STATISTICAL ANALYSIS PLAN

Protocol title:	**A Phase 3, randomized, double-blind, efficacy and safety study comparing SAR442168 to placebo in participants with nonrelapsing secondary progressive multiple sclerosis (HERCULES)**
Protocol number:	**EFC16645**
Compound number (INN/Trademark):	**SAR442168 tolebrutinib/not applicable**
Study phase:	**Phase 3**
Short Title:	**NRSPMS study of BTK inhibitor tolebrutinib (SAR442168) HERCULES**
Statistician:	[redacted]
Statistical project leader:	
Date of issue:	**30-Apr-2024**
Regulatory agency identifier number(s):	
IND:	140884
EudraCT:	2020-000647-30
NCT:	NCT04411641
WHO:	U1111-1246-7768
EUDAMED:	Not applicable
Other:	Not applicable

Date: 30-Apr-2024 Total number of pages: 45

TABLE OF CONTENTS

LIST OF TABLES

VERSION HISTORY

This amended statistical analysis plan (SAP) for study EFC16645 is based on the amended protocol 11 dated 20 November 2023. The first participant was randomized on 23 October 2020. The initial SAP V1 was approved before the interim analysis for futility was conducted. Amended SAP V2 reflected statistical section changes incorporated in amended protocol 10 versus amended protocol 8 and any major changes since amended protocol 10. This amended SAP specifies rule for imputation of missing baseline Gd-enhancing T1-hypertintense lesion presence or absence status and additional analyses to support integrated safety summary.

Major changes in statistical analysis plan

SAP Version	Approval Date	Changes	Rationale
1	10-Feb-2023	Original version changes from protocol amendment 8	
		More covariates, ie, baseline EDSS score and Gd-enhancing T1 lesions at baseline are specified in the Cox model for the primary endpoint and other related time-to-event endpoints	Based on clinical input
		Censoring approach is different: imputation of 6-month confirmation event status is performed only for participants who have an initial onset of disability progression but complete the study at the common study end date with 3-month confirmation and without 6-month confirmation.	Based on FDA feedback
		For subgroup analysis of prior therapy use, the categories are 0, 1, ≥2	Clarification
		Actigraphic analysis baseline is defined at time of enrollment in substudy	Clarification
2	30-Oct-2023	Changes already incorporated in Amended protocol 10: Update the imputation approach for primary analysis: imputation of 6-month confirmation event status is performed only for participants who meet 3-month CDP criteria and continue to meet the criteria for EDSS disability progression through the final study assessment, but do not reach 6-month confirmation.	Based on FDA feedback
		Changes from Amended protocol 10: Change offset variable for analysis of new/enlarging T2 lesions to time from screening MRI to last available MRI scan rather than observation duration.	Clarification

SAP Version	Approval Date	Changes	Rationale
3	30-Apr-2024	Include imputation of missing baseline data of Gd-enhancing T1-hypertintense lesion presence or absence. Impute as absent if all non-missing post baseline values for a participant are 0, otherwise assign randomly based on the baseline proportions in the overall population.	Baseline Gd-enhancing T1-hypertense lesion presence or absence is used as a covariate in the primary analysis of the primary endpoint and also for analysis of other disability related endpoints. Baseline imputation is included so that all participants are included in the analyses. The number of participants affected is minimal with no impact on results.
		Specify additional analyses in Section 5.5 to support the integrated summary of safety	For completeness

Additional minor changes have been made to improve clarity and consistency in both versions 2 and 3

1 INTRODUCTION

1.1 STUDY DESIGN

This is a Phase 3, randomized, double-blind (DB), 2-arm, placebo-controlled, parallel group, multicenter, event-driven (6-month confirmed disability progression [CDP]) trial with a variable treatment duration (minimum of approximately 18 to 24 months) in participants with nonrelapsing secondary progressive multiple sclerosis (NRSPMS).

Approximately 1700 participants will be screened to achieve up to 1290 participants randomly assigned to study intervention.

After a screening period of up to 28 days, participants will be randomly assigned at a 2:1 ratio to receive 60 mg oral tolebrutinib or matching placebo daily. Randomization will be stratified by age at screening (>40 versus ≤40 years) and geographic region (US versus non-US).

A month is defined as a period of 28 days by convention. As such, 6 months is defined as 24 weeks, 24 months as 96 weeks, and 48 months as 192 weeks.

Participants who complete the study intervention period (DB or open-label [OL] tolebrutinib if meeting 6-month CDP) may be offered the option to participate in a long-term safety (LTS) study. For participants who complete the study intervention period (DB or OL, if meeting 6-month CDP) and do not enter the LTS study, a final follow-up visit to collect safety data is performed 4 weeks after the last dose of study intervention. A participant is considered to have completed the study if he/she has completed all periods of the study including the end-of-study (EOS) visit, whether remaining on study intervention or not.

Study primary analysis will be conducted after study completion, when approximately 288 events of 6-month CDP have been observed, and final database lock.

1.2 OBJECTIVES AND ENDPOINTS

Table 1 - Objectives and endpoints

Objectives	Endpoints
Primary	
• To determine the efficacy of tolebrutinib compared to placebo in delaying disability progression in NRSPMS	• Time to onset of 6-month CDP defined as follows: - Increase of ≥1.0 point from the baseline EDSS score when the baseline score is ≤5.0, OR - Increase of ≥0.5 point when the baseline EDSS score is >5.0

Objectives	Endpoints
Secondary	
• To evaluate efficacy of tolebrutinib compared to placebo on clinical endpoints, MRI lesions, cognitive performance, physical function, and quality of life	• Time to onset of sustained 20% increase in the 9-HPT for at least 3 months • Time to onset of sustained 20% increase in the T25-FW for at least 3 months • Time to onset of 3-month CDP as assessed by the EDSS score • Total number of new and/or enlarging T2-hyperintense lesions as detected by MRI, 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 EOS visit • Time to onset of CDI defined as a ≥1.0 point decrease on the EDSS score from baseline confirmed over at least 6 months • Percent change in brain volume loss (BVL) as detected by MRI scans at the EOS compared to Month 6 • Change in cognitive function at the EOS compared to baseline as assessed by SDMT • Change in cognitive function at the EOS compared to baseline as assessed by CVLT-II, where available • Change in MSQoL-54 questionnaire score from baseline through the EOS
• To evaluate safety and tolerability of tolebrutinib	• Adverse events (AEs), serious AEs, AEs leading to permanent study intervention discontinuation, AEs of special interest, and potentially clinically significant abnormalities in laboratory tests, safety scales, ECG, and vital signs during the study period
• To evaluate population pharmacokinetics of tolebrutinib and relevant metabolite(s) in NRSPMS and its relationship to efficacy and safety	• Plasma concentration of tolebrutinib and relevant metabolite(s) (population PK assessment) at Months 6, 9, and 12
• To evaluate pharmacodynamics (PD) of tolebrutinib	• Change in plasma NfL levels at the EOS compared to baseline • Change in lymphocyte phenotype subsets in whole blood at the EOS compared to baseline in a subset of participants • Change in serum immunoglobulin level at baseline compared to the EOS • Change in serum Chi3L1 levels at baseline compared to the EOS

Objectives	Endpoints
Tertiary/exploratory	
• To evaluate efficacy of tolebrutinib on disease progression and activity in NRSPMS, assessed by other clinical and imaging measures and by self-reported assessment	• Time to onset of sustained 20% increase in the 9-HPT for at least 6 months • Time to onset of sustained 20% increase in the T25-FW for at least 6 months • Time to onset of a 4-point decrease in the SDMT, confirmed over at least 3 and 6 months • The proportion of participants with CDI confirmed over at least 6 months • The proportion of participants with CDI confirmed over at least 6 months and maintained until the EOS • Change from baseline to Months 12, 18, and 24 and to the EOS in the EDSS score, T25-FW test, 9-HPT, SDMT, and CVLT-II • Change from baseline to Months 12, 18, and 24 and to the EOS in modified MSFC-3, assessed as the composite of the T25-FW test, 9-HPT, and SDMT • Proportion of participants with NEDA-3 at Months 18, 24, 30, 36, and the EOS • The annualized adjudicated relapse rate (ARR) • Actigraphic analysis of activity counts and indices of change from baseline to the EOS summarized over time (in a subset of participants) • Change from baseline of total volume of T2-hyperintense lesions as detected by brain MRI at Months 18, 24, and the EOS • Total number of new Gd-enhancing T1hyperintense lesions as detected by MRI, defined as the sum of the individual number of new Gd-enhancing T1hyperintense lesions at all scheduled visits starting after baseline up to and including the EOS visit • Change from baseline by visit in the volume of T1-hypointense lesions and cumulative number of new T1 hypointense lesions • MTR recovery at EOS in new MTR lesions detected at Months 6 and 12 • Change in number of phase rim lesions in SWI MRI from baseline through the EOS (subset of centers with capacity of 3T MRI) • Number and volume of slowly evolving lesions (SELs) • Normalized T1 (nT1) intensity evolution in SELs
• To evaluate the treatment effect of tolebrutinib via changes in participants' health-related quality of life (HRQoL), and working capacity	• Change in EQ-5D-5L from baseline by visit over time

1.2.1 Estimands

The primary estimand for the primary endpoint is summarized in Table 2. More details are provided in Section 3.

