STATISTICAL ANALYSIS PLAN

Protocol title:	**A Phase 2 double blind, randomized, placebo controlled study evaluating the effect of SAR443820 on serum neurofilament levels in participants with multiple sclerosis, followed by an open label long-term extension period**
Protocol number:	**ACT16753**
Compound number (INN/Trademark):	**SAR443820** **Not Applicable**
Study phase:	**Phase 2**
Short Title:	**Phase 2 study of SAR443820 in participants with multiple sclerosis (MS)**
Statistician:	[redacted]
Statistical project leader:	[redacted]
Date of issue:	**12-Jul-2024**

Regulatory agency identifier number(s):

IND:	
EudraCT/EU-trial number:	2022-000049-34
NCT:	NCT05630547
WHO:	U1111-1271-1257
EUDAMED:	N/A
Other:	N/A

Date: 12-Jul-2024 Total number of pages: 38

TABLE OF CONTENTS

LIST OF TABLES

VERSION HISTORY

This statistical analysis plan (SAP) for study ACT16753 is based on the Amended Clinical Trial Protocol 04 dated 13-December-2023. The first participant was randomized on 16-Jan-2023. This SAP is approved before the first database lock.

Major changes in statistical analysis plan

SAP Version	Approval Date	Description of statistical changes	Rationale
1	12-Jul-2024	This version of SAP includes the following changes compared to the amended protocol 04: • For the analysis of sNfL, replace baseline MS type (RMS versus PMS) with the randomization stratification factor MS clinical subtype (RRMS, SPMS, or PPMS) in the MMRM model and exclude MS type by visit interaction from the covariates. For the analyses of new Gd-enhancing T1 lesions and new/enlarging T2 lesions, replace baseline MS type (RMS versus PMS) with the randomization stratification factor MS clinical subtype (RRMS, SPMS, or PPMS) in the negative binomial model. • Remove the endpoints 'change from baseline in the total number and volume of non-enhancing lesions at Weeks 12, 24, 36, and 48' for Part A and 'change from baseline in the total number and volume of non-enhancing lesions at Week 96' for Part B as data for these endpoints were not collected • Remove the endpoints 'change from baseline in the intensity (T1) of slowly expanding lesions at Weeks 12, 24, 36 and 48' for Part A and 'change from baseline in the intensity (T1) of slowly expanding lesions to Week 96' for Part B as data for the corresponding endpoints were not collected. • Change the confirmation period of time-to-event endpoints in Part B (ie, time to CDP, time to 20% increase in 9-HPT, time to 20% increase T25-FW test, and time to CCDP) from 12 weeks to 24 weeks.	Clarification

1 INTRODUCTION

1.1 STUDY DESIGN

Study ACT16753 is a Phase 2, randomized, double-blind, placebo-controlled, 2 parallel-arm study evaluating the effect of SAR443820 on serum neurofilament light chain (sNfL) levels in participants with multiple sclerosis, followed by an open-label, long-term extension period.

The study consists of 2 parts (A and B) as follows:

Part A is a 48-week, double-blind, placebo-controlled part, preceded by a screening period of up to 4 weeks before Day 1.

- The screening period will be up to 4 weeks and is designed to evaluate suitability for participation in the study in terms of MS diagnosis, clinical status based on EDSS, and safety screening evaluations. No IMP will be administered in this period.
- The double-blind treatment period will have a total duration of 48 weeks for each participant and will include the following:
  - Participants will be randomized at 1:1 ratio to receive SAR443820 20 mg QD (n = 84) or matching placebo QD (n = 84).
  - Randomization will be stratified by MS clinical subtypes (ie, RRMS, SPMS, or PPMS) in order to ensure a balance between the treatment and placebo arms within each MS clinical subtype. Participants will attend in-clinic study assessments at baseline (Day 1) and Weeks 2, 6, 12, 24, 36, and 48.

All participants who successfully completed Part A will be offered to rollover to Part B. Participants who discontinue IMP during Part A, or who choose not to enter Part B will have their early discontinuation or safety follow-up visit, respectively, up to 2 weeks after the last dose of IMP.

Part B is an open-label, long-term extension of Part A. Part B starts from the end of Part A (Week 48) and continues up to Week 96.

- The assignment of participants to study intervention in Part A will remain double-blinded during Part B unless there is medical need to unblind the study intervention assignment.
- All participants will receive 20 mg QD SAR443820 in Part B.
- Participants will attend in-clinic study assessments at Weeks 52, 72, and 96 (and Week 48 which is the last visit of Part A and first visit of Part B); Week 50, 54, 56, 58, 60, 62, 64, 66, 68, 70, 76, 80, 84, 88, 92 visits for liver chemistry only may be conducted at local labs where available.

1.2 OBJECTIVES AND ENDPOINTS

Table 1 - Objectives and endpoints

Part A

Objectives	Endpoints
Primary	
• To assess the effect of SAR443820 compared to placebo on serum Neurofilament light chain (sNfL)	• Week 48 sNfL levels relative to baseline
Secondary	
• To evaluate efficacy of SAR443820 compared to placebo on imaging and clinical endpoints	Cumulative number of new gadolinium (Gd)-enhancing T1 hyperintense lesions as detected by MRI at Week 48, defined as the sum of the individual number of new Gd-enhancing T1 hyperintense lesions at all scheduled visits starting after baseline up to and including the Week 48 visit Cumulative number of new and/or enlarging T2 hyperintense lesions as detected by MRI, at Week 48, defined as the sum of the individual number of new and/or enlarging T2 lesions at all scheduled visits starting after baseline up to and including the Week 48 visit Time to onset of 12 weeks confirmed disability progression (CDP) from baseline as assessed by the Expanded Disability Status Scale (EDSS) score Time to onset of sustained 20% increase in 9-Hole Peg Test (9-HPT) confirmed over at least 12 weeks Time to onset of sustained 20% increase in timed 25-foot walk test (T25-FW) confirmed over at least 12 weeks Change from baseline in EDSS-Plus at Week 48 Annualized relapse rate (ARR) of RMS population (relapsing SPMS and RRMS) up to Week 48
• To explore effect of SAR443820 compared to placebo on brain volume and chronic lesions	Percent change from baseline in brain volume loss (BVL) as detected by brain MRI at 48 weeks Change from baseline in the volume, number, and normalized intensity (T1) of slowly expanding lesions at Weeks 12, 24, 36, and 48 Change from baseline in the number of phase rim lesions (PRL) at Weeks 12, 24, 36, and 48; will be conducted at 3T capable sites
• To assess the safety and tolerability of SAR443820	Incidence of AE, SAE, TEAE, PCSA in laboratory tests
• To assess pharmacokinetic (PK) of SAR443820	Plasma concentration of SAR443820
Tertiary	
• To assess the effect of SAR443820 compared to placebo on sNfL in	• Change from baseline in sNfL level in MS subpopulations (RRMS, SPMS, PPMS) over 48 weeks

Objectives	Endpoints
subpopulations of multiple sclerosis (MS)	
• To assess the effect of SAR443820 compared to placebo on biomarkers of neurodegeneration, inflammation, and disease progression	• Change from baseline compared to Week 48 in: plasma chitinase-3-like protein 1 (CHI3L1), serum glial fibrillary acidic protein (sGFAP), interleukin-1B (IL1B), IL6, IL8, TNFα, Chemokine (C-C motif) ligands 3 (CCL3), and Chemokine (C-C motif) ligands 4 (CCL4)
• To assess efficacy of SAR443820 compared to placebo on patient-reported outcomes (PROs) assessing the physical and psychological impact of MS from the patient's perspective (MSIS-29v2) and the impact of MS on the individual's walking ability (MSWS-12)	• Change from baseline in MSIS-29v2 physical and psychological domains scoring at Week 12, 24, 36, and 48 • Change from baseline in MSWS-12 at Week 12, 24, 36, and 48

