number of T2 lesions at baseline and number of new or enlarging T2 lesions at each timepoint

  For new lesions counts at W12 and 24 (ie, new GdE T1 lesions, new phase rim lesions, new T1-hypointense lesions and new or enlarging T2 lesions), only MRIs performed within 20 to 40 days after the previous MRI will be considered in the analyses at the considered visit.

Multiple sclerosis relapse endpoints

- Number of relapses (annualized relapse rate) over 12, 24, 48 and 72 weeks. It is defined as the average number of confirmed relapses per year, ie, the total number of confirmed relapses reported during the considered time period divided by the duration in days of the considered period multiplied by 365.25.

  The duration (in days) of the considered period for each participant will be calculated as: last available visit date within the considered period – first IMP administration date in double-blind period + 1.

  MS relapse events are clinical events that met the protocol defined criteria (Section 8.1.2.1) and are reported on the MS relapse e-CRF page. A confirmed relapse will be derived based on the EDSS score (if available) according to the criteria listed in Section 8.1.2.1 of the Protocol: an increase of at least 0.5 points in the EDSS score OR an increase of 1 point on 2 functional scores OR an increase of 2 points on 1 functional score, during the considered period. The first available EDSS score within 2 weeks after the date of onset of MS relapse will be used. If no EDSS is available within this timeframe, the MS relapse will be considered as not confirmed.

- Proportion of relapse-free participants over 12, 24, 48 and 72 weeks. It is defined as the proportion of participants with no confirmed relapses over the considered period. Each period will include all data until the last available visit date within the considered period.

- Change in EDSS score from baseline at each visit (Week 12, 24, 48 and 72).

  EDSS is an ordinal clinical rating scale which ranges from 0 (normal neurologic examination) to 10 (death due to MS) in half-point increments. EDSS consists in rating of 7 functional systems (visual, brainstem, pyramidal [motor], cerebellar [coordination], sensory, cerebral, and bowel/bladder) and ambulation. The EDSS score will be directly captured in the e-CRF.

Patient-reported outcomes

- Change in Multiple Sclerosis Impact Scale (MSIS-29) physical and psychological domains scoring from baseline at Week 12 and 24.

  The physical and psychological impact subscales of the MSIS-29 is composed of 29 items coded using a 4-point Likert scale from 1 (not at all) to 4 (extremely). They range from 0 to 100, with higher scores indicating greater physical or psychological impact. The recall period for MSIS-29 is 2 weeks before completion.

  The physical impact score is computed by summing items number 1-20 inclusive. This score can then be transformed to a score on 0-100 scale using the formula below:

$$\frac{100 \times (\text{observed score} - 20)}{80 - 20}$$

The psychological impact score is computed by summing items number 21-29 inclusive. This score can then be transformed to a score on a 0 -100 scale using the formula below:

$$\frac{100 \times (\text{observed score} - 9)}{36 - 9}$$

For respondents with missing data, but where at least 50% of the items in a domain have been completed (ie, a minimum of 10 items in the physical domain or 5 items in the psychological domain), a respondent–specific mean score computed from the completed items will be computed (1). This respondent–specific mean score will be used as the score for each missing item from the corresponding domain.

- Change in Patient Reported Outcome Measurement Information System (PROMIS)-Fatigue-MS-8 scoring from baseline at Week 12 and 24.

  The PROMIS-Fatigue-MS-8 (Cook & Al, 2012, [2]) has 8 items coded using a 5-point Likert scale from 1 (never/not at all) to 5 (almost always/very much). The recall period for PROMIS-Fatigue-MS-8 is 7 days before completion. All questions must be answered in order to produce a valid score (3). It is converted on a T-score metric, with a mean of 50 and a SD of 10. A higher PROMIS T-score represents more fatigue. The score will be calculated using a concordance table that associates summed scores on the PROMIS-Fatigue-MS with scores on the PROMIS T-score metric (see Section 5.5).

- Patient's Qualitative Assessment of Treatment version 3 (PQATv3) at Week 12

  PQATv3 comprises qualitative and quantitative item on the benefits and disadvantages of treatment, including participant preferences for IV or SC (refer to Protocol Appendix 10.10.3 for the type and list of items). Each item will be analyzed separately. Only change from baseline in quantitative items will be described as part of the CSR.

- Patient global impression of change scale (PGIC)-Fatigue at Week 12 and Week 24

  PGIC-Fatigue comprises one item measuring the change in fatigue symptoms as compared to first study visit on a 5-point Likert scale (1=Much worse, 2=A little worse, 3=No change, 4=A little better, 5=Much better)

- Change in Patient global impression of severity scale (PGIS)-Fatigue from baseline at Week 12 and Week 24

  PGIS-Fatigue comprises one item measuring the overall severity of fatigue on a 5-point Likert scale (1=None, 2=Mild,3= Moderate, 4= Severe, 5= Very severe). A higher PGIS-Fatigue value represents more fatigue severity.

3.4.2 Main analytical approach

For MRI endpoints and MS relapses endpoints (including EDSS), descriptive statistics will be summarized over time at each available timepoint by intervention group in the Efficacy population for the double-blind period and in the open-label extension (OLE) population for the overall study period (double-blind and open-label period). The count and proportion type endpoints will be summarized in categories (frequencies and percentages). The number of new GdE T1 lesions, the number of new or enlarging T2 lesions and the total number of GdE T1 will be described both in categories and as continuous variables.

Mean [+/-SD] over time in the double-blind period will be represented graphically by intervention group in the Efficacy Population for the following endpoints: the number of new GdE T1 lesions, the number of new or enlarging T2 lesions, the total number of GdE T1, the number of SELs, Phase rim lesions parameters and the number of new T1-hypointense lesions.

Boxplots over time in the double-blind period by intervention group will be done for the volume of SELs, T2 lesions and T1-hypointense lesions in the Efficacy Population.

Similar plots may be provided on the overall study period (double-blind and open-label period) in the OLE population.

For PRO endpoints, analyses will be provided in the ITT population. PGIC-Fatigue and PGIS-Fatigue will be described using a qualitative approach (frequencies and percentages for each item). Change from baseline in PGIS-Fatigue will be described quantitatively.

3.5 MULTIPLICITY ISSUES

No adjustment on multiplicity is planned.

3.6 SAFETY ANALYSES

Safety analyses will be presented by randomized intervention group (and in pooled placebo arms in the double-blind study period). Results will be presented separately for Part A and B.

All safety analyses will be performed on:

- the safety population as defined in Section 2 for the double-blind study period (Part A),
- the open-label extension (OLE) population as defined in Section 2 for the open-label study period (Part B).

Unless otherwise specified, the following common rules will be used:

- The analysis of the safety variables will be essentially descriptive, and no testing is planned.
- Safety data in participants who do not belong to the safety populations (eg, exposed but not randomized) will be listed separately.