In all analyses, the comparison of interest will be the comparison of tolebrutinib versus placebo.

Table 2 - Summary of primary estimand for main endpoints

Endpoint Category (estimand)	Estimands			
	Endpoint	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
Primary objective: To evaluate the efficacy of tolebrutinib compared to placebo in delaying disability progression in NRSPMS				
Primary endpoint (treatment policy estimand)	Time from randomization to onset of 6-month CDP	ITT	Regardless of completion of the treatment period (treatment policy)	Treatment effect on time to onset of 6-month CDP, as measured by the hazard ratio for tolebrutinib compared to placebo, analyzed by a Cox proportional hazards model with terms for treatment group, age at screening (>40, ≤40 years), geographical region (US, non-US), baseline EDSS score and baseline Gd-enhancing T1 lesions (present, absent). The comparison of treatment groups will also be assessed using the log-rank test stratified by age at screening (>40, ≤40 years) and geographical region (US, non-US). Incomplete data will be imputed only for participants who complete the study, meet 3-month CDP criteria, and continue to meet criteria for EDSS disability progression through the final study assessment, but do not reach 6-month confirmation. For other participants without a confirmed 6-month CDP or with premature study discontinuation prior to 6-month confirmation after onset of an event, the participants will be censored at the date of last EDSS measurement.

2 ANALYSIS POPULATIONS

The following populations for analyses are defined:

Table 3 - Populations for analyses

Population	Description
Enrolled	All participants who signed the ICF.
Randomly Assigned to Study Intervention/Randomized	All participants with a treatment kit number allocated and recorded in the IRT database, regardless of whether the treatment kit was used or not. Participants treated without being randomized will not be considered randomized and will not be included in any efficacy population.
ITT	The primary efficacy population will be the ITT population, defined as all randomly assigned participants. All efficacy analyses will be conducted according to the treatment group allocated by the randomization schedule, irrespective of the treatment received
Safety	All participants randomly assigned and exposed to study intervention, regardless of the amount of exposure, analyzed according to the treatment actually received. Randomized participants for whom it is unclear whether they took the study medication will be included in the safety population as randomized. The pharmacodynamic (PD) analyses will be performed on the safety population.
Pharmacokinetics	All participants in the safety population with at least one non- missing PK sample after first dose of the study intervention. Participants will be analyzed according to the treatment actually received.

ICF: informed consent form; ITT: intent to treat; PK: pharmacokinetic(s); IRT: interactive response technology

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.

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 treatment group for as-treated analyses will be tolebrutinib 60 mg.

3 STATISTICAL ANALYSES

3.1 GENERAL CONSIDERATIONS

The baseline values of efficacy parameters are generally defined as the last available value prior to the first dose of study medication unless otherwise specified. In most cases the assessment should be prior to the randomization call, but it is possible that some assessments are performed between the actual randomization call and the first intake of investigational medicinal product (IMP). For the Expanded Disability Status Scale (EDSS) score, the baseline value will be taken as the average of the screening and randomization visit values. If one of the values is missing, the non-missing value will be used as baseline. The baseline value of safety parameters is defined as the last available value prior to the first dose of IMP. For the C-SSRS, the baseline value will be the worst assessment between the past 6 months evaluation at screening and the C-SSRS 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 (CIs) will be provided for treatment comparisons.

To include all participants in analyses where baseline Gd-enhancing T1-hyperintense lesion presence or absence is used as a covariate, baseline status (present or absent) will be imputed for participants with a missing baseline count: if subsequent post-baseline counts for a participant are all 0, then impute baseline status as absent, otherwise, impute baseline status randomly with seed = 16645 based on the proportion of presence/absence at baseline in the overall population. Of note, the number of participants affected is minimal with no impact on results.

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 ITT population, and safety data will be analyzed in the safety population, unless specified otherwise.

Unless otherwise specified, analyses will be performed by treatment group and overall for baseline and demographics characteristics.

Observation period

The observation periods for safety will be divided into 4 epochs:

- The **screening** (**pre-treatment) period** is defined as the period from the signed informed consent date up to first administration of the study intervention.
- The **treatment period** (ie, treatment-emergent (TE) period) is defined as the period from the first administration of study intervention to
  - The earlier of last administration of the study intervention + 10 days or first IMP in the LTS study for participants not receiving tolebrutinib OL treatment,
  - The first administration of tolebrutinib OL treatment for participants switching to OL therapy

- The **OL period** (ie, OL TE period), applicable only for participants switching to tolebrutinib OL treatment, is defined as the time from first administration of tolebrutinib OL treatment to the earlier of last administration of OL treatment + 10 days or first IMP in the LTS study
- The **post-treatment period**, if applicable, is defined as the period from the end of treatment period (or end of OL period, if applicable) to the participant's final post-treatment study contact date.

Summaries of OL period data will be presented by original randomized treatment group and overall.

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.

Contingency measures for a regional or national emergency that is declared by a governmental agency

Participants in Ukraine, Russia and Belarus who are lost to follow-up due to national emergency will be considered as prematurely discontinuing the study at the date of last contact (or start of war). For the primary endpoint of time to onset of 6M-CDP, unless already having a confirmed event at the time of last follow-up, participants will be censored at the date of the last EDSS assessment for the primary analysis. Other time-to-event endpoints will be handled similarly. For continuous endpoints as well as safety data, all observed data up to the last contact date will be included for assessment. For count data, all observed data up to the last contact date will be included in the analysis. No imputation will be performed for unobserved data/events that may happen after being lost to follow-up.

3.2 PRIMARY ENDPOINT(S) ANALYSIS

3.2.1 Definition of endpoint(s)

Time to onset of 6-month CDP is defined as the time from randomization to the onset of a sustained increase from baseline in EDSS score (of ≥1.0 point from the baseline EDSS score when the baseline score is ≤5.0, of ≥0.5 points when the baseline EDSS score is >5.0) confirmed after a minimum 6-month interval.

EDSS scores used for potential onset of disability progression may be observed by the Examining Investigator during routine quarterly visits or unscheduled visits. However, confirmation of CDP (after at least 6 months) must be at a routine quarterly visit (with the potential exception in case of contingency measures for a regional or national emergency). Results of all EDSS assessments obtained during routine or unscheduled visits throughout the minimum 6-month period (for the primary endpoint) will serve as the basis for conclusion. Once confirmation is achieved, the potential onset becomes the actual onset for time to onset of CDP.

In case of occurrence of adjudicated relapse, only EDSS assessments measured more than 90 days after relapse onset will be used to determine onset of disability progression. In addition, for the purpose of confirmation, only EDSS scores measured more than 90 days after the onset of an adjudicated relapse will be used. In case of such multiple sclerosis (MS) relapse, the next quarterly EDSS assessment >90 days after relapse onset will be used for CDP confirmation. The minimum increase in score required for progression must also be maintained for any non-confirmatory (ie, intervening) EDSS assessment(s) between the initial (onset) and confirmation EDSS scores.