Part B

Objectives	Endpoints
Primary	
• To assess long-term trends in durability of sNfL	• Week 96 sNfL levels relative to baseline
Secondary	
• To explore the effect of SAR443820 on brain volume and chronic lesions	• Percent change from baseline in BVL as detected by brain MRI at Week 96 • Change from baseline in volume, number and normalized intensity (T1) of slowly expanding lesions to Week 96 • Change from baseline in the number of PRLs (same participants/centers from Part A with 3T capability) at Week 96
• To assess the long-term safety and tolerability of SAR443820	• Incidence of AE, SAE, TEAE, PCSA in laboratory tests, ECG, and vital signs during through Week 96
• To evaluate long-term effect of SAR443820 on disease progression and activity assessed by other clinical and imaging measures on physical function and patient-reported outcomes (PROs)	• Cumulative number of new Gd-enhancing lesions as detected by T1-weighted MRI between 48 and 96 weeks; number of new or enlarging T2-hyperintense lesions on MRI at Week 96 relative to Week 48 • Annualized relapse rate (ARR) of RMS population (relapsing SPMS and RRMS) up to Week 96 • Time to onset of composite CDP (CCDP), confirmed over at least 24 weeks (6-month CCDP), by the EDSS Plus composite (EDSS score increase, OR 20% increase in the T25-FW test, OR 20% increase in the 9-HPT) • Time to onset of 24-week CDP as assessed by the EDSS score • Time to onset of sustained 20% increase in 9-HPT confirmed over at least 24 weeks • Time to onset of sustained 20% increase T25-FW test confirmed over at least 24 weeks

Objectives	Endpoints
	• Change from baseline in EDSS-Plus at Week 96 • Change in MSIS-29v2 physical and psychological domains scoring from baseline through Week 96 • Change in MSWS-12 from baseline through Week 96

1.2.1 Estimands

Primary estimands defined for main endpoints in Part A are summarized in below Table 2. More details are provided in Section 3.

For all these estimands, the comparison of interest will be the comparison of SAR443820 vs. placebo.

Table 2 - Summary of primary estimand for main endpoints

Endpoint Category	Estimands			
	Endpoint	Population	Intercurrent event(s) handling strategy	Population-level summary (analysis and missing data handling)
Primary objective: To assess the effect of SAR443820 compared to placebo on sNfL				
Primary endpoint (hypothetical strategy)	sNfL levels at Week 48 relative to baseline	mITT	Had IMP not been discontinued (hypothetical strategy), had baseline DMT not been permanently modified (hypothetical strategy)	Ratio of the geometric mean ratios compared to baseline between intervention groups from the mixed effect model with repeated measures (MMRM) on log-transformed data. Analysis will exclude data collected after the intercurrent events. Missing data will be handled by MMRM under the missing at random (MAR) assumption

Abbreviations: DMT = disease modifying therapy; IMP = investigational medicinal product; mITT = modified intent-to-treat; sNfL = serum neurofilament light chain

2 ANALYSIS POPULATIONS

The following populations for analyses are defined.

Table 3 - 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.
Exposed	All screened participants who take at least 1 dose of study intervention.
Randomized and treated	All randomized participants who take at least 1 dose of study intervention. Participants will be analyzed according to the intervention allocated by randomization.
Intent-to-treat (ITT)	All randomized participants. Participants will be analyzed according to the intervention allocated by randomization.
Modified ITT (mITT)	All participants from the ITT population who take at least 1 dose of study intervention and with an evaluable primary endpoint. The primary endpoint is evaluable when the following conditions are met: • The participant has a baseline and at least 1 post-baseline sNfL assessment before IMP discontinuation and before permanent modification of the background DMT. • Participants will be analyzed according to the intervention allocated by randomization.
Safety	All randomized participants who take at least 1 dose 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 1 post-baseline PK result with adequate documentation of dosing and sampling dates and times. Participants will be analyzed according to the intervention they actually received.

Abbreviations: DMT = disease modifying therapy; ICF = informed consent form; IRT = interactive response technology; sNfL = serum neurofilament light chain

Participants exposed to study intervention before or without being randomized will not be considered as 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 whether 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 SAR443820 20 mg QD.

3 STATISTICAL ANALYSES

3.1 GENERAL CONSIDERATIONS

In general, continuous data will be summarized using the number of observations available, mean, standard deviation (SD), median, interquartile range (Q1, Q3), minimum, and maximum. Categorical and ordinal data will be summarized using the count and percentage of participants. Efficacy data will be analyzed in the randomized and treated population except the primary endpoint which will be analyzed in the mITT population, and safety data will be analyzed in the safety population, unless specified otherwise.

In Part A, baseline values are defined as follows. The baseline value of NfL is defined as the lower value of screening and Day 1 visits if the number of gd-enhancing T1 lesion at Day 1 is greater than zero, otherwise it will be the average values of the assessments on screening and Day 1 visits prior to the first dose of the study intervention. The baseline values of EDSS and PD biomarkers other than NfL are defined as the average of the screening and Day 1 visit values prior to the first dose of the study intervention. If one of the values for the above assessments at screening and Day 1 is missing, the non-missing value will be used as the baseline value. For the other efficacy and safety assessments, the baseline values are defined as the last available value before the first dose of the study intervention. For the C-SSRS, the baseline value will be the worst assessment between the past 6 months evaluation at screening and the since last visit evaluation on Day 1. For participants randomized but not treated, the baseline value is defined as the last available value before randomization. Unless otherwise indicated, 2-sided p-values and 95% confidence intervals [CI(s)] will be provided for efficacy assessment of treatment differences. Unless otherwise specified, analyses will be performed by initial intervention group, ie, placebo or SAR443820 20 mg QD (and overall for baseline and demographics characteristics).

In Part B, efficacy analyses will be conducted accumulatively according to the treatment group allocated by randomization in Part A (ie, placebo/SAR443820 20 mg QD versus SAR443820 20 mg QD/ SAR443820 20 mg QD). Baseline values defined above for Part A will be used as the baseline in the efficacy analyses for Part B, unless otherwise noted. Safety analyses in Part B are specified in Section 3.6.

Observation period

The observation period for safety will be divided into 4 segments:

- The **pre-treatment** period is defined as the period up to first IMP administration.
- The **on-treatment period** (ie, treatment-emergent (TE) period) **for the Part A** is defined as the period from the first IMP administration to the earliest of 1) last IMP administration + 14 days, 2) first IMP in Part B, 3) death date, or 4) last contact date.
- The **on-treatment period** (ie, treatment-emergent (TE) period) **for the Part B** is defined as the period from the first IMP administration in Part B to the earliest of 1) last IMP administration + 14 days, 2) death date, or 3) last contact date.
- The **post-treatment period**, if applicable, is defined as the period from the end of the on-treatment period to the last contact date.

The on-study period is defined as the time from randomization until the end of the study defined as the last scheduled visit for those who completed the study and the end-of-study date collected on electronic case report form (e-CRF) page "Completion of End of Study" for those who did not complete the study. If death is the end-of-study reason, date of death will be used.

3.2 PRIMARY ENDPOINT(S) ANALYSIS

The primary endpoint in Part A detailed in this section is sNfL levels at week 48 relative to baseline.

The primary endpoint in Part B detailed in this section is sNfL levels at week 96 relative to baseline.

3.2.1 Definition of endpoint(s)

The primary endpoint in Part A and Part B is based on NfL levels measured in serum (sNfL). For Part A, the principal time point is Week 48. For Part B, the principal time point is Week 96.