3.6.1 Extent of exposure

The extent of IMP exposure will be assessed by the duration of IMP exposure and compliance and will be summarized by actual treatment received in the safety population for the double-blind study period (Part A) and in the open-label extension (OLE) population for the open-label study period (Part B).

Duration of IMP exposure

Duration of IMP exposure will be calculated ignoring missed injections/infusions.

Double-blind period (Part A)

For SC arms (actual arm), duration of IMP exposure in weeks is defined as (date of last injection in the double-blind period – date of first infusion (loading dose) + 14)/7.

For IV arms (actual arm), duration of IMP exposure in weeks is defined as (date of last infusion in the double-blind period – date of first infusion + 28)/7.

Open label period (Part B)

For SC arms (actual arm), duration of IMP exposure in weeks is defined as (date of last injection – date of first injection/infusion in the open-label period + 14)/7.

For IV arms (actual arm), duration of IMP exposure in weeks is defined as (date of last infusion – date of first infusion in the open-label period + 28)/7

Both study periods

For each study period and route, duration of IMP exposure will be summarized quantitatively and categorically:

- Double-blind period: 1 day to ≤4 weeks, >4 to ≤8 weeks, >8 to ≤12 weeks, and >12 weeks
- Open-label period: 1 day to ≤4 weeks, >4 to ≤8 weeks, >8 to ≤12 weeks, >12 to ≤18 weeks, >18 to ≤24 weeks, >24 to ≤36 weeks, >36 to ≤48 weeks, …, and >84 weeks

Additionally, the cumulative duration of treatment exposure defined as the sum of the duration of treatment exposure over each study period for all participants (expressed in participant-months) will be provided.

Moreover, the total number of IMP injections/infusions (with at least one vial injected) by participant will be summarized by study period (as a continuous variable).

Treatment compliance

No imputation will be made for participants with missing or incomplete data.

Compliance in SC and IV arms

The percentage of treatment compliance for a participant will be checked through the number of vials used for injections/infusions received divided by the total number of vials planned to be used for injections/infusions for the participant during the studied period (up to the actual last injection/infusion of IMP) and multiplied by 100.

Treatment compliance will be summarized over each study period, quantitatively and in categories: <80%, ≥80%.

Of note, according to the protocol, cases of overdose are reported in the e-CRF as Adverse Events and will be described as part of the AE analyses.

3.6.2 Adverse events

General common rules for adverse events

All adverse events (AEs) will be coded to a lower-level term (LLT), preferred term (PT), high-level term (HLT), high-level group term (HLGT), and associated primary system organ class (SOC) using the Medical Dictionary for Regulatory Activities (MedDRA) version currently in effect at Sanofi at the time of Part A database lock.

The AEs will be analyzed in the following 3 categories:

- Pre-treatment AEs: AEs that developed, worsened or became serious during the pre-treatment period.
- Treatment-emergent adverse events (TEAE)s: AEs that developed, worsened or became serious during the treatment-emergent period
- Post-treatment AEs: AEs that developed, worsened or became serious during the post-treatment period

Similarly, the deaths will be analyzed in the pre-treatment, treatment-emergent and post-treatment periods.

The primary AE analyses will be on TEAEs. Pre-treatment and post-treatment AEs will be described separately.

An AE with incomplete or missing date/time of onset (occurrence, worsening, or becoming serious) will be classified as a TEAE unless there is definitive information to determine it is a pre-treatment or a post-treatment AE.

If the assessment of the relationship to IMP is missing for an AE, this AE will be assumed as related to IMP. If the intensity is missing for 1 of the treatment-emergent occurrences of an AE, the intensity will be imputed with the maximal intensity of the other occurrences. Missing severity will be left as missing.

Multiple occurrences of the same event in the same participant will be counted only once in the tables within a treatment phase.

The AE tables will be sorted as indicated in Table 5.

Table 5 - Sorting of AE tables

AE presentation	Sorting rules
SOC and PT	By the internationally agreed SOC order and decreasing frequency of PTs[a, b]
SMQ/CMQ and PT	By decreasing frequency of SMQs/CMQs and PTs[a]
PT	By decreasing frequency of PTs[a]

a Sorting will be based on the incidence in SAR441344 IV intervention group
b The table of all TEAEs presented by primary SOC and PT will define the presentation order for all other tables (eg, treatment-emergent SAE) presented by SOC and PT, unless otherwise specified.

Analysis of all adverse events

The overview of TEAE with the details below will be generated:

- Any TEAE
- Any treatment emergent SAE
- TEAE leading to death
- Any TEAE leading to permanent treatment discontinuation
- Any treatment emergent AESI

The AE summaries of Table 6 will be generated with number (%) of participants experiencing at least one event. The all TEAE summary by Primary SOC and PT will be performed by trial impact (disruption) due to COVID-19.

Table 6 - Analyses of adverse events

Type of AE	MedDRA levels
All TEAE	Primary SOC and PT
	PT
Common TEAE (≥4% in at least one treatment group)	Primary SOC and PT
TEAE related to IMP as per Investigator's judgment	Primary SOC and PT
TEAE by maximal intensity	Primary SOC and PT
Treatment emergent SAE	Primary SOC and PT
Treatment emergent SAE related to IMP as per Investigator's judgment	Primary SOC and PT
TEAE leading to permanent IMP discontinuation	Primary SOC and PT
TEAE leading to death[b]	Primary SOC and PT
Pretreatment AE	Overview[a]
Post-treatment AE	Overview[a]
Post-treatment SAE	Primary SOC and PT

a Will include the following AE categories: any AEs, any serious AEs, any AEs leading to death, any AEs leading to permanent IMP discontinuation
b Death as an outcome of the AE as reported by the Investigator in the AE page

Analysis of deaths

In addition to the analyses of deaths included in Table 6 the number (%) of participants in the following category will be provided:

- Deaths during the treatment-emergent and post-treatment periods by main reason for death

Deaths in non-randomized participants or randomized but not treated participants will be listed separately

Analysis of adverse events of special interest (AESIs)

Adverse events of special interest (AESIs) will be selected for analyses as indicated in Table 7. Number (%) of participants experiencing at least one event will be provided for each event of interest, by SOC and PT. Tables will be sorted as indicated in Table 5.