Note: 6-month confirmation must be ≥154 days after onset. Given that visits are scheduled every 12 weeks with a time window of ± 7 days for conducting EDSS assessment, the minimum per protocol allowed time between 2 EDSS assessments scheduled 24 weeks apart is 154 days (6 months = 24 weeks = 168 days minus a 7-day window for each of the 2 visits, ie, 168 - 14 = 154).

3.2.2 Main analytical approach

The primary analysis of the primary endpoint, time to onset of 6-month CDP, will be performed on the ITT population. The null hypothesis is no treatment difference between tolebrutinib and placebo, and the alternative hypothesis is that there is a between-treatment difference. The study will be declared positive if the null hypothesis for the primary endpoint of 6-month CDP for tolebrutinib compared to placebo is rejected at the 2-sided 0.05 significance level.

In this primary ITT analysis:

- For participants who complete the study without an initial onset of disability progression, the participant will be censored at the date of last EDSS assessment.
- For participants who have an initial onset of disability progression but complete the study without 3-month confirmation, the participant will be censored at the date of last EDSS assessment.
- For participants who prematurely discontinue the study before 6-month confirmation of an onset of disability progression, regardless of having an initial onset, the participant will be censored at the date of last EDSS assessment.
- For participants who complete the study, meet 3-month CDP criteria, and continue to meet criteria for EDSS disability progression through the final study assessment, but do not reach 6-month confirmation, the 6-month CDP event status of the participant will be determined by an imputation approach. The details of this imputation approach are provided in Section 3.2.3.

Death

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 event will be calculated as

- (date of EDSS assessment at a tentative onset of disability progression – date of randomization + 1), or
- (date of death – date of randomization + 1) if a tentative onset date does not exist.

If a participant dies for other reasons before a 6-month CDP, the participant's event time will be censored at the date of last EDSS assessment.

Analytical Approach

The time to onset of 6-month CDP will be modeled by a Cox proportional hazards model with robust variance estimation (1). The covariates included in the model are treatment group, age at screening (>40, ≤40), region (US, non-US), baseline EDSS score and baseline gadolinium (Gd)-enhancing T1 lesions (present, absent). The hazard ratio, its 95% confidence interval and the p-value for comparing tolebrutinib to placebo will be estimated from this model. Comparison between tolebrutinib and placebo will also be assessed by a log-rank test stratified by age at screening (>40, ≤40 years) and region (US, non-US). Kaplan-Meier (KM) plots of the cumulative incidence rate will be provided by treatment group to depict the course of onset of 6-month CDP over time. The proportion of participants with events at given time points (eg, Months 6, 12, 18, 24, 30, …) and median time to 6-month CDP will be calculated using the KM estimates.

3.2.3 Multiple imputation for missingness due to end of study

For participants who meet 3-month CDP criteria, and continue to meet the criteria for EDSS disability progression through the final study assessment, but do not reach 6-month confirmation, the 6-month CDP event status (ie, confirmed or not confirmed) of the participant will be determined by a logistic regression model as follows:

$$Y_i \sim Bernoulli(p_{ij}), i = 1, \dots, n_j \; ; \; j = 1,2$$

$$logit(p_{ij}) = ln\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \beta_{3j}x_{3ij} + \beta_{4j}x_{4ij}$$

where Y_{ij} is the indicator for event status for participant i from group j; 1 = 6-month CDP, 0 = no 6-month CDP, p_{ij} is the probability that participant i from group j with a 3-month CDP has a 6-month CDP event, x_{1ij} is the age at screening (>40, ≤40 years), x_{2ij} is the geographic region (US, non-US), x_{3ij} is the baseline EDSS score, x_{4ij} is baseline Gd-enhancing T1 lesions (present, absent), and β_{kj} are the model parameters ($k = 0,1,2,3,4$; $j = 1,2$). The logistic model parameters β_{kj} will be estimated from participants with a 3-month CDP and who do not have missing value due to end of study. Separate logistic models will be estimated within each treatment arm. Then, these estimates will be used to calculate the predicted probability for participants who meet 3-month CDP criteria and continue to meet criteria for EDSS disability

progression through the final study assessment, and missingness due to end of study. Finally, the missing value for participants with an initial onset of disability progression and missingness for 6-month CDP due to end of study will be imputed via a randomly generated value from a Bernoulli distribution with p_{ij}.

Using the above methodology, 1000 complete data sets will be generated with imputed missing data via PROC MI in SAS®. Each of the data sets will be analyzed using the Cox proportional hazards model as described above. Log-rank test and KM plot will also be provided based on 1000 complete data sets based on the method in (2). The SAS MIANALYZE procedure will be used to generate statistical inferences by combining results from the 1000 analyses using Rubin's formula.

Sample SAS code is provided below:

```
proc mi data=DATAIN out=DATAOUT seed=16645 nimpute=1000;
  by  Treatment_group;
  var age_strata geographic_region EDSS_base Gd_T1;
  fcs logistic (Confirmation = age_strata geographic_region EDSS_base Gd_T1/
link=glogit);
run;

PROC MIANALYZE data= DATAOUT;
   modeleffects Estimate;
   stderr StdErr;
run;
```

DATAIN only includes participants who meet the imputation criteria mentioned above. Details of multiple imputation can be found in SAS support document (3).

3.2.4 Sensitivity analysis

The following sensitivity analyses, without any imputation, will be performed to evaluate the impact of missing data to the primary analysis results:

1. Participants with a potential onset of disability progression but without 6-month confirmation due to missing data, regardless of reason, will be considered to have confirmed progression with the time to onset being the time to the initial EDSS increase. Of note, a potential onset is invalidated if there is any non-confirmatory EDSS in the confirmation period that does not meet the minimum change required for progression.
2. Participants with missing or incomplete 6-month CDP will be treated as censored at the last EDSS assessment date.

For the above analyses, the same Cox proportional hazards model as described in Section 3.2.2, without imputation, will be used to estimate the hazard ratio, 95% CI, and p-value on tolebrutinib compared to placebo.

3.2.5 Subgroup analyses

To assess the homogeneity of the treatment effect across various subgroups, analyses will be performed on the primary endpoint across the following (no subgroup analysis will be performed if fewer than 5 participants in one category):

- Geographic region (US, non-US; Eastern Europe, Western Europe, North America, rest of the world [ROW])
- Age at screening (>40, ≤40 years)
- Sex (Male or Female)
- Race (White, Black or African American, Asian, Other).
- Baseline Gd-enhancing T1 lesions (present, absent)
- Baseline EDSS score (≤4.5, >4.5; ≤5.5, >5.5)
- Prior disease modifying therapy use (0, 1, ≥2)
- Duration since relapsing-remitting multiple sclerosis (RRMS) symptom onset (≤5, >5 to ≤10, >10 years)
- Adjudicated relapse during the study (yes, no)

Since adjudicated relapse during the study is a post randomization event, this subgroup analysis is purely exploratory.

The treatment effect (tolebrutinib compared to placebo) for the primary endpoint will be provided, as well as the corresponding 95% CI, for each subgroup separately, using the same methods as described in Section 3.2.2. Forest plots of hazard ratios and the corresponding 95% CIs will be provided.

Treatment by subgroup interaction and its p-value will be derived from a Cox proportional hazards model with terms for treatment group, age at screening (>40, ≤40), region (US, non-US), baseline EDSS score, baseline Gd-enhancing T1 lesions (present, absent), subgroup (if different than the aforementioned covariates) and treatment by subgroup interaction as covariates. If quantitative treatment by subgroup 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 are for assessing efficacy of tolebrutinib compared to placebo. 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 3.7.2 (PD).