3.2.2 Main analytical approach

3.2.2.1 Part A

The primary endpoint in Part A will be analyzed with the primary estimand defined according to the following attributes:

- Endpoint: sNfL levels at Week 48 relative to baseline
- Treatment condition: SAR443820 will be compared to placebo, on top of background therapy where applicable
- Analysis population: mITT population
- Intercurrent events (IE):
  - The IMP discontinuation IE will be handled with the hypothetical strategy. The primary endpoint will be assessed based on assessments performed before IMP discontinuation.
  - The modification in the background DMT (either dosage or medication) IE will be handled with the hypothetical strategy. The primary endpoint will be assessed based on assessments performed before the permanent modification of the background DMT (as distinguished from rescue therapy, which is intermittently limited to 2 to 5 days, after which the original background DMT will be restored at its pre-rescue dose).
- Population-level summary: ratio of the geometric mean ratios compared to baseline between intervention groups.

The primary analytical approach will use a mixed effect model with repeated measures (MMRM) on natural (base 'e') log-transformed sNfL levels. The model will include the fixed categorical effects of the treatment group (SAR443820 or placebo), MS clinical subtype (RRMS, SPMS, or

PPMS), visit (Weeks 12, 24, 36, 48), treatment-by-visit interaction, log-transformed baseline sNfL levels and log-transformed baseline sNfL levels-by-visit interaction. An unstructured covariance structure will be used to model the within-subject errors. The Kenward-Roger (1) approximation will be used to estimate denominator degrees of freedom and adjust standard errors. Missing data will be handled by MMRM under the missing at random (MAR) assumption. The geometric mean ratios of sNfL levels at Week 48 relative to baseline in each treatment group, as well as the ratio of the geometric mean ratios under treatment with SAR443820 compared to placebo will be estimated (point estimates and 2-sided 90% confidence intervals [CI], respectively) from the MMRM model using weights for each stratum equal to the overall proportion of participants in each stratum (ie, "population weight") through exponentiated least squared means and exponentiated least squared means differences. Plots of geometric mean ratios of sNFL levels relative to baseline in each treatment group with 90% CIs over time will be provided.

If this MMRM model fails to converge, the following variance-covariance structures will be tested in this order:

- Heterogeneous Toeplitz
- Heterogeneous AR(1)
- Heterogeneous CS
- Toeplitz
- AR(1)
- CS.

The first (co)variance structure yielding convergence will be used as the primary analysis. When a variance-covariance structure other than unstructured is used, the denominator degree of freedom will be estimated using the between within method (DDFM=BW in SAS PROC MIXED) (2)

3.2.2.2 Part B

The primary endpoint of Part B, ie, sNfL levels at Week 96 relative to baseline, will be analyzed using the same statistical methods described for the primary endpoint of Part A (see Section 3.2.2.1) for descriptive purpose only. Only data collected on or before IMP discontinuation will be included in the MMRM model. Plots of geometric mean ratios of sNfL levels relative to baseline in each treatment group with 90% CIs over time will be provided for better visualization.

3.2.3 Sensitivity analysis for primary endpoint in Part A

[redacted]

- [redacted]

3.2.4 Supplementary analyses

Participants who permanently discontinue study intervention, including those whose background DMT is permanently modified, will be encouraged to remain in the study or to attend a safety follow-up visit within 2 weeks following discontinuation of study intervention. The sNfL levels collected after the discontinuation of study intervention or permanently modified the background DMT will be included in a supplementary analysis for the primary endpoint in Part A to assess the efficacy of SAR443820 compared to placebo in an intention-to-treat (ITT) setting, ie, the assessment of the treatment policy strategy, using mITT population. The same MMRM model on log-transformed sNfL levels as specified in Section 3.2.2.1 for the main analysis of the primary endpoint in Part A will be applied. For participants who withdraw from the study before Week 48, sNfL levels will be missing after the study discontinuation. Missing data will be handled by MMRM under the missing at random (MAR) assumption.

3.2.5 Subgroup analyses for primary endpoint in Part A

To assess the homogeneity of the treatment effect across various subgroups, analyses will be performed on the primary endpoint in Part A across the following subgroups (categories with fewer than 10% participants will be combined with other categories; for factors with only two categories, no subgroup analysis will be performed if fewer than 10% participants in one category):

- MS clinical subtypes
  - RRMS, SPMS, PPMS
  - Pooled RMS (relapsing SPMS + RRMS), non-relapsing SPMS, PPMS
  - Pooled RMS (relapsing SPMS + RRMS), pooled PMS (non-relapsing SPMS + PPMS)
- Background DMT (with concomitant DMT, without concomitant DMT)
- Age at screening (<50, ≥50 years)
- Sex (Male or Female)
- Baseline BMI (<median, ≥median), median value refers to the median baseline BMI among all randomized participants.
- Baseline EDSS score (<4, ≥4)

The estimated geometric mean ratio in each treatment group (ie, SAR443820 versus placebo) and the relative reduction under treatment with SAR443820 compared to placebo will be provided (point estimates and 2-sided 90% confidence intervals [CI], respectively) for each subgroup separately, using the same method as specified in Section 3.2.2.1 for the main analysis of the

primary endpoint in Part A. Forest plots of geometric mean ratio and corresponding 90% CIs comparing SAR443820 to placebo within each subgroup will be provided.

Treatment by subgroup interaction and its p-value will be derived from a MMRM with terms for treatment, MS clinical subtypes (RRMS, SPMS, or PPMS), visit (Weeks 12, 24, 36, 48), treatment-by-visit interaction, log-transformed baseline sNfL levels, log-transformed baseline sNfL levels-by-visit interaction, subgroup (if different than the aforementioned covariates), and subgroup-by-treatment interaction as covariates. If a quantitative subgroup by treatment interaction is detected with nominal p-value < 0.1 for any subgroup factor, a further investigation will be performed to evaluate possible qualitative interaction.

3.3 SECONDARY ENDPOINT(S) ANALYSIS

The secondary endpoints detailed in this section covers the secondary efficacy endpoints mentioned in Table 1 under Section 1.2 for Part A and Part B.

Other secondary endpoints analyses are defined in Section 3.6.2 (AE, SAE), Section 3.6.1 (laboratory abnormalities) and Section 3.7.1 (PK).

3.3.1 Definition of endpoint(s)

Time to onset of 12-week CDP is defined as the time from randomization to the onset of an increase in EDSS score (defined as an increase of ≥1.0 point from the baseline EDSS score when the baseline score is ≤5.5 or an increase of ≥0.5 points from the baseline EDSS score when the baseline score is >5.5) confirmed after a 12-week interval.

- The initial onset increase from baseline EDSS score must be observed at a routine quarterly visit.
- All intermediate EDSS scores (EDSS scores obtained after onset of progression and before the confirmatory assessment), if any, must maintain at least the minimum increase.
- Confirmatory EDSS assessment must be obtained at least 12 weeks after onset and at a routine quarterly visit.

Note: 12-week confirmation must be ≥78 days after onset. Given that visits are scheduled every 12 weeks with a time window of ± 3 days for conducting EDSS assessment in Part A, the minimum per protocol allowed time between 2 EDSS assessments scheduled 12 weeks apart is 78 days (12 weeks=84 days minus a 3-day window for each of the 2 visits, ie, 84-6 = 78).

Time to onset of 24-week CDP is defined similarly with time to onset of 12-week CDP. The only difference is that for time to onset of 24-week CDP, the increase in EDSS score has to be confirmed over 24 weeks. Accordingly, the minimum per protocol allowed time between confirmation and onset is 154 days (6 months=24 weeks=168 days minus a 7-day for each of the 2 visits, ie, 168-14 = 154). Both the initial onset and the confirmation must be observed at scheduled visits, and all intermediate EDSS scores if any must maintain at least the minimum increase.