Table 7 - Selections for AESIs

AESIs	Selection[a]
Increase in alanine transaminase (ALT)	e-CRF specific tick box on the AE page
Pregnancy	e-CRF specific tick box on the AE page
Symptomatic overdose (serious or nonserious) with IMP	e-CRF specific tick box on the AE page
Arterial and/or venous thrombotic or embolic event	CMQsn00081 (Embolic and thrombotic events Narrow)
Anaphylaxis	CMQsn00021 (Anaphylactic reaction Narrow)
Severe infusion related reaction	CMQsn00214 (Hypersensitivity Narrow) - AEs with severe intensity only
Severe IMP injection or infusion site reaction	CMQ00002 (injection site reaction) - AEs with severe intensity only
Severe infections including opportunistic infections	CMQ10176 (GLB_INFECTIONS) - AEs with severe intensity only
Tuberculosis or suspected tuberculosis leading to initiation of medications	CMQ10217 (GLB_TUBERCULOSIS)
Diagnosed and biologically proven SARS-CoV-2 infection	CMQsn00237 (COVID-19 Narrow)

a The list of terms may be adjusted according to MedDRA version changes

In addition, the following variables will be tabulated by intervention group for the local injection site reactions TEAEs:

- Toxicity grade of the event (mild, moderate, severe, very severe) by type of reaction (pain, tenderness, erythema/redness, swelling, itching, other)
- Number of events divided by the number of double-blind IMP administrations received for double-blind period, or by the number of open-label administrations received for open-label period

For AESIs “Arterial and/or venous thrombotic or embolic event” and “Severe infections including opportunistic infections”, a specific output will be provided including the number (%) of participants having experienced at least one adverse event and the relative risk of having an event in SAR groups as compared to placebo, for each of these two AESI categories. The percentage difference of AEs in SAR groups as compared to placebo and the exact CI of the difference will be displayed as well.

Finally, the pain verbal descriptor scale (VDS) assessing pain at injection site (self-assessment) will be summarized by visit and time-point (before IMP administration, after completion of the IMP administration, and 2 hours post-administration) for each intervention group.

Of note, multiple sclerosis relapses are exempt from being reported as AEs except when they meet the definition of an SAE. Multiple sclerosis relapses are collected on the eCRF and analyzed as part of the efficacy analyses.

3.6.3 Additional safety assessments

3.6.3.1 Laboratory variables, vital signs and electrocardiograms (ECGs)

The following laboratory variables, vital signs and electrocardiogram (ECG) variables will be analyzed in the double-blind and in the open-label study periods separately.

They will be converted into standard international units and international units will be used in all listings and tables.

- Hematology:
  - Red blood cells and platelets and coagulation: hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, red blood cell count, %Reticulocytes, platelet count
  - White blood cells: white blood cell count, neutrophils, lymphocytes, monocytes, basophils, eosinophils
- Coagulation:
  - International normalized ratio (INR)
  - Activated partial thromboplastin time (aPTT)
- Clinical chemistry:
  - Metabolism: serum glucose
  - Electrolytes: potassium, sodium, calcium, chloride, bicarbonate, phosphorous
  - Renal function: creatinine, eGFR, blood urea nitrogen (BUN). eGFR will be derived using the Modification of the Diet in Renal Disease (MDRD) equation:

    GFR (mL/min/1.73 m^2) = 175 × (Creatinine in mg/dL)-1.154 × (Age)-0.203 × (0.742 if female) × (1.212 if African American)

  - Liver function: alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, Creatine phosphokinase, total and direct bilirubin, albumin, total protein
  - Pregnancy test: Serum β-human chorionic gonadotropin (all female participants)
- Urinalysis:
  - Urinalysis for quantitative analysis: pH, specific gravity, proteins, ketones, bilirubin, urobilinogen, nitrite, erythrocytes, and leukocytes count and glucose

  Urine parameters will be summarized through listings of participants with abnormalities.
- Vital signs: heart rate, systolic and diastolic blood pressure according to position (sitting, standing, supine), weight, height (at screening only), respiratory rate, temperature.
- ECG variables: heart rate, PR, QRS duration, QT, and corrected QTc (Correction Method Unspecified as collected during on-site automatic reading, as well as correction according to Bazett and Fridericia formulae) and ECG assessments will be described as normal or abnormal. The following formulas will be used: $QTcB = QT/RR^{0.5}$ and $QTcF = QT/RR^{0.33}$, with RR=HR/60.

Data below the lower limit of quantitation/detection limit (LLOQ) will be replaced by half of the LLOQ, data above the upper limit of quantification will be replaced by ULOQ value.

For hematological, coagulation and biochemistry parameters, the central laboratory reference ranges will be used for central laboratory measurements and local reference ranges for local measurements.

Quantitative analyses

For all laboratory variables, vital signs, and ECG variables above, descriptive statistics for results and changes from baseline will be provided for each planned timepoint (after re-allocation of visits according to time-windows defined in Section 5.4), including the last value and the worst value (minimum and/or maximum value depending on the parameter) during the on-treatment period. These analyses were performed using central measurements only for laboratory variables. ECG variables are based upon the automatic reading of the device at each site.

For hemoglobin, hematocrit, neutrophils, lymphocytes, platelet count and alanine aminotransferase: mean changes from baseline with the corresponding standard error will be plotted over time by intervention group.

Analyses according to PCSA

Potentially clinically significant abnormality (PCSA) analyses will be performed based on the PCSA list currently in effect at Sanofi at the time of Part A database lock. These analyses will be performed using both central and local measurements for laboratory variables (either scheduled or unscheduled). For parameters for which no PCSA criteria are defined, similar analyses will be done using the normal range, if applicable.

Analyses according to PCSA will be performed based on the worst value during the treatment-emergent period, using all measurements (whether local or central, scheduled, nonscheduled or repeated).

For laboratory variables and vital signs above, the incidence of participants with at least one PCSA during the treatment-emergent period will be summarized regardless of the baseline level and according to the following baseline status categories:

- Normal/missing
- Abnormal according to PCSA criterion or criteria

Summary statistics (including number, mean, median, Q1, Q3, standard deviation, minimum and maximum) of all ECG variables (raw values and changes from baseline) will be calculated for each planned visit by intervention group. In addition, mean changes from baseline with the corresponding standard errors will be plotted over time in each intervention group.

For ECG, the incidence of participants with at least one abnormal ECG during the treatment-emergent period will be summarized regardless of the baseline level and according to the following baseline status categories:

- Normal/missing
- Abnormal

Additional analyses for drug-induced liver injury

The following additional analyses will be performed for drug-induced liver injury:

- Time to onset of the initial alanine aminotransferase (ALT) or aspartate aminotransferase (AST) elevation (>3 x ULN) and total bilirubin elevation (>2 x ULN) during the treatment-emergent period will be analyzed using Kaplan-Meier method.
- A graph of the distribution of peak values of ALT versus peak values of total bilirubin during the treatment-emergent period will be provided.
- For each liver function test (eg, ALT), participants having experienced a PCSA (eg, ALT >5 ULN) will be summarized using the following categories: Returned to baseline PCSA status (or returned to value ≤ ULN in case of missing baseline) before last IMP dose, Returned to baseline PCSA status after last IMP dose, Never returned to baseline PCSA status, No assessment after elevation. This summary will be performed by categories of elevation (ALT >3, >5, >10, >20 ULN).