3.3.1 Secondary endpoint(s)

3.3.1.1 Definition of endpoint(s)

The 9-hole peg test (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. The total time to complete the task is recorded. An increase of >20% from the baseline in the 9-HPT is considered meaningful worsening (4). To transform the data so that higher Z-scores correspond to an improved outcome and lower Z-scores correspond to deterioration, the inverse of the test values are taken prior to computing the Z-score. The baseline mean and standard deviation of all participants are used to create Z-scores:

$$Z = \frac{\frac{1}{2}\left(\frac{1}{trialarm,left} + \frac{1}{trialarm,right}\right) - Baseline\ Mean}{Baseline\ SD}$$

The times from the two trials for each hand are averaged, converted to the reciprocals of the mean times for each hand and then the two reciprocals are averaged, for example,

$$\frac{1}{trialarm,left} = \frac{1}{average\ of\ the\ two\ times\ of\ the\ left\ arm\ trails}$$

In case one of the two times for the same arm is missing, the available time will be used. In case 9-HPT was assessed for only one hand for reason other than “physical limitation”, then the Z-score will be set to missing. For a participant who could not complete the 9-HPT due to a “physical limitation”, the maximum time (300 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.

The Timed 25-Foot Walk (T25-FW) test 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). An increase of >20% from baseline in the T25-FW test is considered meaningful worsening (4). The baseline mean and SD of all participants are used to create Z-scores:

$$Z = (-1) \times \frac{Average\ score - Baseline\ Mean}{Baseline\ SD}$$

The negative of the Z-score is used so that improvement is a positive number and worsening a negative number. 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. If results from both trials are missing due to reasons other than “physical limitation”, the Z-score will be set to missing. In case of values below (above) the lower (upper) bound, the value will be set to the respective bound.

Time to onset of 3-month CDP is defined as the time from randomization to the onset of a sustained increase from baseline in EDSS score (of ≥1.0 point from the baseline EDSS score when the baseline score is ≤5.0, of ≥0.5 points when the baseline EDSS score is >5.0) confirmed after a minimum 3-month interval. The confirmation of 3-month CDP follows the same criteria as that of 6-month CDP.

Note: 3-month confirmation must be ≥70 days after onset. Given that visits are scheduled every 12 weeks with a time window of ± 7 days for conducting EDSS assessment, the minimum per protocol allowed time between 2 EDSS assessments scheduled 12 weeks apart is 70 days (3 months = 12 weeks = 84 days minus a 7-day window for each of the 2 visits, ie, 84 - 14 = 70).

Confirmed disability improvement is defined as a ≥1 point decrease from baseline in the EDSS score lasting at least 6 months.

Symbol Digit Modalities Test (SDMT) is used to assess processing speed, divided attention, visual scanning, tracking and motor speed. The number of correct substitutions and number of items completed within a 90 second interval (max = 110) are recorded. A decrease of 4 points from baseline on the SDMT is considered meaningful worsening (5). The baseline mean and standard deviation of all participants are used to create the Z-score for SDMT number of correct substitutions:

$$Z = \frac{SDMT\ score - Baseline\ Mean}{Baseline\ SD}$$

Z-scores will be used for analysis.

The California Verbal Learning Test®-II (CVLT-II) is a verbal learning and memory test. It consists of recall and recognition of a list of 16 words. For each assessment, 5 trials are completed. The total correct, total intrusions, and total repetitions across the 5 trials combined are recorded and converted to a standardized score by sex within each of 7 age groups (6). Total Correct Recall Trials 1–5 is scaled to a normalized T-score metric, which has a mean of 50 and SD of 10, with higher values indicating better performance. Total intrusions and total repetitions are normed on a linear Z-score metric, with a mean of 0, SD of 1.0, and increment values of plus or minus 0.5. Higher positive Z-scores on these measures reflect greater deficits. Standardized scores will be used for analysis.

Multiple Sclerosis Quality of Life-54 (MSQoL-54) is a multidimensional health-related quality of life measure that combines both generic and MS-specific items into a single instrument (7). The recall time, when relevant, is for the previous 4 weeks. It is composed of 54 items. There are 2 derived summary scores, physical health composite and mental health composite, 12 subscale scores: physical function, role limitations-physical, role limitations-emotional, pain, emotional wellbeing, energy, health perceptions, social function, cognitive function, health distress, overall quality of life, and sexual function, and 2 single-item measures: satisfaction with sexual function and change in health. The total number of items in each subscale is listed as the divisor for each subtotal. Due to missing data, the divisor can be less than that if not every item within a given scale has been answered. Each subscale score can then be calculated by summing the non-missing item responses and dividing by the number of non-missing item responses. Physical health

composite score is calculated as a weighted sum of physical function (weight = 0.17), health perceptions (weight = 0.17), energy (weight = 0.12), role limitations-physical (weight = 0.12), pain (weight = 0.11), sexual function (weight = 0.08), social function (weight = 0.12) and health distress (weight = 0.11). Mental health composite score is calculated as a weighted sum of health distress (weight = 0.14), overall quality of life (weight = 0.18), emotional well-being (weight = 0.29), role limitations-emotional (weight = 0.24) and cognitive function (weight = 0.15). All scores are transformed linearly to a common 0 (worst) to100 (best) score range.

3.3.1.2 Main analytical approach

3.3.1.2.1 Time to event endpoints

Time to onset of 3-month CDP will be analyzed using the same statistical methods as for the primary analysis of time to onset of 6-month CDP without imputation.

Time to onset of sustained 20% increase in the 9-HPT (T25-FW) for at least 3 months and time to onset of 6-month confirmed disability improvement (CDI) will be analyzed using similar statistical methods as for the primary analysis of time to onset of 6-month CDP, but without imputation. Participants with onset without confirmation will be treated as censored at the last endpoint assessment.

3.3.1.2.2 Number of new and/or enlarging T2 lesions

Number of new and/or enlarging T2 hyperintense lesions over the study period will be analyzed using a negative binomial regression model with robust variance estimation in the ITT population. The robust variance can be estimated by specifying the participant identifier in the repeated statement using SAS PROC GENMOD (version 9.4 or later). The model will include the total lesion count across all magnetic resonance imaging (MRI) scans at all scheduled post-randomization visits during the study period as the response variable, with treatment group, age at screening (>40, ≤40 years), geographic region (US, non-US), baseline EDSS score, and baseline number of T2 lesions as covariates. Log transformed observation duration from screening MRI to last available MRI scan will be the offset variable. The estimated mean number of lesions per year for each treatment group and the corresponding two-sided 95% CI will be estimated from the negative binomial model where 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”). The relative reduction on tolebrutinib compared to placebo, its 95% CI, and p-value will be provided. No imputation for missing data due to early study withdrawal will be performed.

3.3.1.2.3 Continuous endpoints

Percent change in brain volume at the EOS compared to Month 6 will be analyzed using a mixed-effect model with repeated measures (MMRM) approach in the ITT population for estimating the treatment effect with appropriate transformation if necessary (eg, log-transform). Estimates derived from the MMRM approach will be weighted by overall proportion of participants in

each stratum or categorical factor. The model will include percent change from Month 6 values at scheduled visits up to and including the EOS visit as the response variable, and treatment group, age at screening (>40, ≤40 years), geographic region (US, non-US), visit, treatment-by-visit interaction, Month 6 brain volume, and Month 6 brain volume-by-visit interaction as covariates; have an unstructured variance-covariance matrix; and be fitted using restricted maximum likelihood estimation. If this model fails to converge, the following variance-covariance structures will be tested in this order:

- Heterogeneous Toeplitz (Toeplitz with non-equal variances)
- Heterogeneous AR(1) (first-order autoregressive with non-equal variances)
- Heterogeneous CS (compound symmetry with non-equal variances)
- Toeplitz (equal variances and a separate correlation for each level of separation between the time points)
- AR(1) (first-order autoregressive, equal variances and exponentially decreasing correlations)
- CS (compound symmetry, equal variances and equal pairwise correlations across fixed time points).

The first (co)variance structure yielding convergence will be used as the primary analysis. Comparisons between treatment arms will be made using least squares means (LSM) contrasts at the EOS with denominator degrees of freedom estimated using the Kenward-Roger approximation (8). When a variance-covariance structure other than unstructured is used, the denominator degrees of freedom will be estimated using the between within method (DDFM=BW in SAS PROC MIXED) (9).