If a participant dies due to MS, it will be considered a confirmed disability progression regardless of the baseline EDSS or the change in EDSS. The time to onset will be calculated as (date of EDSS assessment at a tentative onset of the event – date of randomization + 1) or (date of death – date of randomization + 1) if a tentative onset date does not exist. Death for other reasons than MS will not be considered a disability progression and the participant's event time will be censored at the date of last EDSS assessment.

The 9-HPT is a brief, standardized, quantitative test of upper extremity function (lower bound: 10 sec; upper bound: 300 sec). Two consecutive trials with the dominant hand are immediately followed by two consecutive trials with the non-dominant hand. Let D_1 and D_2 be the time to complete the 9-HPT for the dominant hand for the two trials and let ND_1 and ND_2 be the time to complete the 9-HPT for the non-dominant hand for the two trials. The overall 9-HPT is calculated as:

$$9HPT = \frac{2}{\left(\frac{2}{D_1 + D_2}\right) + \left(\frac{2}{ND_1 + ND_2}\right)}$$

For a participant who could not complete the 9-HPT due to a "physical limitation", the maximum time (300 sec) will be used. In case of missing data for other reason than inability to complete the test, missing data will be treated as follows:

- In case only 1 trial was done for a hand, the average of the two trials will be replaced by the time for the available trial.
- In case 9-HPT was assessed for only one hand, the overall 9-HPT will not be calculated and considered missing.

A higher value of overall 9-HPT is indicative of higher disability. In case of values below (above) the lower (upper) bound, the value will be set to the respective bound.

The T25-FW is a quantitative mobility and leg function performance test based on a timed 25-foot walk; time in seconds to walk 25 feet (lower bound: 2.2 sec; upper bound: 180 sec). For each assessment, 2 trials are performed. The score is the average of the times from the 2 trials (or single score if only 1 completed). For a participant who could not complete the 25-FW due to a "physical limitation", the maximum time (180 sec) will be used. Missing values for other reasons will not be imputed. In case of values below (above) the lower (upper) bound, the value will be set to the respective bound. A higher value of T25-FW is indicative of higher disability.

Time to onset of 24-week CCDP, as assessed by composite endpoint EDSS-Plus (EDSS score, or T25-FW test, or 9-HPT), is to be confirmed over at least 24 weeks (3). Disability criteria for the individual components are defined as follows:

- Disability progression on the EDSS is defined as an increase of ≥1.0 point from the baseline EDSS score when the baseline score is ≤5.5 or an increase of ≥0.5 points from the baseline EDSS score when the baseline score is >5.5.
- Disability progression on the T25-FW test is defined as an increase (worsening) of ≥20% from the baseline score.

- Disability progression on the 9-HPT is defined as an increase (worsening) of ≥20% from the baseline score.

Change from baseline in EDSS-Plus refers to the following three different endpoints:

- Change from baseline in EDSS score
- Change from baseline in 9-HPT
- Change from baseline in T25-FW.

Analyses of the three endpoints will be conducted separately.

MSIS-29v2 evaluates the specific physical and psychological impact of MS from a participant's perspective (4). This patient reported outcome instrument has 2 subscales: 1) a physical impact score (20 items) and 2) a psychological impact score (9 items). The physical and psychological impact subscales of the MSIS-29v2 range from 0 to 100, with higher scores indicating greater physical or psychological impact. 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:

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

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:

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

For participants 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), then the missing item score will be imputed by the average score of the available items in the domain. For each domain, the sum score is computed as the sum of scores for all completed items and the imputed scores for the missing items and then transformed to a score on a 0-100 scale using the formulae mentioned above.

MSWS-12 measures the impact of walking impairment in participant with MS (5). This patient reported outcome instrument has 12 items with a global score ranging from 0 to 100. A higher score indicates better quality of life. The sum score of MSWS-12 is computed by summing all 12 items and then transformed to a score on a 0 – 100 scale using the formula below:

$$\frac{100 \times (\text{sum score} - 12)}{60 - 12}$$

Similarly to MSIS-29v2, for MSWS-12 measurements that are partially missing, if at least 50% of the items have been completed (a minimum of 6 items), the missing item score will be imputed by the average score of the available items. The sum score is computed as the sum of scores for all

completed items and the imputed scores for the missing items and then transformed to a score on a 0-100 scale using the formula mentioned above.

3.3.2 Main analytical approach

All secondary efficacy endpoints in Part A and Part B will be analyzed using the randomized and treated population. Participants who permanently discontinue the IMP or permanently modify the background DMT will be asked and encouraged to stay in the study for the remaining visits. In this case, all measurements collected after the treatment status change during the planned treatment periods will be included in the analysis.

Categorical endpoints with count data

Categorical endpoints with count data include:

Part A

- Average number of new gadolinium- (Gd-) enhancing T1-hyperintense lesions per scan from randomization to Week 48
- Annualized rate of new/enlarging T2-hyperintense lesions from randomization to Week 48
- Annualized relapse rate (ARR) of RMS population (relapsing SPMS and RRMS) from randomization to Week 48

Part B

- Number of new gadolinium- (Gd-) enhancing T1-hyperintense lesions from Week 48 to Week 96
- Annualized rate of new/enlarging T2-hyperintense lesions from Week 48 to Week 96
- ARR of RMS population (relapsing SPMS and RRMS) from randomization to Week 96

A negative binomial regression model with robust variance estimation will be used to analyze these categorical endpoints with count data. The robust variance can be estimated by specifying the participant identifier in the repeated statement using SAS PROC GENMOD (version 9.4 or later).

For analyzing the average number of new gadolinium- (Gd-) enhancing T1-hyperintense lesions per scan and annualized rate of new/enlarging T2-hyperintense lesions, the model will include the total count of the respective type of lesion across all scheduled MRI scans during the observation duration as the response variable, with treatment group, baseline lesions count and MS clinical subtype (RRMS, SPMS, or PPMS) as covariates. For Part B endpoints, the baseline lesions count will be replaced by the Week 48 lesions count. The offset variable will be the log transformed number of MRI scans during the observation duration for T1 lesions and log transformed observation duration for T2 lesions. The observation duration for Part A is from randomization to the last MRI scan on or prior to Week 48. The observation duration for Part B is from the end of Week 48 to last MRI scan on or prior to Week 96. The average number of lesions per scan (95% CI) for Gd-enhancing T1-hyperintense lesions and the annualized rate (95% CI) of new/enlarging T2-hyperintense lesions for each treatment group will be estimated from the model along with the relative risk (95% CI, p-value) for SAR443820 compared to placebo. The adjusted mean value

within treatment group is based on a sample size dependent weight within each categorical covariate (ie, weight by overall proportion of participants in each stratum or categorical factor, also called “population weight”).

For ARR of RMS population (relapsing SPMS and RRMS), a similar negative binomial regression model will be used. The model will include the total number of relapses per participant occurring in the planned treatment period as the response variable, with treatment group as the covariate. Log transformed observation duration from randomization to Week 48 for participants complete Part A, or from randomization to the last contact date for participants discontinued in Part A will be the offset variable for the ARR of RMS population in Part A. Log transformed duration from randomization to Week 96, except when a participant withdraws early from the study, in which case, the log transformed duration from randomization to the last contact of the participant before Week 96 will be the offset variable for the ARR of RMS population in Part B. The estimated ARR for each treatment group and corresponding 2-sided 95% CI will be derived from the negative binomial model where the adjusted mean value within treatment group is based on population weight. The relative reduction in ARR with SAR443820 compared to placebo, its 2-sided 95% CI and p-value will be provided.