3.7 OTHER ANALYSES

Unless otherwise specified, results from this Section will be presented separately for Part A and B. In this Section, intervention groups considered in the analyses over the double-blind period will be: Placebo (pooling SC and IV placebo arms), SC SAR441344 300 mg, and IV SAR441344 1200 mg. Over the open-label period, statistical outputs will be displayed by randomized intervention group (Placebo SC, Placebo IV, SAR441344 300 mg, and IV SAR441344 1200 mg).

3.7.1 Other variables and/or parameters

3.7.1.1 PK analyses

SAR441344 concentrations at all protocol planned time points will be reported using descriptive statistics (number of available data, mean, geometric mean, SD, median, minimum and maximum) for each SAR441344 dose group, in the PK population. Additional PK parameters such as Cmax, tmax, t1/2z and AUC at steady state will be estimated using a population PK approach. These parameters will be presented in a separate, standalone report.

3.7.1.2 Immunogenicity analyses

Participant's ADA status, response variable (see definitions below) will be summarized on the ADA population, across the whole study (over both double-blind and open-label periods).

Analyses over the whole study (Part A and B) will be performed on participants from the ADA population.

Analyses over the open-label period will be performed on participants from the ADA-OLE population.

Kinetics of ADA responses will be described over the whole study period only, for participants from the ADA population with treatment-induced ADA and for participants with treatment-boosted ADA, separately. Time to ADA onset and duration of ADA will be described with minimum, Q1, median, Q3 and maximum statistics.

Sample status (negative, positive, inconclusive) and titers will also be described overtime using descriptive statistics.

Peak titer will be described with minimum, Q1, median, Q3 and maximum statistics for participants with treatment-induced ADA and for participants with treatment-boosted ADA, separately, if any.

The impact of positive immune response on efficacy, PK and safety variables may be further explored, depending on ADA incidence.

Antidrug antibodies were assessed:

- At baseline (before the first IMP administration in the double-blind period)
- During the double-blind period at Week 4, Week 8, Week 12, or early termination and during follow-up (12 and 24 weeks after early termination) for participants not proceeding into open-label period
- During the open-label period at Week 16, Week 20, Week 24, Week 36, Week 48, Week 60, Week 72, Week 84 or early termination and during follow-up (12 and 24 weeks after early termination)

Participant's ADA status

- Participants with **pre-existing ADA**s correspond to participants with ADAs present in samples drawn before first administration of intervention. Participants with missing ADA sample at baseline will be considered as without pre-existing ADA.
- Participants with **treatment-emergent ADA** correspond to participants with at least one treatment-induced/boosted ADA.
  - Participants with **treatment-induced ADA**s correspond to participants with ADAs that developed during the treatment-emergent (TE) period and without pre-existing ADA (including participants without pre-treatment samples).
  - Participants with **treatment-boosted ADA**s correspond to participants with pre-existing ADAs that are boosted during the TE period to a significant higher titer than the baseline. A 2-fold serial dilution schema is used during titration, so at least a 4-fold increase will be considered as significant.
- Participants with **unclassified ADA** correspond to participants with pre-existing ADAs that cannot be classified as treatment-boosted ADA because of missing titer(s) (ie, a positive ADA sample during the TE period in a participant with pre-existing ADA but with missing titer at this sample or at baseline).
- Participants **without treatment-emergent ADA** correspond to participants without treatment-induced/boosted ADA and without any inconclusive sample nor unclassified ADA during the TE period.
- Participants **with inconclusive ADA** are defined as participants which cannot irrefutably be classified as with or without treatment-emergent ADA.
- Participants with **post-treatment ADA** correspond to participants with at least one treatment-induced/boosted ADA after the on-treatment period (residual and post-treatment period).

Kinetics of ADA response (over the whole treatment period in Part A and B in participants from ADA population having a duration of exposure greater than 6 months)

Kinetics of ADA response will be derived for participants with treatment-induced/boosted ADA considering ADA samples collected during the TE period and post-treatment period.

- **Time to onset of ADA response** is defined as the time period between the first IMP administration and the first treatment-induced/boosted ADA.
- **Duration of ADA response** is defined as the time between the first treatment-induced/ boosted ADA and the last treatment-induced/boosted ADA, irrespective of negative samples or positive samples not reaching the boosted threshold in-between. ADA duration will be summarized only for participants with persistent ADA response.
  - A positive sample (boosted positive sample for participants with pre-existing ADA) occurring after the TE period will be considered as treatment-induced/boosted ADA if a previous treatment-induced/boosted ADA occurred during the TE period and less than 16 weeks before this sample; indeed, a treatment-induced/boosted ADA occurring

at the end of the TE period could have positive samples after the TE period due to half-life of the immunoglobulin G

- **Persistent ADA response** is defined by treatment-induced/boosted ADA with a duration of ADA response of at least 16 weeks.
- **Transient ADA response** is defined by treatment-induced/boosted ADA with a duration of ADA response of less than 16 weeks and the last sample of the TE period is not treatment-induced/boosted.
- **Indeterminate ADA response** is defined by treatment-induced/boosted ADA that are neither persistent nor transient.

ADA response variable:

- **ADA incidence** is defined as the proportion of participants found to have seroconverted (treatment-induced ADAs) or boosted their pre-existing ADA response (treatment-boosted ADAs) at any time point during the TE period.

3.7.1.3 Quality of life analyses

Please refer to Section 3.4.1 for the description of PRO endpoints and Section 3.4.2 analytical approach (descriptive statistics only).

3.7.1.4 Biomarker analyses

Additional biomarker analyses other than described below might be described and reported separately. In particular, additional statistical analyses for RNA biomarkers will be detailed in a dedicated/separate SAP.

Specific data handling rules are needed for the analysis of biomarkers data. These are detailed in Section 5.4.

3.7.1.4.1 Score of activation of CD40/CD40L pathway

The CD40/CD40L pathway activation score ("activation score") will be estimated from the expression levels of 83 genes identified internally by Sanofi and which have been found impacted by CD40L or SAR441344 in the presence of anti-CD3/CD28 in an ex vivo stimulation study. The gene expression levels will be internally measured in Sanofi with Nanostring technology. After background correction and positive control normalization, the data will be provided as count variables.

Before estimating the activation score, quality control (QC) will be assessed based on the distribution of missing data and values under the detection limit. Variability in sample input will be corrected using the housekeeping genes from the dataset by following the Nanostring guidelines:

1. Most stable housekeeping genes will be selected

2. The geometric mean of the selected housekeeping genes will be estimated for each sample.
3. The arithmetic mean of these geometric means will then be calculated.
4. A normalizing factor will be estimated for each sample with the ratio between this arithmetic mean and the geometric mean
5. The counts for every gene will be multiplied by the sample-specific normalization factor.

Variability correction can alternatively be done using the RUV (Remove Unwanted Variation) approach with the housekeeping genes. Finally, the count data will be log2 transformed. The PCA plot of expression data will inspect potential outliers and batch effects (if relevant).