Change from baseline in SDMT number of correct substitutions, CVLT-II total correct and MSQoL-54 physical and mental health composite scores at the EOS will be analyzed using an analogous MMRM approach in the ITT population with the change from baseline values for the respective endpoint at scheduled visits up to and including the EOS visit as response variables, and treatment group, age at screening (>40, ≤40 years), geographic region (US, non-US), visit, treatment-by-visit interaction, baseline value for the endpoint being assessed and baseline value-by-visit interaction as covariates.

For each endpoint, difference in LSM, the corresponding 95% CI, and p-value will be estimated from the MMRM model for the comparison of tolebrutinib to placebo using weights for each stratum equal to the overall proportion of participants in each stratum (ie, “Population weight”). Plots of LSM [±standard error (SE)] over time will be provided. Participants who discontinue study intervention before the study end will be asked and encouraged to return for all remaining study visits and the additional off-treatment values measured up through the EOS will be included in the primary analysis. For participants who withdraw from the study before the EOS, values will be missing after study discontinuation and no imputation will be performed.

SDMT, CVLT-II and MSQoL-54 data are assumed a priori to be normally distributed. Given the large sample size, the impact of violation of normality should not be great. In order to assess the impact of deviation from the normality assumption, a sensitivity analysis using rank analysis of

covariance (ANCOVA), with the same covariates as for the MMRM analysis except VISIT and related interaction terms, will be performed to test the significance of the treatment effect.

Given that participants who reach the primary endpoint of 6M-CDP are eligible to switch to OL tolebrutinib, including data after the switch could impact the analysis of secondary endpoints: number of new/enlarging T2 lesions per year, time to onset of sustained 20% increase in the 9-HPT (T25-FW) for at least 3 months, and percent change in brain volume at EOS compared to Month 6. Sensitivity analyses excluding data after initiation of OL therapy will be performed for these endpoints.

No other sensitivity, supplementary, or subgroup analyses will be performed for the secondary endpoints.

3.4 TERTIARY/EXPLORATORY ENDPOINT(S) ANALYSIS

3.4.1 Definition of endpoint(s)

Multiple Sclerosis Functional Composite (MSFC)-3 is a multidimensional, three-component performance scale to assess the degree of impairment in MS patients. Here, the MSFC-3 score is defined as the average of the T25-FW, 9-HPT, and SDMT Z-scores. If any one of the T25-FW, 9-HPT, and SDMT Z-scores is missing, then MSFC-3 score is set to missing.

No evidence of disease activity-3 (NEDA-3) is defined as absence of all of the following:

- 6-month CDP
- Active MRI lesions (both new or enlarged T2-hyperintense lesions and Gd-enhancing T1-hyperintense lesions)
- Adjudicated MS relapse

EQ-5D-5L is a standardized health-related quality of life questionnaire that provides a simple, generic measure of health for clinical and economic appraisal. It consists of 2 pages, the EQ-5D descriptive system and the EQ visual analogue scale (VAS). The EQ-5D descriptive system comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. EQ-5D-5L health status will be converted into a single index value by using the value sets based on UK population and according to the crosswalk algorithm developed by B. van Hout (10). The index score is anchored at 1 (full health) with 0 (negative values) corresponding to states being as bad or worse than dead. The EQ VAS records the respondent's self-rated health on a vertical VAS ranging from 0 (worst imaginable health state) at the bottom to 100 (best imaginable health state) at the top. Overall, higher scores indicate high health utility.

For the purposes of this study, MS relapse is defined as a monophasic, acute or subacute onset of, new neurological symptoms or worsening of previous neurological symptoms with an objective change on neurological examination. Symptoms must:

1. be attributable to MS,
2. last for ≥24 hours, with or without recovery,
3. be present at normal body temperature (ie, no infection, excessive exercise, or excessively high ambient temperature), and
4. be preceded by ≥30 days of clinical stability (including no previous MS relapse).

Confirmation of MS relapse will be based on EDSS score (provided by the Examining Investigator), based on the following definition:

- A confirmed MS relapse is one accompanied by a clinically relevant change in the EDSS score performed by the Examining Investigator, ie, an increase of at least 0.5 points in the EDSS score, an increase of 1 point on 2 functional scores, or an increase of 2 points on 1 functional score, excluding changes involving bowel/bladder and cerebral functional score compared to the previously available rating (the last EDSS rating that did not occur during a relapse).

3.4.2 Main analytical approach

Summary statistics including mean, SD, median, interquartile range (Q1, Q3), minimum, and maximum will be provided for the following continuous data endpoints at scheduled visits over the study period including the EOS. Only non-missing data will be included. For change from baseline, only participants with baseline and post-baseline data will be summarized.

- EDSS score and change from baseline
- Number of Gd-enhancing T1 lesions and change from baseline
- Number of new Gd-enhancing T1 lesions
- Number of new/enlarging T2 lesions per 6-month interval
- T2 lesion volume and change from baseline
- Cumulative number of new T1-hypointense lesions
- T1-hypointense lesion volume and change from baseline.
- Number of phase rim lesions and change from baseline
- Normalized brain volume at screening and percent change in brain volume at post baseline time points
- Number of slowly evolving lesions at EOS compared to screening
- Volume of slowly evolving lesions at EOS compared to screening
- SDMT number of correct substitutions Z-score and change from baseline
- SDMT number of completed items and change from baseline
- CVLT-II standardized scores for total number correct and change from baseline
- CVLT-II standardized scores for total number of intrusions and change from baseline

- CVLT-II standardized scores for total number of repetitions and change from baseline
- MSFC-3 score and change from baseline
- T25-FW raw score and Z-score for leg function/ambulation and change from baseline
- 9-HPT raw score and Z-score for arm/hand function and change from baseline
- MSQoL-54 score and change from baseline for the 2 derived composite scores, 12 subscale scores and 2 single item scores
- EQ-5D-5L health state utility value score and change from baseline
- EQ VAS score and change from baseline

Plots of mean (±SD) and/or mean change (±SD) and/or median over time will also be provided for better visualization.

Time-to-event endpoints

Time-to-event endpoints include:

- Time to onset of sustained 20% increase in the 9-HPT for at least 6 months,
- Time to onset of sustained 20% increase in the T25-FW for at least 6 months
- Time to onset of a 4-point decrease in the SDMT confirmed over at least 3 months
- Time to onset of a 4-point decrease in the SDMT confirmed over at least 6 months

Participants with no event will be censored at the date of the last assessment for the endpoint. No imputation will be performed for missing data. These endpoints will be analyzed using similar statistical methods as for the primary analysis of time to onset of 6-month CDP.

Binary endpoints

Binary endpoints include:

- Proportion of participants with CDI confirmed over at least 6 months
- Proportion of participants with CDI confirmed over at least 6 months and maintained until the EOS
- Proportion of participants with NEDA-3 at Months 18, 24, 30, 36, and EOS

A Cochran–Mantel–Haenszel (CMH) test stratifying by age at screening (>40, ≤40 years) and region (US, non-US) will be used and the relative risk ratio on tolebrutinib compared to placebo with 95% CI will be estimated.

Count endpoints

Count endpoints include:

- Annualized adjudicated relapse rate
- Total number of new Gd-enhancing T1-hyperintense lesions

A negative binomial model will be used to analyze total count occurring during the observation period, adjusting for treatment group, age at screening (>40, ≤40 years), region (US, non-US), baseline EDSS score and baseline Gd-enhancing T1 lesion (present, absent). Log transformed post-randomization observation duration or number of scans will be the offset variable. The estimated mean rates/numbers for each treatment group with the corresponding 95% CI, and the relative reduction on tolebrutinib compared to placebo with the 95% CI will be estimated.

Continuous endpoints

Continuous endpoints include:

- Change from baseline in MSFC-3 Z-score at Months 12, 18, 24, and EOS
- Change from baseline in total volume of T2-hyperintense lesions at Months 18, 24, and EOS
- Change from baseline in EQ-5D-5L descriptive system index score and VAS score by visit

An MMRM method will be used for these endpoints, including covariates of treatment group, age at screening (>40, ≤40 years), region (US, non-US), visit, treatment-by-visit interaction, the baseline score, and baseline-by-visit interaction. Least squares means ± SE for each treatment group, as well as LSM difference with the 95% CI will be estimated.