Time-to-event endpoints

Time-to-event endpoints defined in Part A include:

- Time to onset of 12-week CDP from baseline as assessed by the EDSS score
- Time to onset of sustained 20% increase in 9-HPT confirmed over at least 12 weeks
- Time to onset of sustained 20% increase in T25-FW confirmed over at least 12 weeks

Time-to-event endpoint defined in Part B include:

- Time to onset of 24-week CDP from baseline as assessed by the EDSS score
- Time to onset of sustained 20% increase in 9-HPT confirmed over at least 24 weeks
- Time to onset of sustained 20% increase in T25-FW confirmed over at least 24 weeks
- Time to onset of CCDP, confirmed over at least 24 weeks (6-month CCDP), by the EDSS Plus composite (EDSS score increase, or 20% increase in the T25-FW test, or 20% increase in the 9-HPT)

These time to event endpoints will be analyzed by a Cox proportional hazards model. The covariates included in the model are treatment group and MS clinical subtype (RRMS, SPMS, PPMS). The hazard ratio between treatment groups, its 95% confidence interval and the p-value will be estimated from this model with robust variance estimation (6). Comparison between treatment groups will also be assessed by a log-rank test stratified by MS clinical subtype (RRMS, SPMS, PPMS). Kaplan-Meier (KM) plots of the cumulative incidence rate will be provided by treatment group to depict the course of occurrence of event over time. The proportion of participants with events at given time points (eg, 12, 24, 36, 48, etc.) will be calculated using the KM estimates.

Continuous endpoints

Continuous endpoints include:

Part A

- Change from baseline in EDSS score at Week 48
- Change from baseline in 9-HPT at Week 48
- Change from baseline in T25-FW at Week 48
- Percent change from baseline in BVL at Week 48
- Change from baseline in the volume, number, and normalized T1 intensity of SELs at Weeks 12, 24, 36 and 48
- Change from baseline in the number of PRL at Weeks 12, 24, 36 and 48.

Part B

- Percent change from baseline in BVL at Week 96
- Change from baseline in the volume, number, and normalized T1 intensity of SELs to Week 96
- Change from baseline in the number of PRL at Week 96
- Change from baseline in EDSS at Week 96
- Change from baseline in 9-HPT at Week 96
- Change from baseline in T25-FW at Week 96
- Change from baseline in MSIS-29v2 physical and psychological domains score at Week 96
- Change from baseline in MSWS-12 at Week 96

Summary statistics including mean, standard deviation, median, interquartile range (Q1, Q3), minimum, and maximum will be provided for these continuous endpoints at scheduled visits. Only non-missing data will be included. For change/percent change from baseline, only participants with baseline and at least one post-baseline assessment will be summarized.

Normalized T1 intensity of SELs may be explored using linear mixed model. The model will include the treatment group (SAR443820 or placebo), MS clinical subtype (RRMS, SPMS, or PPMS), time (from randomization to the date of MRI sample being collected), treatment-by-time interaction, baseline value of normalized T1 intensity of SELs as the fixed effects. Both intercept and the coefficient for time are assumed to have participant-level random effects and these two random effects are assumed to be independent.

The other continuous endpoints will be analyzed using a similar MMRM approach as described for the primary endpoint of Part A in Section 3.2.2.1, except that the log-transformed baseline sNfL levels will be replaced by the baseline values of the corresponding endpoint. Appropriate transformation will be used if necessary (eg, log-transform). Plots of least squares means

(±standard error) over time will be provided. Plots of mean (± SE) and/or mean change (± SE) over time will also be provided.

No sensitivity or subgroup analyses will be conducted.

3.4 TERTIARY/EXPLORATORY ENDPOINT(S) ANALYSIS IN PART A

3.4.1 Definition of endpoint(s)

The tertiary endpoints detailed in this section are mentioned in Table 1 under Section 1.2 for Part A. Refer to Section 3.3.1 for the details of MSIS-29v2 and MSWS-12 assessments.

3.4.2 Main analytical approach

All tertiary efficacy endpoints in Part A will be analyzed using the randomized and treated population. All measurements collected after the treatment status change during the planned treatment period for participants who discontinue the IMP or permanently modify the background DMT will be included in the analysis.

Tertiary endpoints are all continuous endpoints. Summary statistics including mean, standard deviation, median, interquartile range (Q1, Q3), minimum, and maximum will be provided at scheduled visits in Part A. Only participants with baseline and at least one post-baseline assessment will be summarized. A similar MMRM approach as described for the primary endpoint of Part A in Section 3.2.2.1 will be used, except that the log-transformed baseline sNfL levels will be replaced by the baseline values of the corresponding endpoint. Appropriate transformation will be used if necessary (eg, log-transform). Plots of least squares means (±standard error) over time will be provided. Plots of mean (± SE) and/or mean change (± SE) over time will also be provided.

No sensitivity or subgroup analyses will be conducted.

3.5 MULTIPLICITY ISSUES

Not applicable.

3.6 SAFETY ANALYSES

All safety analyses will be performed on the safety population as defined in Section 2, unless otherwise specified, using the following common rules:

- 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 population (eg, exposed but not randomized) will be provided separately.

In Part A, the safety analyses will be carried out with participants by the actual treatment received, irrespective of the treatment the participant has been randomized to.

Part A + Part B Safety Analyses

For categorical safety measures, such as AE, PCSA, shift summaries, data throughout both Part A and Part B will be summarized in three groups:

- Placebo/SAR443820 group includes participants who actually received placebo in Part A and SAR443820 in Part B. For participants in this group, only the data during SAR443820 exposure in Part B will be summarized. The baseline will be defined as the last available value prior to the first dose of SAR443820 in Part B.
- SAR443820/SAR443820 group includes participants who actually received SAR443820 in Part A, regardless of whether they entered Part B. For participants in this group, all data during SAR443820 exposure will be summarized. The baseline will be defined as the last available value prior to the first dose of the study intervention (same baseline defined for Part A in Section 3.1).
- Overall SAR443820 group includes participants from Placebo/SAR443820 group and SAR443820/SAR443820 group.

For continuous safety measures collected at longitudinal time point, such as labs and vital signs, data throughout both Part A and Part B will be summarized in two groups:

- Placebo/SAR443820 group includes participants who actually received placebo in Part A, regardless of whether they enter Part B. For participants in this group, all data during the on-treatment period of Part A and Part B will be summarized at each protocol scheduled visit. The baseline will be defined as the last available value prior to the first dose of study intervention (same baseline defined for Part A in Section 3.1).
- SAR443820/SAR443820 group includes participants who actually received SAR443820 in Part A, regardless of whether they entered Part B. For participants in this group, all data during the on-treatment period of Part A and Part B will be summarized at each protocol scheduled visit. The baseline will be defined as the last available value prior to the first dose of study intervention (same baseline defined for Part A in Section 3.1).

3.6.1 Extent of exposure

The extent of IMP exposure will be assessed by the duration of IMP exposure and summarized within the safety population.

Duration of IMP exposure

Duration of IMP exposure in Part A is defined as last IMP administration date in Part A – first IMP administration date + 1 day, regardless of unplanned intermittent discontinuations. If the date of the last dose of IMP in Part A is missing, the duration of IMP exposure in Part A will be left as missing.

Duration of cumulative SAR443820 exposure in Part A + Part B is defined as the last dose date of SAR443820 – first dose date of SAR443820 +1 day, regardless of unplanned intermittent discontinuations. If the last dose date of SAR443820 is missing, this duration will be left as missing.