The activation score will be estimated based on the gene level expression difference between 83 genes found upregulated and downregulated by SAR441344.

The 83 genes activation score will be estimated for samples using the following steps:

1. Median centering the expression levels for each gene
2. Summing the expression levels of the genes found upregulated by the treatment
3. Summing the expression levels of the genes found downregulated by the treatment
4. Taking the difference of both the quantities and dividing by the gene number

The CD40/CD40L pathway activation scores at baseline will be estimated before the Part A database lock. Descriptive statistics of the activation scores at baseline will be provided in the Efficacy population (and ITT population as well if the number of participants in the Efficacy and ITT Populations differs more than 10%).

Once all the samples from baseline, Week 4, Week 12, and Week 24 are analyzed, the gene expression levels provided for the Part A database lock will be renormalized with the newly available data. The activation scores will therefore be re-estimated for all patients at each visit (including baseline) before the Part B database lock. Descriptive statistics of the CD40/CD40L pathway activation scores over time will be provided in the ITT (double-blind period) and OLE population (over both double-blind and open-label periods), including both the baseline calculated for the Part A database lock and baseline re-estimated at the Part B database lock.

3.7.1.4.2 Responder population identification and characterization with biomarker data

It is hypothesized that the efficacy of SAR441344 may depend on the activation of CD40/CD40L pathway. Subgroup analysis of the primary endpoint will be conducted based on the CD40/CD40L pathway activation score (please refer to Section 3.7.2 for more details).

In consequence, below exploratory analyses will be performed to assess CD40/CD40L pathway activation score as predictor of efficacy and to identify the activation score threshold above which participants would preferentially benefit from SAR441344 treatment.

As described in Section 3.2.3, the expected number of new GdE T1 lesions at Week 12 will be modeled through a negative binomial regression with the baseline GdE T1 activity

(presence/absence) and treatment as effects and score by treatment interaction, in the Efficacity Population.

The linear relationship between the new GdE T1 lesions rate and the activation score in each arm will be assessed with a plot representing the activation score against partial residuals summed with the predicted lesions count (when the other covariates are held constant). If a non-linear trend is detected, a spline on the activation score may be used in the model.

In addition, the participants will be divided into subgroups using the activation score tertiles. If there are enough placebo patients in each subset, the treatment effect compared with placebo will be estimated in each activation score subgroup. The effect will be assessed with the negative binomial regression described above but fitted within each subgroup (one model per subgroup). The rate ratio for each treatment will be calculated with a 95% CI interval. Plots of the treatment effects against the activation score subgroups will be provided to examine the variation in the treatment effect across the score.

Performance metrics for predicting the outcome will be reported, such as mean squared errors or brier score. Resampling methods such as bootstrap or cross-validation will be used to mitigate the overfitting issue. The fitted model will also be used to simulate data and compare their distributions with the observed one.

Determination of a threshold in the activation score

The non-responder patients are assumed to have a non-substantial relative reduction of new GdE T1 lesions count per month in the intervention group compared to the placebo group. The relative reduction for the non-responders is assumed to be less than 70% (and 90% as a second approach). For each relative reduction (70% and 90%), a threshold in the activation score can be calculated by solving the following equation:

$$\frac{E[GdE\ T1\ lesions\ count \mid treatment, covariates]}{E[GdE\ T1\ lesions\ count \mid placebo, covariates]} < 0.25 \qquad (or\ 0.1\ in\ the\ second\ approach)$$

Each variable can be calculated from the above model. Bootstrap 95% CI will be provided with the threshold estimation.

The rate ratio in SAR441344 groups when compared to placebo (and 95% CI) (Y-axis) depending on CD40/CD40L pathway activation score (X-axis) will be plotted, in the Efficacy Population, by intervention group. Horizontal reference lines corresponding to rate ratios of 0.1 (TV) and 0.25 (LRV) will be overlaid on the plots.

In addition, the proportion of patients with/without new lesions at Week 12 will be calculated in the two subgroups defined by the determined threshold (ie, patients having baseline activation score of CD40/CD40L pathway below and above the threshold at baseline).

Other signature(s) / threshold(s) may also be defined in the Biomarker Statistical Analysis Plan.

3.7.1.4.3 Blood-based biochemical biomarkers

- Change in plasma neurofilament light chain (NfL) at Week 12, Week 24… through EOS, as compared to baseline
- Change in plasma chitinase-3-like 1 (CHI3L1) at Week 12, Week 24…through EOS compared to baseline
- Change in plasma sTREM2 at Week 12, Week 24…through EOS compared to baseline
- Change in plasma CXCL13 at Week 12, Week 24…through EOS compared to baseline
- Change in plasma sCD27 at Week 12, Week 24…through EOS compared to baseline
- Change in serum immunoglobin levels (IgG, IgM) at Week 8, Week 12, Week 24…through EOS compared to baseline
- Soluble CD40L (sCD40L) in plasma raw value and change from baseline by visit (baseline, Week 12, and during 24 weeks of follow-up after EOT)

Descriptive statistics

Descriptive statistics (mean, SD, SEM, median, Q1, Q3, and range) will be provided for these endpoints at each available timepoint by intervention group in the Safety population for the double-blind period ("as-treated" approach) and in the OLE population over both the double-blind and open-label period. For log-normally distributed biomarkers, geometric means will be provided in addition. For NfL and CXCL13, descriptive statistics of the percent change from baseline will be provided as well at each available timepoint by intervention group in the Safety population. Missing values will not be imputed.

Boxplots of raw data and absolute change from baseline will also be provided by intervention group over time, over the double-blind period and then over both the double-blind and open-label period.

Pharmacodynamic model

Restricted maximum likelihood estimation based on ANCOVA model will be used to obtain NfL concentration estimates (geometric mean concentration) at Week 12 in the Safety population. The response variable will be the log-transformed values of the NfL level as the NfL level is expected to follow log-normal distribution. The model will include treatment as categorical effect and log-transformed NfL baseline concentration as continuous fixed effect. Baseline adjusted exponentiated least-squares means estimates (adjusted geometric means of NfL value) at Week 12 will be provided for both SAR441344 groups and placebo group with their corresponding 95% CIs. The exponentiated SAR441344 intervention group differences with placebo are the geometric mean ratios (GMR) which will be reported at Week 12 with the corresponding 95% CIs, and p-values comparing the treatment groups (for descriptive purposes only).

In addition, a similar model will be done but using "log-transformed (NfL) - log-transformed (baseline NfL)" as response variable. Baseline adjusted exponentiated least-squares means estimates (adjusted geometric means of the reduction ratio) at Week 12 will be provided for both

SAR441344 groups and placebo and 95% CIs. Both models are equivalent in terms of treatment comparison.

If needed, outlying NfL values collected close to relapses may be removed from the model as a second step.