Since volume data are often non-normally distributed, analysis of T2 lesion volume will be performed on log-transformed volume data. MSFC-3 and EQ-5D-5L data are assumed a priori to be normally distributed. Given the large sample size, the impact of violation of normality should not be great. In order to assess the impact of deviation from the normality assumption, a sensitivity analysis using rank ANCOVA, with the same covariates as for the MMRM analysis except VISIT and related interaction terms, will be performed to test the significance of the treatment effect.

Other endpoints

Data for magnetization transfer ratio (MTR) recovery at EOS in new MTR lesions detected at Months 6 and 12 and normalized T1 intensity in SELs will be listed.

Actigraphy sub-study data will be collected and processed by an external vendor, Koneksa. Analysis and reporting of this sub-study will be performed separately from the CSR.

Missing data will not be imputed in the analysis of tertiary endpoints, nor will sensitivity or subgroup analyses be conducted.

3.5 MULTIPLICITY ISSUES

To strongly control the Type 1 error rate for the study, a hierarchical testing procedure will be applied at a 2-sided 5% significance level, ie, each hypothesis will be formally tested only if the preceding one is significant at the 5% level. If statistical significance is achieved for the primary endpoint, a selective set of secondary endpoints will be tested following the hierarchical testing procedure shown below:

- Time to onset of 3-month CDP as assessed by the EDSS score
- Number of new and/or enlarging T2 hyperintense lesions per year
- Time to onset of sustained 20% increase in the 9-HPT for at least 3 months
- Time to onset of sustained 20% increase in the T25-FW for at least 3 months
- Time to onset of CDI confirmed over at least 6 months
- Percent change in brain volume as detected by MRI scans at the EOS compared to Month 6

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 descriptive with no testing planned. The summary of safety results will be presented by treatment group.
- Safety data in participants who do not belong to the safety population (eg, exposed but not randomized) will be provided separately.

3.6.1 Extent of exposure

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

Duration of IMP exposure

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

Duration of IMP exposure will be summarized quantitatively and categorically in the DB treatment period and OL period respectively: >0 to ≤24 weeks, >24 to ≤48 weeks, >48 to ≤72 weeks, >72 to ≤96 weeks, >96 weeks.

Additionally, the cumulative duration of treatment exposure (expressed in participant-years) will be provided.

Treatment compliance

A given administration will be considered noncompliant if the participant did not take the planned dose of treatment as required by the protocol, ie, once daily oral dose. No imputation will be made for participants with missing or incomplete data.

Percentage of compliance for a participant will be defined as the number of days that the participant was compliant divided by the total number of days of planned administration during the DB treatment period and OL period respectively.

Treatment compliance will be summarized quantitatively and categorically: <80%, ≥80%.

3.6.2 Adverse events

General common rules for adverse events

All adverse events (AEs) will be graded according to National Cancer Institute Common Terminology for Adverse Events (NCI-CTCAE version 5.0, 27 Nov 2017) and 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: AEs that developed, worsened, or became serious during the DB TE period
- OL AEs: AEs that developed, worsened, or become serious during the OL TE period
- Post-treatment AEs: AEs that developed, worsened, or became serious during the post-treatment period

Similarly, if occurring, deaths will be summarized in the pre-treatment, DB TE, OL, and post-treatment periods.

The primary focus of AE reporting will be on TEAEs. Pre-treatment, OL, 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, OL, 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 severity is missing for 1 of the TE occurrences of an AE, the severity will be imputed with the maximal severity of the other occurrences. If the severity is missing for all the occurrences, the severity will be left as missing.

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 tolebrutinib treatment group incidence, alphabetic order in case of equal frequency.

Analysis of all adverse events

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

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

Similar summaries for OL emergent AEs will also be provided. An additional overview summary including the number and rate of events will be provided in the DB treatment period and OL period respectively.

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 discontinuation and severe AEs will be provided.

Table 5 - Analyses of adverse events

Type of AE	MedDRA levels
All TEAE	Primary SOC and PT
	PT (incidence ≥2%)
Severe (grade ≥3) TEAEs	Primary SOC and PT
	PT
TEAE related to IMP as per Investigator's judgment	Primary SOC and PT
	PT
Treatment emergent SAE	Primary SOC and PT
	PT

Type of AE	MedDRA levels
Treatment emergent SAE related to IMP as per Investigator's judgment	Primary SOC and PT
Treatment emergent AESIs	AESI category and PT
TEAE leading to permanent study intervention discontinuation	Primary SOC and PT
	PT
TEAE leading to death (death as an outcome of the AE as reported by the Investigator in the AE page)	Primary SOC and PT
Pretreatment AE	Overview[a]
	Primary SOC and PT
Post-treatment AE	Overview[a]
	Primary SOC and PT

a Will include the following AE categories: any AEs, any severe (grade ≥3) AEs, any serious AEs, any AEs leading to death

Analysis of deaths

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 TE, OL, and post-treatment periods
- Deaths in non-randomized/non-enrolled participants or randomized but not treated participants

Analysis of adverse events of special interest (AESIs)

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.

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 related to IMP)
Increase in ALT >3 x ULN (confirmed)	Dedicated CRF page (AECAT="ALT INCREASE DATA") and AESI marked "Y"
ECG observation of atrial fibrillation or atrial flutter	CMQ30003 for selection and AESI check box marked
Severe infection (NCI CTCAE Grade 3 or above), that may or may not meet seriousness criteria (eg, a Grade 3 opportunistic infection)	SOC of Infections and Infestations, Grade 3 or above and AESI marked "Y"

AESIs	Selection
Moderate or severe hemorrhagic events (NCI CTCAE Grade 2 or above), including, but not limited to, symptomatic bleeding, bleeding in a critical area or organ such as the CNS, or intraocular bleeding	SMQNAME='Haemorrhage terms (excl laboratory terms)" for selection, Grade 2 or above, and AESI marked "Y"
Thrombocytopenia, platelet count <75 x 10^9/L	PT="Thrombocytopenia" 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 units, if applicable.

- Hematology:
  - Red blood cells and platelet: red blood cell count, platelet count, hemoglobin, hematocrit
  - White blood cells: white blood cell count, neutrophils, lymphocytes, monocytes, basophils, eosinophils
- Clinical chemistry:
  - Metabolism: glucose, total protein, albumin, creatine phosphokinase, lipase
  - Electrolytes: sodium, potassium, chloride, calcium, bicarbonate
  - Renal function: creatinine, creatinine clearance, blood urea nitrogen. (Creatinine clearance will be derived with the equation of Cockcroft and Gault using weight assessed at the same visit as creatinine.)
  - Liver function: alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase, total and direct bilirubin
- Urinalysis for quantitative analysis: specific gravity, pH
- Vital signs: heart rate, systolic and diastolic blood pressure, weight, temperature
- ECG variables: heart rate, PR, QRS, QT, and corrected QTcF (according to Fridericia). In case the ECG machine does not automatically calculate QTcF, manual calculation using nomogram or automatic website calculator is acceptable.

For laboratory data reported as <xx or >yy, a numeric value of xx/2 or yy, respectively, will be used for quantitative analysis.

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 analysis visit window during the DB treatment period and OL period respectively. These analyses will be performed using central measurements only for laboratory variables.

For each laboratory parameter, vital signs parameter and ECG parameter, mean changes from baseline with the corresponding SE 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.

Analyses according to PCSA (normal range) will be performed based on the worst value during the DB treatment-emergent period and OL 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 DB treatment-emergent period and OL period will be summarized respectively 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:

- Time to onset of the initial ALT or AST elevation (>3 x ULN) and total bilirubin elevation (>2 x ULN) during the treatment-emergent period will be analyzed using the KM method.
- The KM method will also be used to analyze time to onset of the following initial elevation ALT >3 x ULN, ALT >5 x ULN, ALT >8 x ULN, ALT >10 x ULN, AST >3 x ULN, AST >5 x ULN, AST >8 x ULN, AST >10 x ULN, total bilirubin >2 x ULN during the treatment-emergent period, respectively.
- 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 x 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 x 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 Columbia Suicide Severity Rating Scale (C-SSRS) during DB treatment period and OL period will be summarized respectively. A shift table for baseline versus during treatment and OL responses will be provided respectively according to the categories of no suicidal ideation or behavior, suicidal ideation, and suicidal behavior.