Duration of IMP exposure in Part A will be summarized quantitatively and categorically by the treatment group participants actually received in Part A (Placebo versus SAR443820):

- > 0 and ≤ 12 weeks
- > 12 and ≤ 24 weeks
- > 24 and ≤ 36 weeks
- > 36 and ≤ 48 weeks
- > 48 weeks

Duration of cumulative SAR443820 in Part A + Part B will also be summarized quantitatively and categorically by actual treatment (Placebo/SAR443820, SAR443820/SAR443820 and overall) within the safety population:

- > 0 and ≤ 12 weeks
- > 12 and ≤ 24 weeks
- > 24 and ≤ 36 weeks
- > 36 and ≤ 48 weeks
- > 48 and ≤ 72 weeks
- > 72 and ≤ 96 weeks
- > 96 weeks

Additionally, the cumulative duration of treatment exposure (expressed in participant-years) will be provided for Part A and Part A + Part B separately.

Treatment compliance

A given administration will be considered noncompliant if the participant did not take the number of administrations as required by the protocol. No imputation will be made for participants with missing or incomplete data.

Percentage of treatment compliance for a participant will be defined as the number of administrations that the participant was compliant divided by the total number of administrations that the participant was planned to take from the first administration of IMP up to the actual last administration of IMP.

Treatment compliance will be summarized quantitatively and categorically (<80%, ≥80%) for Part A and Part A + Part B using the similar format as stated above for exposure.

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 database lock.

The AEs will be analyzed in the following 4 categories:

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

Similarly, if occurring, 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. 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 period, using the maximum (worst) severity by treatment phase.

The AE tables will be sorted as indicated in Table 4.

Table 4 - Sorting of AE tables

AE presentation	Sorting rules
SOC and PT	By the internationally agreed SOC order and decreasing frequency of PTs[a]
AESI category and PT	By AESI category (protocol order) and decreasing frequency of PTs[a]
PT	By decreasing frequency of PTs[a]

a Sorting will be based on the SAR443820 treatment group incidence, alphabetic order in case of equal frequency.

Analysis of all adverse events in Part A and Part A + Part B

The overview of TEAE will be generated presenting the number (%) of participants with:

- Any TEAE
- Any severe TEAE
- Any treatment-emergent SAE
- TEAE leading to death
- Any TEAE leading to permanent intervention discontinuation of IMP
- Any treatment emergent AESI
- Any TEAE considered by the investigator as related to IMP

An additional overview summary including the number and rate of events will be provided.

The AE summaries of Table 5 will be generated with number (%) of participants experiencing at least one event.

Listings of SAEs from randomized participants, AEs from participants who are treated but not randomized, AEs leading to treatment/study discontinuation and severe AEs will be provided.

Table 5 - Analyses of adverse events

Type of AE	MedDRA levels
All TEAEs	Primary SOC and PT
Common TEAE (≥2% in any group)	PT
TEAEs related to IMP as per Investigator's judgment	Primary SOC and PT
	PT
TEAE by maximal intensity	Primary SOC and PT
Treatment emergent SAEs	Primary SOC and PT
	PT
Treatment emergent SAEs related to IMP as per Investigator's judgment	Primary SOC and PT
Treatment emergent AESIs	AESI category and PT
TEAEs leading to permanent intervention discontinuation	Primary SOC and PT
	PT
TEAEs leading to death[b]	Primary SOC and PT
Pretreatment AEs	Overview[a]
	Primary SOC and PT
Post-treatment AEs	Overview[a]
	Primary SOC and PT

Type of AE	MedDRA levels

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

Analysis of deaths in Part A and Part A + Part B

In addition to the analyses of deaths included in Table 5, the number (%) of participants in the following categories will be provided:

- Deaths during the treatment-emergent period and post-treatment period
- Deaths in non-randomized participants or randomized but not treated participants

Analysis of adverse events of special interest (AESIs) in Part A and Part A + Part B

Adverse events of special interest (AESIs) will be selected for analyses as indicated in Table 6. Number (%) of participants experiencing at least one event will be provided for each event of interest. Tables will be sorted as indicated in Table 4. A listing of AESIs will also be provided.

Table 6 - Selections for AESIs

AESIs	Selection
Pregnancy of a female participant or female partner of a male participant	Dedicated CRF page (AECAT="PREGNANCY DATA")
Symptomatic overdose (serious or non-serious) with IMP	Dedicated CRF page (AECAT="OVERDOSE DATA"; must be symptomatic and AESI marked "Y"
Convulsions, seizures	CMQsn00079 for selection and AESI marked "Y"
Serious infections meeting SAE definitions criteria	SOC of Infections and Infestations, both AESER and AESI marked "Y"
Increase in ALT > 3 x ULN	Dedicated CRF page (AECAT="ALT INCREASE DATA") and AESI marked "Y"

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. They will be converted into standard international units and conventional unit, if applicable.

- Hematology:
  - Red blood cells and platelets and coagulation: hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, red blood cell count, platelet count
  - White blood cells: white blood cell count, neutrophils, lymphocytes, monocytes, basophils, eosinophils

- Clinical chemistry:
  - Metabolism: glucose, total protein, albumin, creatine phosphokinase
  - Electrolytes: sodium, potassium, calcium
  - Renal function: creatinine, blood urea nitrogen
  - Liver function: alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, total and direct bilirubin
  - Pregnancy test: Serum β-human chorionic gonadotropin (all female participants)
  - Hepatitis screen: human immunodeficiency virus antibody, hepatitis B surface antigen, hepatitis C virus antibody
- Urinalysis:
  - Urinalysis for quantitative analysis: pH, specific gravity, proteins, glucose, blood, ketones, bilirubin, urobilinogen, nitrite, leukocyte esterase (by dipstick)
- Vital signs: heart rate, systolic and diastolic blood pressure, weight, respiratory rate, temperature
- ECG variables: heart rate, PR, QRS, QT, and corrected QTc (according to Bazett and Fridericia)

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 (ULOQ)will be replaced by ULOQ value for quantitative analysis.

Quantitative analyses

When relevant, for laboratory variables, vital signs and ECG variables above, descriptive statistics for results and changes from baseline will be provided for each analysis visit window during the on-treatment period for Part A and Part A + Part B. These analyses will be performed using central measurements only (when available) for laboratory variables.

For each laboratory parameter, vital sign parameter and ECG parameter, mean changes from baseline with the corresponding standard error will be plotted over time.

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 the database lock. For parameters for which no PCSA criteria are defined, similar analyses will be done using the normal range, if applicable.

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

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

- Normal/missing
- Abnormal according to PCSA (normal range) criterion or criteria

Additional analyses for drug-induced liver injury

The following additional analyses will be performed for drug-induced liver injury in Part A and Part A + Part B:

- Time to onset of the initial ALT elevation >3 x ULN, the initial AST elevation >3 x ULN, and the initial ALT or AST elevation >3 x ULN with total bilirubin elevation >2 x ULN during the treatment-emergent period will be presented by treatment group with Kaplan-Meier curves.
- 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).

3.6.3.2 Analysis of suicidality assessment

The number (%) of participants with suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent based on the C-SSRS during treatment will be summarized. A shift table for baseline versus during treatment responses in Part A and Part A + Part B will be provided according to the categories of no suicidal ideation or behavior, suicidal ideation and suicidal behavior.

3.7 OTHER ANALYSES

3.7.1 Pharmacokinetic (PK) analyses

Plasma concentration of SAR443820 will be summarized by study intervention in the PK population using arithmetic and geometric means, standard deviation, standard error of the mean, coefficient of variation, minimum, median and maximum per sampling time. For drug-treated participants, where concentration values are below the lower limit of quantification (LLOQ), one-half of the LLOQ will be used. Values will be expressed in the tables with no more than three decimal places. These analyses will be performed for all participants received SAR443820 and by baseline body weight (by quartiles).