CXCL13 will be analyzed using an ANCOVA on the absolute change from baseline in CXCL13 at Week 12, including treatment as categorical effect and baseline as continuous covariate, in the Safety population. The least-squares means estimates of the changes from baseline at Week 12 in each treatment group and their corresponding two-sided 95% CIs will be reported, as well as the differences between SAR441344 groups and placebo in least square mean change from baseline (and 95% CIs) and p-values comparing the treatment groups (for descriptive purposes only). The analysis of the residuals will be primarily based on plots of studentized residuals. In case of major deviation to the statistical assumptions (normality of the distribution for instance), a log-transformation of the data will be tried and CXCL13 will be analyzed similarly to NfL concentration. If needed, in case statistical assumptions are still not met after log-transformation, a rank-based analysis of covariance (rank ANCOVA) model adjusted for baseline will be used instead to compare the CXCL13 between SAR441344 intervention groups and placebo.

3.7.1.4.4 Other biomarkers

Other biomarker analyses will be reported separately.

3.7.2 Subgroup analyses

A cut-off for the CD40/CD40L pathway activation score which may enable to identify participants who preferentially benefit from SAR441344 will be determined by the 60th percentile of the baseline activation score (40% of the participants will have a baseline activation score greater or equal to the cut-off). This pre-defined threshold will be used as a basis to perform subgroup analyses within this study.

Subgroup analyses based on the pre-defined threshold

Subgroup analyses of the primary efficacy endpoint will therefore be performed to assess the homogeneity of the treatment effect at Week 12 across the subgroups based on CD40/CD40L pathway activation scores:

- CD40/CD40L pathway activation score < pre-defined threshold (60th percentile)
- CD40/CD40L pathway activation score ≥ pre-defined threshold (60th percentile)

SAR441344 effect on the primary endpoint in each subgroup will be estimated in the Efficacy Population using a similar model as that utilized for the primary analysis (see Section 3.2.2) but including additionally the categorized CD40/CD40L pathway activation score (score above pre-defined threshold Y/N) and categorized CD40/CD40L pathway activation score by treatment interaction. SAR441344 effect will be estimated through the categorized CD40/CD40L pathway activation score by treatment interaction.

The mean number of new lesions per month at Week 12 in SAR441344 intervention groups in each subgroup, as well as the relative reduction compared to placebo group will be estimated (estimates, rate ratios and 95% CIs) using linear function of the least squares means obtained through the negative binomial regression model and exponential transformation. The LS-means coefficients will be proportional to those found in the analysis data set (OBSMARGINS SAS option). In case of model fitting issue, other options such as eg, an exact Poisson regression model will be used.

Subset analyses based on pre-defined threshold in SAR441344 participants and all placebo participants:

An exploratory analysis assessing the treatment effect at Week 12 in the subset of participants from the Efficacy Population including SAR441344 participants with CD40/CD40L pathway activation scores greater than the pre-defined threshold, and all placebo participants will be performed: SAR441344 effect on the primary endpoint in the above defined subset of participants will be estimated using a negative binomial regression including the baseline GdE T1 activity (presence/absence) and treatment as factors. In case of model fitting issue, other options such as eg, an exact Poisson regression model will be used.

The mean number of new GdE T1 lesions per month at Week 12 in SAR441344 intervention groups in the above defined subset of participants, as well as the relative reduction compared to placebo group will be estimated (estimates, rate ratios and 95% CIs) using linear function of the least squares means obtained through the negative binomial regression model and exponential transformation. The LS-means coefficients will be proportional to those found in the analysis data set (OBSMARGINS SAS option). The same asymmetric confidence intervals as described in Section 3.2.2 will be also provided for Quantitative Decision-Making methodology.

In addition, the SAR441344 effect in the above defined subset of participants will be also estimated in the ITT Population using the same logistic regression model described in Section 3.2.4 but only including the baseline GdE T1 activity (presence/absence) as a covariate.

Analyses based on the threshold identified as per Section 3.7.1.4.2 in SAR441344 participants and all placebo participants:

If a different responder population is identified as per Section 3.7.1.4.2, similar analyses as described above will be performed, and in the scope of the subset analysis, asymmetric confidence intervals will be also provided for Quantitative Decision-Making methodology (as described in Section 3.2.2). The outcome of this Quantitative Decision-Making methodology will be provided for information purposes only.

3.8 INTERIM ANALYSES

No interim analysis is planned since the analysis of the double-blind study period is considered as the primary analysis. No multiplicity adjustment for multiple analyses is needed.

Two analyses will be conducted:

1. Analysis of the double-blind treatment period, after completion of the 12-week double-blind period for all participants, and analysis of the main efficacy and safety data available in the open-label treatment period (Part A database lock). The efficacy analysis in the open-label treatment period will include summary descriptive statistics of the number of new GdE T1 lesions, the number of new or enlarging T2 lesions and the total number of GdE T1 lesions over time. The safety analysis in the open-label treatment period will consist of the analysis of adverse events (AE overview, TEAE and SAE analysis). No individual data listing will be provided at this stage in order to maintain the blind linked to initial randomization to the study team (outside of Biostatistics & Programming).
2. Analysis of the open-label treatment period based on complete and cleaned data (final database lock). Safety table and listings on the double-blind period may be updated in order to include possible newly reported double-blind period data (which were collected after Part A database lock).

3.9 CHANGES TO PROTOCOL-PLANNED ANALYSES

No major statistical changes were included in the protocol amendment(s).

4 SAMPLE SIZE DETERMINATION

The sample size was derived by applying the frequentist approach of the Quantitative Decision Making method described by Quan et al (4). This approach is based on the comparison of the lower (LL) and upper limits (UL) of an asymmetric CI of the treatment difference, with target (TV) and lower reference (LRV) values.

Using the following features:

- TV: 90% reduction
- LRV: 70% reduction
- Significance criterion: 86% (P(Non-negative conclusion| LRV)≈14%)
- Relevance criterion: 15% (P(Negative conclusion| TV)≈15%)

And, assuming the following:

- The number of new GdE T1 lesions by month follows a negative binomial distribution with a mean of 0.1 in the SAR441344 group and a mean of 1 in the placebo group (that is, 90% reduction of mean lesions count in the SAR441344 group relative to that of the placebo group).
- The dispersion parameter of the negative binomial distributions equals 2 (estimated from Week 12 placebo data from the vatelizumab [SAR339658] DRI13839 study).

Then, with 40 evaluable participants in the SAR441344 group, and 20 evaluable participants in the placebo group, the probability for the asymmetric CI of the relative reduction in mean new GdE T1 lesions count at Week 12 in participants receiving SAR441344 compared to that of those receiving placebo to be located in the “Positive” conclusion area is approximately 71%, and the probability to be located in the “Negative” conclusion area is approximately 15%. Conclusion rules are detailed in Section 3.2.2 (Table 4).