3.7 OTHER ANALYSES

3.7.1 Other variables and/or parameters

3.7.1.1 PK analyses

Tolebrutinib and M2 metabolite pharmacokinetic (PK) individual parameters (at least Cmax, tmax, AUC0-24) and corresponding descriptive statistics (mean, geometric mean, median, SD, coefficient of variation, minimum, and maximum) will be determined by population PK analysis by the PKDM department using sampling times at Months 6, 9, and 12.

Tolebrutinib and M2 metabolite PK parameters will be reported in a standalone report outside of the CSR.

3.7.2 Pharmacodynamic (PD) analyses

Blood samples for lymphocyte phenotyping will be collected for a subset of up to 200 randomized participants. Flow cytometry will be used for evaluation of change (and % change) from baseline in lymphocyte count and phenotyping subset. Additional pharmacodynamics evaluations include assessment of neurofilament light chain (NfL) levels in plasma, chitinase-3-like protein 1 (Chi3L1), and immunoglobulin (Ig) levels in serum.

These parameters will be summarized by actual treatment group on the safety population using the following descriptive statistics: mean, geometric mean, median, SD, coefficient of variation, minimum, and maximum.

3.8 INTERIM ANALYSES

A non-binding interim futility analysis will be performed when approximately 50% of the primary endpoint events are observed. The futility criterion will be determined by calculating the predictive power, which is a weighted average of the conditional power (11). Futility of the study may be declared only when the predictive power for the primary endpoint is below 10%, which is equivalent to the condition that the hazard ratio between tolebrutinib and placebo derived from the primary analysis is greater than or equal to 0.92. The futility interim analysis will be conducted by an independent statistics group and reviewed by the DMC. Neither the Sponsor's team nor the Investigator's staff will have access to the treatment information at the individual participant level or group level before the study is formally unblinded after study completion or after the DMC recommendation and Sponsor agreement for stopping the trial. The details of the interim analysis will be included in the Data Monitoring Committee (DMC) SAP.

4 SAMPLE SIZE DETERMINATION

Approximately 1700 participants will be screened to achieve up to 1290 randomly assigned to study intervention (2:1 randomization ratio of tolebrutinib to placebo). This study is planned as an event-driven trial based on 6-month CDP. The study will continue until approximately 288 events are projected to have occurred to provide 80% power to detect a 30% risk reduction in 6-month CDP with tolebrutinib compared to placebo (2-sided $\alpha = 0.05$). The following assumptions were used for the calculations: 2-year placebo event rate of 23.6%; annual discontinuation rate of 10%; constant hazard rates using a log-rank test; estimated enrollment period of 24 months with the last randomized participant followed for 24 months.

Actual recruitment and disability event rates may vary; therefore, it is possible to stop with a reduced sample size of around 1100 and still maintain study power by extending the duration of the trial, if needed, to reach approximately the same number of events. For example, it could be that the last randomized participant is followed for around 30 months instead of 24, if necessary. Additional recruitment of participants will not substantially reduce the trial duration. In any case, the power is based on the number of events, and the actual trial duration will be dependent on the observed placebo rate and treatment effect.

5 SUPPORTING DOCUMENTATION

5.1 APPENDIX 1 LIST OF ABBREVIATIONS

9-HPT:	9-hole peg test
AE:	adverse event
AESIs:	adverse events of special interest
ALT:	alanine aminotransferase
ANCOVA:	analysis of covariance
AST:	aspartate aminotransferase
ATC:	anatomic therapeutic category
BVL:	brain volume loss
CDI:	confirmed disability improvement
CDP:	confirmed disability progression
CI(s):	confidence intervals
CMH:	Cochran-Mantel-Haenszel
C-SSRS:	Columbia Suicide Severity Rating Scale
CVLT-II:	California Verbal Learning Test-II
DB:	double-blind
DMC:	Data Monitoring Committee
ECG:	electrocardiogram
e-CRF:	electronic case report form
EDSS:	expanded disability status scale
EOS:	end of study
Gd:	Gadolinium
GLMM:	generalized linear mixed model
HLGT:	high-level group term
HLT:	high-level term
IMP:	investigational medicinal product
KM:	Kaplan-Meier
LLT:	lower-level term
LSM:	least squares means
MedDRA:	Medical Dictionary for Regulatory Activities
MMRM:	mixed-effect model with repeated measures
MRI:	magnetic resonance imaging
MS:	multiple sclerosis
MSFC-3:	multiple sclerosis functional composite
MSQoL-54:	multiple sclerosis quality of life-54
MTR:	magnetization transfer ratio
NCI-CTCAE:	National Cancer Institute Common Terminology for Adverse Events
NEDA-3:	no evidence of disease activity-3
NRSPMS:	non-relapsing secondary progressive multiple sclerosis
OL:	open-label
PCSA:	potentially clinically significant abnormality

PD:	pharmacodynamic
PK:	pharmacokinetics
PT:	preferred term
ROW:	rest of the world
RRMS:	relapsing-remitting multiple sclerosis
SDMT:	symbol digit modalities test
SE:	standard error
SOC:	system organ class
SPMS:	secondary progressive multiple sclerosis
T25-FW:	timed 25-foot walk
TE:	treatment-emergent
TEAE:	treatment-emergent adverse event
VAS:	visual analogue scale
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 subjects 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 and not exposed, if applicable
- Randomized and exposed
  - Completed the DB study intervention period
  - Did not complete the DB study intervention period including main reason for discontinuation and reason for withdrawal by subject
  - Switched to the OL tolebrutinib treatment
  - Completed the OL period
  - Did not complete the OL period including main reason for discontinuation and reason for withdrawal by subject
- Completed the study
- Did not complete the study including main reason for study discontinuation
- Status at last contact (alive, dead)
- Entered LTS study

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 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 (DB and OL) and with early study discontinuation will be provided by country and site. Listings of other reasons for treatment 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. 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 demographics 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 quantitative variable and in categories (≤40, >40; 18 to 60)
- 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)
- Region (US, non-US; North America, Western Europe, Eastern Europe, ROW)
- 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)
- # Standard alcohol drinks on typical day when drinking (1 or 2, >2)

Disease characteristics at baseline

- Time since symptom onset of relapsing-remitting multiple sclerosis (RRMS) in years
- Time since diagnosis of secondary progressive multiple sclerosis (SPMS) in years
- Time since most recent relapse in years
- Baseline EDSS score as a quantitative variable (average of screening and baseline scores when both available)
- Disease burden at baseline (baseline EDSS ≤4.5 versus >4.5; EDSS ≤5.5 versus >5.5)
- Baseline Gd-enhancing T1 lesion (absent, present)

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

- Prior medications are those the participant received prior to first IMP intake. Prior medications can be discontinued before first administration or can be ongoing during treatment period.
- Concomitant medications are any medications received by the participant concomitantly to the IMP during the DB treatment and OL periods.
- Post-treatment medications are those the participant received in the period running from the end of the concomitant medications period up to the end of the study.
- 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 MS disease modifying therapy 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 medications will be summarized for the randomized and exposed population, by anatomic and therapeutic level. The summaries will be

sorted by decreasing frequency of anatomic category (ATC) based on incidence in the tolebrutinib 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 and similar quarterly windows will be applied if the treatment period extends beyond 48 months.