The population PK analyses will be presented separately from the main clinical study report.

3.7.2 Pharmacodynamic (PD) analyses

PD parameters include sNfL, CHI3L1, sGFAP (excluding China), IL1β, IL6, IL8, TNFα, CCL3 and CCL4. NfL data will be analyzed under primary endpoint as described in Section 3.2. The rest of PD parameters will be summarized in Part A by actual treatment group in the safety population using the following descriptive statistics: mean, geometric mean, median, standard deviation, coefficient of variation, minimum, and maximum. The corresponding change from baseline values will also be provided. Plots of mean (±SD) and/or mean change (±SD) over time will also be provided for better visualization.

3.8 INTERIM ANALYSES

No interim analysis is planned.

4 SAMPLE SIZE DETERMINATION

Sample size calculations were performed to ensure reasonable accuracy for the estimation of the ratio of the SAR443820 arm sNfL geometric mean relative to the placebo arm sNfL geometric mean. The width (when a GMR of [redacted] is observed, which corresponds to a 30% reduction relative to placebo) of the 2-sided 90% CI, using a coverage correction to ensure coverage probability of [redacted]%, is displayed for various sample sizes in Table 7:

Table 7 - Sample size determination

	Log-transformed data					Corresponding geometric mean ratio (GMR)			
N	**Difference in means**	**Standard deviation (SD)**	**90% confidence interval (CI) half width***	**Lower limit**	**Upper limit**	**GMR**	**Lower limit**	**Upper limit**	**Total width (asymmetric)**
[redacted]									

(*): with a coverage probability (probability that the observed interval will be no longer than the specified width) of 90%.

For an observed[redacted] reduction relative to placebo, a sample size of 168 participants (to account for an anticipated 17% dropout rate) will provide an upper limit below 0.90 (10% of reduction) for the 2-sided 90% CI, which is deemed sufficient.

5 SUPPORTING DOCUMENTATION

5.1 APPENDIX 1 LIST OF ABBREVIATIONS

AE:	adverse event
AESIs:	adverse events of special interest
ATC:	anatomic therapeutic category
CHI3L1:	chitinase-3-like protein 1
CI:	confidence interval
ECG:	electrocardiogram
e-CRF:	electronic case report form
HLT:	high level term
IE:	intercurrent event
ITT:	intent-to-treat
LLT:	lower-level term
MAR:	Missing at Random
MedDRA:	medical dictionary for regulatory activities
mITT:	modified intent-to-treat
MMRM:	Mixed model with repetead mesuares
PCSA:	potentially clinically significant abnormality
PT:	preferred term
SAP:	statistical analysis plan
SD:	standard deviation
sNfL:	serum neurofilament light chain
SOC:	system organ class
TE:	treatment-emergent
TEAE:	treatment-emergent adverse event
WHO-DD:	World Health Organization-drug dictionary

5.2 APPENDIX 2 PARTICIPANT DISPOSITIONS

The number (%) of participants included in each of the analysis populations listed in Table 3 will be summarized.

Screened participants are those with a signed informed consent. Screen failures are defined as participants who consent to participate in the study but are not subsequently randomized. The number of screened participants will be summarized along with the number (%) of screen failures overall and by reason(s).

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

- Exposed but not randomized, if applicable

- Randomized
- Randomized but not exposed, if applicable
- Randomized and exposed
  - Completed the Part A treatment period
  - Did not complete the Part A treatment period including main reason for permanent study intervention discontinuation and reason for intervention withdrawal by participant
- Completed the Part A study period
- Did not complete the Part A study period including main reason for study discontinuation
- Enter the Part B
  - Completed the Part B treatment period
  - Did not complete the Part B treatment period including main reason for permanent study intervention discontinuation and reason for intervention withdrawal by participant
- Completed the Part B study period
- Did not complete the Part B study period including main reason for study discontinuation
- Status at last contact (alive, dead)

Reasons for permanent study intervention and study discontinuation, “adverse event” and “other reasons” will be split as related versus not related to Covid-19, if applicable.

For all categories of participants (except for screened and nonrandomized) percentages will be calculated using the number of randomized participants as the denominator.

In addition, the number (%) of participants screened, screened-failed, randomized, with permanent study intervention discontinuation and with early study discontinuation for Part A/Part B study period will be provided by country and site. Listings of other reasons for study intervention discontinuation, and participants who are exposed but not randomized will be provided.

Protocol deviations

Critical and major protocol deviations (automatic or manual) will be summarized in the randomized population as well as displayed separately as related versus not related to Covid-19 if applicable in Part A and Part A + Part B, respectively. A listing of critical and major protocol deviations will be provided.

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

Demographics, baseline characteristics, medical surgical history

The following demographic and baseline characteristics, smoking and alcohol history, disease characteristics at baseline, and medical and surgical history will be summarized using descriptive statistics in the randomized population.

Demographic and baseline characteristics:

- age in years as a quantitative variable and in categories (≤40, >40)
- sex (Male, Female)
- race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Multiple, Unknown, Not reported)
- ethnicity (Hispanic or Latino, Not Hispanic or Latino, Unknown, Not reported)
- weight in kg as a quantitative variable
- BMI in kg/m^2 as a quantitative variable and in categories (<25, ≥25 to <30, ≥30)

Smoking and alcohol habits:

- smoking history (Never, Current, Former)
- cessation prior to screening for Former smokers in months
- cigarettes per day for smokers
- frequency of alcohol drinking in the past 12 months (Never, Occasional, At least monthly, At least weekly, At least daily)
- Number of standard alcohol drinks on typical day when drinking (1 or 2, >2)

Disease characteristics at baseline:

- MS type (PPMS, SPMS non relapsing, SPMS relapsing, RRMS)
- time since symptom onset in years
- time since diagnosis in years as a quantitative variable and in categories (<1, ≥1 to 5, ≥5 to 10, ≥10)
- time since most recent relapse in years
- number of relapses within the past 1 year as a quantitative variable and in categories (0, 1, 2, ≥3)
- number of relapses within the past 2 years as a quantitative variable and in categories (0, 1, 2, ≥3)
- baseline EDSS score as a quantitative variable

- time since most recent MRI prior to screening in months
- number of active Gd-enhancing T1 lesions on last MRI prior to screening as a quantitative variable and in categories (0, 1, 2, ≥3)

Baseline safety and efficacy parameters (apart from those listed above) will be presented along with the safety and efficacy summaries.

Relevant medical (or surgical) history collected in the eCRF 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 database lock and will be summarized by primary SOC and HLT (internationally agreed SOC order and decreasing frequency of HLTs in the Overall group).

Prior or concomitant medications

All medications will be coded using the version of the World Health Organization-Drug Dictionary (WHO-DD) currently in effect at Sanofi at the time of database lock.

- Prior medications are those the participant used prior to first IMP intake. Prior medications can be discontinued before first administration or can be ongoing during treatment period.
- Concomitant medications are any interventions received by the participant concomitantly to IMP during the treatment period.
- Post-treatment medications are those the participant took in the period running from the end of the concomitant medications period up to the end of the study, if applicable.
- 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.

Any prior DMT use will be summarized by the standardized medication name in the randomized population, sorted by decreasing frequency in the Overall group. Additionally, prior, concomitant, and post-treatment (if applicable) medications will be summarized for the randomized and exposed population by anatomic and therapeutic level. Concomitant medications will be summarized in Part A and Part A + Part B respectively. The summaries will be sorted by decreasing frequency of anatomic category (ATC) based on incidence in the SAR442168 group. 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.