Assuming that IV and SC placebo arms can be pooled for comparison with either SAR441344 1200 mg IV q4w arm or SAR441344 300 mg SC every 2 weeks arm, and as no adjustment for multiplicity of doses is planned, approximately 160 people will be screened to achieve 120 participants randomly assigned to study intervention and 100 evaluable participants for an estimated total of 40 and 10 evaluable participants per SAR441344 and placebo intervention in each of the dose groups respectively.

It is planned to perform subgroup analysis according to CD40/CD40L pathway activation as measured by mRNA signature in peripheral blood cells. With this sample size and assuming 40% of participants classified as participants with the activated pathway at baseline (CD40/CD40L+) and assuming the absence of effect on primary endpoint in the placebo group, the comparison of 16 participants CD40/CD40+ receiving SAR441344 versus 20 participants receiving placebo will result in an overall probability for a positive conclusion of approximately 53% and a negative conclusion of less than 15%. Assuming a 95% reduction of mean lesion counts versus placebo in this sub-population, the probability of positive conclusion increases to 64% and negative conclusion decreases to 5% with 31% probability of intermediate outcome.

5 SUPPORTING DOCUMENTATION

5.1 APPENDIX 1 LIST OF ABBREVIATIONS

AE:	adverse event
AESIs:	adverse events of special interest
ALT:	alanine aminotransferase
BUN:	blood urea nitrogen
CI:	confidence interval
ECG:	electrocardiogram
HLT:	high level term
LLT:	lower-level term
MDRD:	modification of the diet in renal disease
MedDRA:	medical dictionary for regulatory activities
PCSA:	potentially clinically significant abnormality
PT:	preferred term
SOC:	system organ class
TEAE:	treatment-emergent adverse event
ULOQ:	Upper Limit of Quantification

5.2 APPENDIX 2 PARTICIPANT DISPOSITIONS

The number (%) of participants included in each of the analysis populations listed in Table 2 will be summarized.

Screen failures are defined as participants who consent to participate in the study but are not subsequently randomized. The number (%) of screen failures and reasons for screen failures will be provided in the screened population.

The number (%) of participants in the following categories will be provided:

- Randomized participants
- Randomized but not exposed participants
- Randomized and exposed participants
- Participants who completed the double-blind treatment period as per protocol
- Participants who did not complete the double-blind treatment period as per protocol and main reason for permanent intervention discontinuation.
- Participants who completed the double-blind study period as per protocol.
- Participants who did not complete the double-blind study period as per protocol and main reason for study discontinuation.

- Participants who completed the double-blind study period as per protocol but did not enter in the open-label treatment period
- Participants who completed the open-label treatment period as per protocol
- Participants who did not complete the open-label treatment period as per protocol and main reason for permanent intervention discontinuation.
- Participants who completed the open-label study period as per protocol.
- Participants who did not complete the open-label study period as per protocol and main reason for study discontinuation.

Reasons for permanent study intervention and study discontinuation "adverse event" and "other reasons" will be split as related versus not related to COVID-19. A specific reason within "other reasons" related to emergency situation (eg, war in Ukraine) will be included as well.

The number (%) of exposed and not randomized participants will be listed.

In addition, the number (%) of participants screened, screened-failed, randomized, with permanent treatment discontinuation (by treatment period) and with early study discontinuation will be provided by country and by country/site.

A summary of visits impacted by COVID-19 pandemic will be provided along with the description of the impact (visit not done, visit partially done on site/by phone, visit done but delayed).

Protocol deviations

Critical and major protocol deviations (automatic or manual) will be summarized in the randomized population, by treatment period (double-blind and open label).

In addition, they will be displayed separately as related versus not related to COVID-19, by treatment period, in the randomized population. The list of predefined protocol deviations can be found in the eTMF (Trial management Section).

5.3 APPENDIX 3 DEMOGRAPHICS AND BASELINE CHARACTERISTICS, PRIOR OR CONCOMITANT MEDICATIONS

Demographics, baseline characteristics, medical surgical history

The following demographics and baseline characteristics, medical and surgical history and multiple sclerosis characteristics at baseline will be summarized using descriptive statistics in the randomized population.

Demographic and baseline characteristics

- age in years as quantitative variable and in categories (<18, ≥18 to <25, ≥25 to <35, ≥35 to <45, ≥45 to ≤55, >55)
- gender (Male, Female)
- race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Island, Not reported, Unknown). If several races are collected for a participant, a category of "Multiple" will be displayed.
- ethnicity (Hispanic or Latino, not Hispanic or Latino, Not reported, Unknown)
- body mass index (BMI) in kg/m² (quantitative and qualitative variable: <18; ≥18-<25; ≥25-<30; ≥25-<30; ≥30-<35; ≥35)

Medical and surgical history will be coded to a LLT, PT, HLT, HLGT, and associated primary SOC using the MedDRA version currently in effect at Sanofi at the time of Part A database lock.

Baseline characteristics related to multiple sclerosis history include:

- Multiple Sclerosis Type (Relapsing Remitting Multiple Sclerosis, Secondary Progressive Multiple Sclerosis)
- Number of relapse(s) within past year, as quantitative variable and in categories (0,1,2, ≥3)
- Number of relapse(s) within past 2 years, as quantitative variable and in categories (0,1,2, ≥3)
- Number of Gd-enhancing brain lesion(s) on the last MRI in the past 6 months prior to screening
- Time since first symptoms of RMS [years], calculated from the date of informed consent
- Time from diagnosis of RMS [years], calculated from the date of informed consent
- Time since the most recent relapse [months], calculated from the date of informed consent

Other baseline safety and efficacy parameters (not listed above) will be presented along with the safety and efficacy summaries.

Prior or concomitant medications

All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD) using the version currently in effect at Sanofi at the time of Part A database lock. Medications to be reported in the e-CRF Prior/Concomitant Medications form are:

- Medications received during the 4 weeks before randomization
- All prior MS treatments
- All prior treatments considered clinically important
- Medications taken during the study including MS related treatment

Medications will be reported using the following definitions:

- Prior medications are those the participant received within 4 weeks prior to first dose of double-blind IMP. Prior medications can be discontinued before first administration or can be ongoing during treatment period.
- Double-blind concomitant medications are any medications received by the participant concomitantly to the IMP, from first dose of double-blind IMP to the last dose of double-blind IMP + 180 days. For participants entering in the open-label period, concomitant medications will be truncated at the day before first open-label IMP administration in the open-label period.
- Post-treatment double-blind medications are those the participant received in the period running from the end of the double-blind concomitant medications period up to the end of the study.
- Open-label concomitant medications are defined as any medication received by the participant concomitantly with the open-label IMP, from the first open-label IMP administration to the last open-label IMP administration + 180 days.
- Post-treatment open-label medications are those the participant received in the period starting from 181 days after the last open-label IMP administration.
- A given medication can be classified as a prior medication and/or as a concomitant medication and/or as post-treatment medication. If it cannot be determined whether a given medication was taken prior or concomitantly or post, it will be considered as prior, concomitant, and post-treatment medication.