Table 7 - Analyses window definition of efficacy variables

AVISIT	Target Day	EDSS		MRI		T25FW, 9-HPT, SDMT, CVLT-II		MSQoL-54, EQ-5D-5L	
		Start Day	End Day	Start Day	End Day	Start Day	End Day	Start Day	End Day
Baseline	1	-40	1	-40	1	-40	1	-40	1
Month 3	84	2	126			2	126		
Month 6	168	127	210	2	252	127	210	2	252
Month 9	252	211	294			211	294		
Month 12	336	295	378	253	420	295	378	253	420
Month 15	420	379	462			379	462		
Month 18	504	463	546	421	588	463	546	421	588
Month 21	588	547	630			547	630		
Month 24	672	631	714	589	840	631	714	589	756
Month 27	756	715	798			715	798		
Month 30	840	799	882			799	882	757	924
Month 33	924	883	966			883	966		
Month 36	1008	967	1050	841	1176	967	1050	925	1092
Month 39	1092	1051	1134			1051	1134		
Month 42	1176	1135	1218			1135	1218	1093	1260
Month 45	1260	1219	1302			1219	1302		
Month 48	1344	1303	1386	1177	1512	1303	1386	1261	1428
EOS		Start Day: last visit window upper limit +1 End Day: EOS visit date Nominal visit (9000)							
FU for participants not entering LTS		≥Date of EOS+1 Nominal visit (8010)							
pEOT		Start Day: date of discontinuation+1 End Day: start of the next scheduled visit window Map to the closest scheduled visit							

Table 8 - Analysis window definition of safety variables

AVISIT	Target Day	Hematology, biochemistry except liver function tests		Vital Signs, C-SSRS, B-HCG Test		Weight		ECG		Urinalysis		Coagulation	
		Start Day	End Day	Start Day	End Day	Start Day	End Day	Start Day	End Day	Start Day	End Day	Start Day	End Day
Screening	-1	-40	-1	-40	-1	-40	-1	-40	-1	-40	-1	-40	-1
Day 1	1	1	1	1	1	1	1						
Month 1	28	2	42										
Month 2	56	43	70										
Month 3	84	71	98	2	126	2	126	1	126	1	210		
Month 4	112	99	126										
Month 5	140	127	154										
Month 6	168	155	210	127	210	127	252	127	210				
Month 9	252	211	294	211	294			211	294				
Month 12	336	295	378	295	378	253	420	295	504	211	420		
Month 15	420	379	462	379	462								
Month 18	504	463	546	463	546	421	588			421	588		
Month 21	588	547	630	547	630								
Month 24	672	631	714	631	714	589	756	505	840	589	756		
Month 27	756	715	798	715	798								
Month 30	840	799	882	799	882	757	924			757	924		
Month 33	924	883	966	883	966								
Month 36	1008	967	1050	967	1050	925	1092	841	1176	925	1092		
Month 39	1092	1051	1134	1051	1134								
Month 42	1176	1135	1218	1135	1218	1093	1260			1093	1260		
Month 45	1260	1219	1302	1219	1302								
Month 48	1344	1303	1386	1303	1386	1261	1386	1177	1386	1261	1386		
EOS		Start Day: last visit window upper limit +1; End Day: EOS visit date Nominal visit (9000)											
Liver Function test		Nominal visit at Month 0.5, 1.25, 1.5, 1.75, 2.25, 2.5, 2.75, 7, 8, 10, 11.											
FU for participants who do not enter LTS		>=Date of EOS+1 Nominal visit (8010)											
pEOT		Start Day: date of discontinuation+1; End Day: start of the next scheduled visit window Map to the closest scheduled visit											

Note: Analysis windows in open-label period can follow the same rule except that the first open-label dose date is Day 1 in that period

Table 9 - Analysis window definition of PD variables

AVISIT	Target Day	Plasma samples (NfL), serum samples (Chi3L1)		Serum samples (Ig levels)		Lymphocyte phenotypes by flow cytometry		Blood sample for archiving	
		Start Day	End Day	Start Day	End Day	Start Day	End Day	Start Day	End Day
Screening	-1							-40	-1
Day 1	1	1	1	1	1	1	1		
Month 1	28								
Month 2	56								
Month 3	84	2	126						
Month 4	112								
Month 5	140								
Month 6	168	127	252	2	252				
Month 9	252								
Month 12	336	253	504	253	504				
Month 15	420								
Month 18	504								
Month 21	588								
Month 24	672	505	840	505	840				
Month 27	756								
Month 30	840								
Month 33	924								
Month 36	1008	841	1176	841	1176				
Month 39	1092								
Month 42	1176								
Month 45	1260								
Month 48	1344	1177	1386	1177	1386				
EOS		Start Day: last visit window upper limit + 1; End Day: EOS visit date Nominal visit (9000)							
pEOT		Start Day: date of discontinuation + 1; End Day: start of the next scheduled visit window Map to the closest scheduled visit							

Unscheduled visits

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 PCSAs, and the shift summaries for safety. They will also be included in the by-visit summaries (central lab only for laboratory data and ECG) if they are re-allocated to scheduled visits based on the analysis windows defined above.

5.5 APPENDIX 5 ANALYSES SUPPORTING THE INTEGRATED SUMMARY OF SAFETY

Below are the additional analyses supporting the integrated summary of safety.

- Time to study intervention discontinuation for any reason (except switch to OL due to 6M-CDP) and due to AE using the KM method
- Exposure to study intervention by duration in participant years (≥24, ≥48, ≥72, ≥96, ≥120, ≥144, ≥168 weeks)
- Exposure to study intervention by subgroups in participant years
  - Age at screening (>40, ≤40 years)
  - Sex (Male or Female)
  - Age by sex (>40 & male, >40 & female, ≤40 & male, ≤40 & female)
  - Race (White, Black or African American, Asian, Other)
  - Weight in kg (<70, ≥70 to <90, ≥90)
  - BMI in kg/m^2 (<25, ≥25 to <30, ≥30)
  - Baseline EDSS (<4, ≥4)
  - Geographic region (US, non-US; Eastern Europe, Western Europe, North America, rest of the world [ROW])
  - Prior disease modifying therapy use (yes, no)
- Most frequent TEAEs sorted by relative risk and the corresponding forest plot
- Most frequent TEAEs sorted by risk difference and the corresponding forest plot
- Summary of TEAEs by primary SOC
- Summary of TEAEs by primary SOC, HLT and PT
- TEAEs leading to permanent study intervention discontinuation by exposure intervals (based on 24-week intervals)
- Incidence rate of post-treatment SAEs related to IMP
- Summary of TE AESIs by category, primary SOC and PT
- Summary of other TEAE groupings of interest by category, primary SOC and PT
- Incidence rate and relative risk ratio for TE AESIs and other selected AE groupings (COVID-19, suicidal behavior, malignancy) and the corresponding forest plot
- Incidence rate and risk difference for TE AESIs and other selected AE groupings (COVID-19, suicidal behavior, malignancy) and the corresponding forest plot
- Detailed summary for each AESI and other AE groupings
- Summaries of PTs occurring in at least 2/3/5% of participants on tolebrutinib and with a greater incidence compared to placebo by primary SOC and PT

- Boxplots and shift plots of the maximum values of ALT, AST, ALP and total bilirubin during treatment
- Incidence rates of TEAEs, TE SAEs, TEAEs leading to permanent study intervention discontinuation and AESIs/other AE group events by subgroups with corresponding forest plots

6 REFERENCES

1. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. J Am Stat Assoc.1989;84(408):1074-78.

2. Zhao Y, Herring AH, Zhou H, Ali MW, Koch GG. A multiple imputation method for sensitivity analyses of time-to-event data with possibly informative censoring. J Biopharm Stat. 2014;24(2):229-53.

3. Yuan YC. Multiple imputation for missing data: Concepts and new development. Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference; 2000.

4. Schwid SR, Goodman AD, McDermott MP, Bever CF, Cook SD. Quantitative functional measures in MS: what is a reliable change? Neurology. 2002;58(8):1294-96.

5. Benedict RH, DeLuca J, Phillips G, LaRocca N, Hudson LD, Rudick R, et al. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. Mult Scler. 2017;23(5):721-33.

6. Delis DC, Kramer JH, Kaplan E, Ober BA. (2000). CVLT-II, California verbal learning test: Adult version : manual.

7. Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A Health-Related Quality of Life Measure for Multiple Sclerosis. Quality Life Res. 1995;4(3):187-206.

8. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics. 1997;53:983-97.

9. Lu K, Mehrotra VD. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. Statistics in Medicine. 2010;29:474-88.

10. Van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15:708-15.

11. Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials Chapman & Hall / CRC. 1999.

Signature Page for VV-CLIN-0642962 v3.0
efc16645-16-1-9-sap

Approve & eSign	[redacted]

Approve & eSign	[redacted]