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 and PD variables.

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.

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 for efficacy in Part A

Table 10 - Analyses window definition for safety in Part A

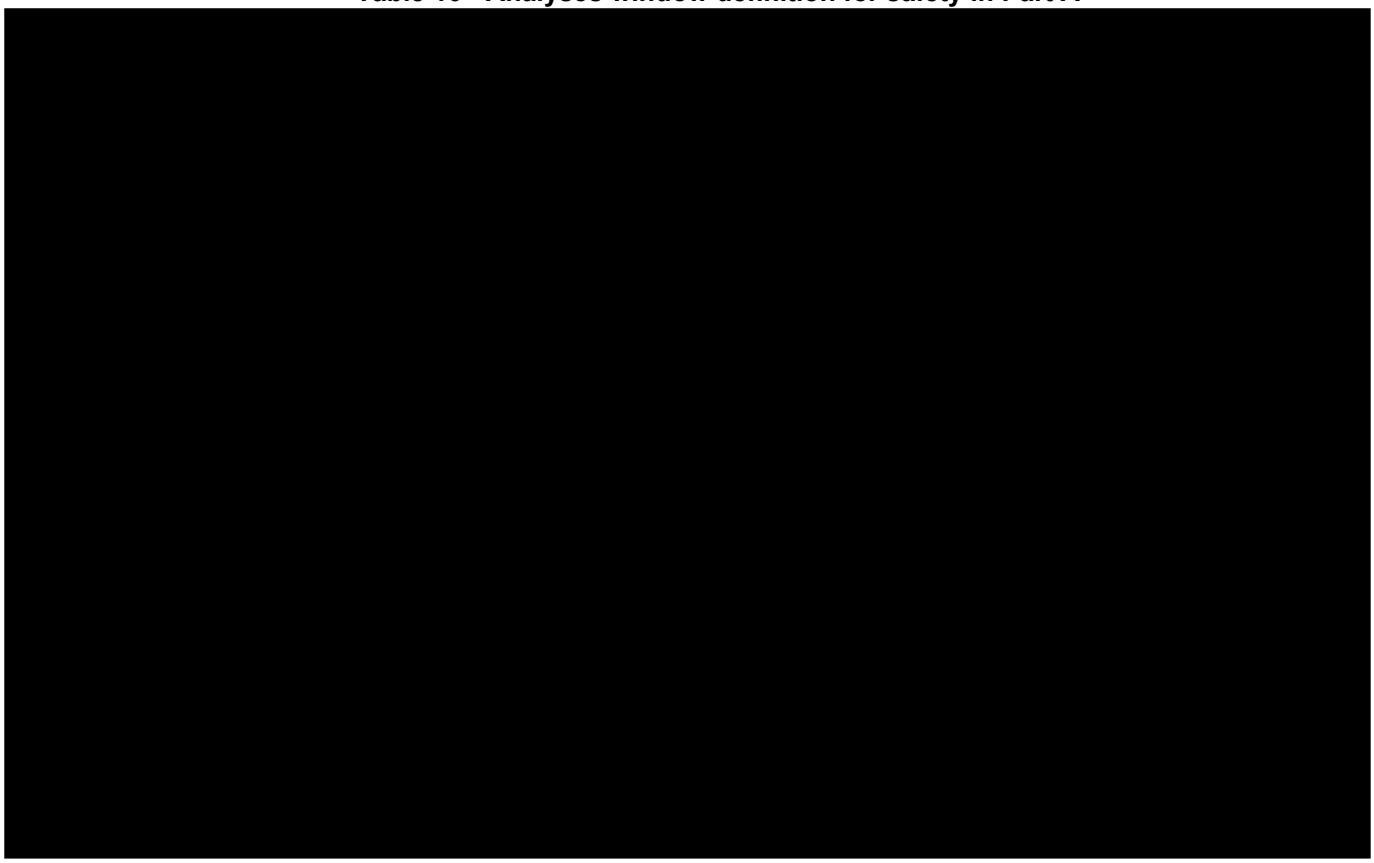

Table 11 - Analyses window definition for safety in Part B

Table 12 - Analyses window definition for PD in Part A

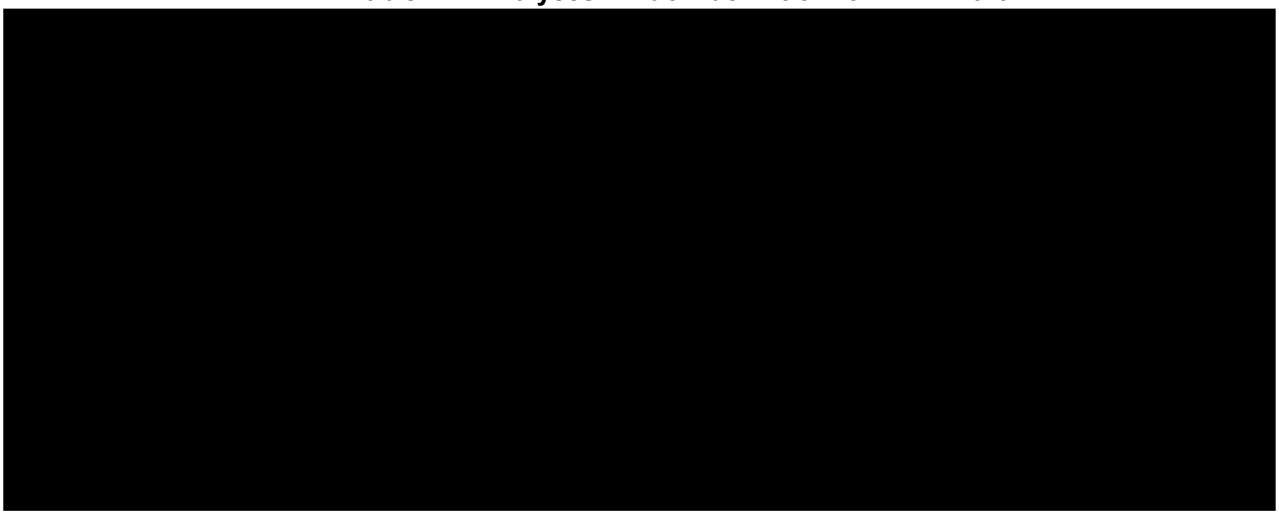

Table 13 - Analyses window definition for PD in Part B

Unscheduled visit measurements of laboratory data, vital signs and ECG will be used for computation of baseline, the last on-treatment value, analysis according to PCSA, and the shift summaries for safety. They will also be included in the by-visit summaries (central lab only for laboratory data) if they are re-allocated to scheduled visits based on the analysis windows defined above.

6 REFERENCES

1. Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. Computational Statistics & Data Analysis. 2009 May 15;53(7):2583-95.

2. Kaifeng Lu, Mehrotra DV. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. Stat Med. 2010;29(4):474-88.

3. Cadavid D, Cohen JA, Freedman MS, Goldman MD, Hartung HP, Havrdova E, et al. The EDSS-Plus, an improved endpoint for disability progression in secondary progressive multiple sclerosis. Multiple Sclerosis Journal. 2017 Jan;23(1):94-105.

4. McGuigan C, Hutchinson M. The multiple sclerosis impact scale (MSIS-29) is a reliable and sensitive measure. J Neurol Neurosurg Psychiatry. 2004;75(2):266-9.

5. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). Neurology. 2003;60(1):31-6.

6. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. J Am Stat Assoc. 1989;84(408):1074-8.

Signature Page for VV-CLIN-0665575 v1.0
act16753-16-1-9-sap

Approve & eSign	[redacted] Clinical [redacted]

Approve & eSign	[redacted] Clinical [redacted]