The prior and concomitant and post-treatment medications will be summarized for the randomized population, by anatomic and therapeutic level. The summaries will be sorted by decreasing frequency of anatomic category (ATC) and therapeutic class based on incidence in SAR441344 IV intervention groups. In case of equal frequency, alphabetical order will be used. Participants will be counted once in each ATC category (anatomic or therapeutic) linked to the medication.

All concomitant and post-treatment medications recorded during the open-label period will be summarized by overall participants in the open-label extension population.

Prohibited medications include other MS disease-modifying treatments and medications listed as part of exclusion criteria E06 (refer to protocol Section 5.2). They will be listed as part of protocol deviations.

5.4 APPENDIX 4 DATA HANDLING CONVENTIONS

Analysis windows for time points

The following analysis windows will decide how the scheduled and/or unscheduled visits will be used in the by-visit analyses of efficacy, safety, PK and ADA and biomarkers variables. For participants with premature study discontinuation, follow-up measurements will be reallocated to the corresponding analysis window and reported together with other measurements.

A measurement (scheduled or unscheduled) will be used if it is available and measurement date is within the analysis window.

After applying these time windows, if multiple assessments are associated to the same time point, the closest from the targeted study day will be used. If the difference is a tie, the value after the targeted study day will be used. If multiple valid values exist within a same day, then the first value of the day will be selected when time is available, otherwise the scheduled exam will be used.

If there is no measurement for a given parameter in an analysis window, data will be considered missing for the corresponding visit.

Table 8 - Analyses window definition

Study period	Scheduled visit post baseline	Targeted study day	Analysis window in study days
Double-blind period	Week 2 (Visit 3)	15	Not defined[a]
	Week 4 (Visit 4)	29	15 to 42
	Week 6 (Visit 5)	43	Not defined[a]
	Week 8 (Visit 6)	57	43 to 70
	Week 10 (Visit 7)	71	Not defined[a]
	Week 12 (Visit 8)	85	71 to 98 for participants not entered in the open-label period For MRIs: 71 to 3 days after the study day corresponding to the first open-label administration for participants who entered in the open-label period For other endpoints: 71 to the study day corresponding to the first open-label administration for participants who entered in the open-label period
Open-label period/safety data and PK	Week 14 (Visit 9)	99	Not defined[a]
	Week 16 (Visit 10)	113	Study day corresponding to the first open-label administration +1 to 126
	Week 18 (Visit 11)	127	Not defined[a]
	Week 20 (Visit 12)	141	127 to 154
	Week 22 (Visit 13)	155	Not defined[a]
	Week 24 (Visit 14)	169	155 to 182
	Week w (additional SC injection q4w visits after W24 to EOS)	w*7 +1	Not defined[a]
	Week w (q4w visits after W24 to EOS)	w*7 +1	w*7 +1-14 to w*7 +14

Study period	Scheduled visit post baseline	Targeted study day	Analysis window in study days
Open-label period/Efficacy data, PROs data and Ig levels and Plasma sample for NfL, CHI3L1, and other biomarkers	Week 20 (Visit 12)[b]	141	For MRIs: 4 days after the study day corresponding to the first open-label administration to 154 For other endpoints: Study day corresponding to the first open-label administration + 1 to 154
	Week 24 (Visit 14)	169	155 to 252 for MRI parameters 127 to 252 for other parameters
	Week 48 (Visit 20)	337	253 to 420
	Week 72 (Visit 26)	505	421 to 588
Open-label period/ PBMC substudy and Whole blood biomarkers	Week 24 (Visit 14)	169	Any result from the open-label period will be allocated to W24

Study days are calculated considering Day 1 as the day of first administration of intervention in double-blind period (or the day of randomization for participant not exposed).

a These visits were planned for SC arms only and no exams were performed during these visits

b For MRI parameters only

Unscheduled and EOT/EOS visits

Unscheduled visit measurements of efficacy, laboratory data, vital signs, biomarkers and ADA will be used, in particular for computation of baseline, the last on-treatment value, analysis according to PCSAs, and the shift summaries for safety. They will also be included in the by-visit summaries if they are re-allocated to scheduled visits.

Biomarkers

- The baseline biomarker parameter is defined as the last non-missing value prior to the first dose of IMP unless otherwise specified.
- Only on-treatment results will be used (as defined in Section 3.1).
- EOS visits will only be included (if available) in descriptive statistics and not in pharmacodynamic models.
- As for laboratory data, biomarkers data below the lower limit of quantitation/detection limit (LLOQ) will be replaced by half of the LLOQ, data above the upper limit of quantification will be replaced by ULOQ value. Data above the upper limit of quantification (ULOQ) will be replaced by ULOQ value.
- For all biomarkers (including Nanostring) patients experiencing an early EOT due to inflammatory AE (ie, SOC 'Infections and infestations' and leading to permanent IMP discontinuation): all values post-inflammatory AE will be considered as missing in order to remove potential misleading effect.

5.5 APPENDIX 5 PROMIS CONVERSION TABLE FROM RAW SCORE TO T-SCORE

Based on (2), the following conversion table will be used:

PROMIS-Fatigue$_{MS}$ raw score	Equivalent PROMIS T-score
8	34.1
9	39.3
10	41.8
11	43.6
12	45.1
13	46.4
14	47.6
15	48.7
16	49.8
17	50.9
18	52.0
19	53.0
20	54.1
21	55.1
22	56.2
23	57.3
24	58.3
25	59.3
26	60.4
27	61.5
28	62.5
29	63.6
30	64.7
31	65.8
32	66.9
33	68.1
34	69.3
35	70.5
36	71.9
37	73.4
38	75.1
39	77.5
40	80.9

6 REFERENCES

1. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The multiple Scleroris Impact Scale (MSIS-29): A new patient-based outcome measure. Brain. 2001;124(Pt 5):962-73.

2. Cook KF, Bamer AM, Roddey TS, Kraft GH, Kim J, Amtmann D. A PROMIS fatigue short form for use by individuals who have multiple sclerosis. Qual Life Res. 2012;21(6):1021-30.

3. HealthMeasures. PROMIS Fatigue Scoring Manual [Online]. 2022 Mar 06 [cited 2022 Jun 27]. Available from: URL: https://staging.healthmeasures.net/images/PROMIS/manuals/PROMIS_Fatigue_Scoring_Manual.pdf

4. Quan H, Chen X, Lan Y, Luo X, Kubiak R, Bonnet N, et al. Applications of Bayesian analysis to proof-of-concept trial planning and decision making. Pharm Stat. 2020;19(4):468-81.

Signature Page for VV-CLIN-0607107 v3.0
act16877-16-1-9-sap

Approve & eSign	[redacted]
Approve & eSign	[redacted]